NASA-CR-162414
19800002000

The Deep Space Network Progress Report 42-53

July and August 1979

October 15, 1979

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The Deep Space Network Progress Report 42-53

July and August 1979

October 15, 1979

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 NASA Contract No. NAS7-100.

Preface

Beginning with Volume XX, the Deep Space Network Progress Report changed from the Technical Report 32- series to the Progress Report 42- series. The volume number continues the sequence of the preceding issues. Thus, Progress Report 42-20 is the twentieth volume of the Deep Space Network series, and is an uninterrupted follow-on to Technical Report 32-1526, Volume XIX.

This report presents DSN progress in flight project support, tracking and data acquisition (TDA) research and technology, network engineering, hardware and software implementation, and operations. Each issue presents material in some, but not all, of the following categories in the order indicated.

Description of the DSN

Mission Support
- Ongoing Planetary/Interplanetary Flight Projects
- Advanced Flight Projects

Radio Astronomy

Special Projects

Supporting Research and Technology
- Tracking and Ground-Based Navigation
- Communications–Spacecraft/Ground
- Station Control and Operations Technology
- Network Control and Data Processing

Network and Facility Engineering and Implementation
- Network
- Network Operations Control Center
- Ground Communications
- Deep Space Stations
- Quality Assurance

Operations
- Network Operations
- Network Operations Control Center
- Ground Communications
- Deep Space Stations

Program Planning
- TDA Planning

In each issue, the part entitled "Description of the DSN" describes the functions and facilities of the DSN and may report the current configuration of one of the seven DSN systems (Tracking, Telemetry, Command, Monitor & Control, Test & Training, Radio Science, and Very Long Baseline Interferometry).

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

Contents

DESCRIPTION OF THE DSN

MISSION SUPPORT

Ongoing Planetary/Interplanetary Flight Projects

RADIO ASTRONOMY

SUPPORT RESEARCH AND TECHNOLOGY

Tracking and Ground-Based Navigation

Station Control and Operations Technology

NETWORK FACILITY ENGINEERING AND IMPLEMENTATION

Network

Deep Space Stations

OPERATIONS

Network Operations

Deep Space Stations

Network Functions and Facilities

N. A. Renzetti
Office of Tracking and Data Acquisition

The objectives, functions, and organization of the Deep Space Network are summarized; deep space station, ground communication, and network operations control capabilities are described.

The Deep Space Network was established by the National Aeronautics and Space Administration (NASA) Office of Space Tracking and Data Systems and is under the system management and technical direction of the Jet Propulsion Laboratory (JPL). The network is designed for two-way communications with unmanned spacecraft traveling approximately 16,000 km (10,000 miles) from Earth to the farthest planets and to the edge of our solar system. It has provided tracking and data acquisition support for the following NASA deep space exploration projects: Ranger, Surveyor, Mariner Venus 1962, Mariner Mars 1964, Mariner Venus 1967, Mariner Mars 1969, Mariner Mars 1971, and Mariner Venus-Mercury 1973, for which JPL has been responsible for the project management, the development of the spacecraft, and the conduct of mission operations; Lunar Orbiter, for which the Langley Research Center carried out the project management, spacecraft development, and conduct of mission operations; Pioneer, for which Ames Research Center carried out the project management, spacecraft development, and conduct of mission operations; and Apollo, for which the Lyndon B. Johnson Space Center was the project center and the Deep Space Network supplemented the Manned Space Flight Network, which was managed by the Goddard Space Flight Center. The network is currently providing tracking and data acquisition support for Helios, a joint U.S./West German project; Viking, for which Langley Research Center provided the Lander spacecraft and project management until May, 1978, at which time project management and mission operations were transferred to JPL, and for which JPL provided the Orbiter spacecraft; Voyager, for which JPL provides project management, spacecraft development, and is conducting mission operations; and Pioneers, for which the Ames Research Center provides project management, spacecraft development, and conduct of mission operations. The network is adding new capability to meet the requirements of the Galileo mission to Jupiter, for which JPL is providing the Orbiter spacecraft, and the Ames Research Center the probe. In addition, JPL will carry out the project management and the conduct of mission operations.

The Deep Space Network (DSN) is one of two NASA networks. The other, the Spaceflight Tracking and Data Network (STDN), is under the system management and technical direction of the Goddard Space Flight Center (GSFC). Its function is to support manned and unmanned Earth-orbiting satellites. The Deep Space Network supports lunar, planetary, and interplanetary flight projects.

From its inception, NASA has had the objective of conducting scientific investigations throughout the solar system. It was recognized that in order to meet this objective, significant supporting research and advanced technology development must be conducted in order to provide deep space telecommunications for science data return in a cost effective

manner. Therefore, the Network is continually evolved to keep pace with the state of the art of telecommunications and data handling. It was also recognized early that close coordination would be needed between the requirements of the flight projects for data return and the capabilities needed in the Network. This close collaboration was effected by the appointment of a Tracking and Data Systems Manager as part of the flight project team from the initiation of the project to the end of the mission. By this process, requirements were identified early enough to provide funding and implementation in time for use by the flight project in its flight phase.

As of July 1972, NASA undertook a change in the interface between the Network and the flight projects. Prior to that time, since 1 January 1964, in addition to consisting of the Deep Space Stations and the Ground Communications Facility, the Network had also included the mission control and computing facilities and provided the equipment in the mission support areas for the conduct of mission operations. The latter facilities were housed in a building at JPL known as the Space Flight Operations Facility (SFOF). The interface change was to accommodate a hardware interface between the support of the network operations control functions and those of the mission control and computing functions. This resulted in the flight projects assuming the cognizance of the large general-purpose digital computers which were used for both network processing and mission data processing. They also assumed cognizance of all of the equipment in the flight operations facility for display and communications necessary for the conduct of mission operations. The Network then undertook the development of hardware and computer software necessary to do its network operations control and monitor functions in separate computers. A characteristic of the new interface is that the Network provides direct data flow to and from the stations; namely, metric data, science and engineering telemetry, and such network monitor data as are useful to the flight project. This is done via appropriate ground communication equipment to mission operations centers, wherever they may be.

The principal deliverables to the users of the Network are carried out by data system configurations as follows:

(1) The DSN Tracking System generates radio metric data, i.e., angles, one- and two-way doppler and range, and transmits raw data to Mission Control.

(2) The DSN Telemetry System receives, decodes, records, and retransmits engineering and scientific data generated in the spacecraft to Mission Control.

(3) The DSN Command System accepts spacecraft commands from Mission Control and transmits the commands via the Ground Communication Facility to a Deep Space Station. The commands are then radiated to the spacecraft in order to initiate spacecraft functions in flight.

(4) The DSN Radio Science System generates radio science data, i.e., the frequency and amplitude of spacecraft transmitted signals affected by passage through media such as the solar corona, planetary atmospheres, and plantetary rings, and transmits this data to Mission Control.

(5) The DSN Very Long Baseline Interferometry System generates time and frequency data to synchronize the clocks among the three Deep Space Communications complexes. It will generate universal time and polar motion and relative Deep Space Station locations as by-products of the primary data delivery function.

The data system configurations supporting testing, training, and network operations control functions are as follows:

(1) The DSN Monitor and Control System instruments, transmits, records, and displays those parameters of the DSN necessary to verify configuration and validate the Network. It provides the tools necessary for Network Operations personnel to control and monitor the Network and interface with flight project mission control personnel.

(2) The DSN Test and Training System generates and controls simulated data to support development, test, training and fault isolation within the DSN. It participates in mission simulation with flight projects.

The capabilities needed to carry out the above functions have evolved in the following technical areas:

(1) The Deep Space Stations, which are distributed around Earth and which, prior to 1964, formed part of the Deep Space Instrumentation Facility. The technology involved in equipping these stations is strongly related to the state of the art of telecommunications and flight-ground design considerations, and is almost completely multimission in character.

(2) The Ground Communications Facility provides the capability required for the transmission, reception, and monitoring of Earth-based, point-to-point communications between the stations and the Network Operations Control Center at JPL, Pasadena, and to the JPL Mission Operations Centers. Four communications disciplines are provided: teletype, voice, high-speed, and wideband. The Ground Communications Facility uses the capabilities provided by common carriers throughout the world, engineered into an integrated system by Goddard Space Flight Center, and controlled from the

communications Center located in the Space Flight Operations Facility (Building 230) at JPL.

The Network Operations Control Center is the functional entity for centralized operational control of the Network and interfaces with the users. It has two separable functional elements; namely, Network Operations Control and Network Data Processing. The functions of the Network Operations Control are:

(1) Control and coordination of Network support to meet commitments to Network users.

(2) Utilization of the Network data processing computing capability to generate all standards and limits required for Network operations.

(3) Utilization of Network data processing computing capability to analyze and validate the performance of all Network systems.

The personnel who carry out the above functions are located in the Space Flight Operations Facility, where mission operations functions are carried out by certain flight projects. Network personnel are directed by an Operations Control Chief. The functions of the Network Data Processing are:

(1) Processing of data used by Network Operations Control for control and analysis of the Network.

(2) Display in the Network Operations Control Area of data processed in the Network Data Processing Area.

(3) Interface with communications circuits for input to and output from the Network Data Processing Area.

(4) Data logging and production of the intermediate data records.

The personnel who carry out these functions are located approximately 200 meters from the Space Flight Operations Facility. The equipment consists of minicomputers for real-time data system monitoring, two XDS Sigma 5s, display, magnetic tape recorders, and appropriate interface equipment with the ground data communications.

Voyager Support

J. E. Allen and H. E. Nance
Deep Space Network Operation Section

This is a continuation of the Deep Space Network report on tracking and data acquisition for the Voyager project. This report covers the post-encounter period for Voyager 1 and the encounter period for Voyager 2, from April 1979 through July 1979.

I. Voyager Operation – Status

A. Voyager 1

Voyager 1 completed the post-encounter activities on 9 April 1979. During the remainder of the reporting period the activities were kept at a low level so that the majority of the support facilities could be committed to the preparation and support of the Voyager 2 encounter.

B. Voyager 2

Voyager 2 made its closest approach to Jupiter and its largest satellites on the afternoon of 9 July 1979. This event, like that of Voyager 1, was preceded by the observatory phase from 24 April through 23 May, and far-encounter phases from 23 May through 7 July 1979. A movie phase was completed during 27 to 29 May. After completion of the close-encounter phase the activities moved to the post-encounter activities which extend to 28 August 1979. Solar conjunction activities were observed by both spacecraft during the period 8 through 29 August 1979.

II. DSN Operations

A. Quick Turn Around

The period from January through August 1979, was the first time in the history of the DSN that it was required to support two projects, both in their prime mission phase, with the same priority on commitment of the tracking facilities. These two projects were Voyager and Pioneer Venus. The past philosophy of providing tracking coverage to an encountering spacecraft and exclude tracking of any other spacecraft for a period of time had to be abandoned and a new philosophy adapted. This new philosophy included the tracking of at least two projects by a DSN station and providing maximum tracking time per spacecraft. To accomplish these objectives, the "quick turnaround" was devised for the 64–m network along with a new plan for station maintenance.

The basic premise of the new philosophy was that the involved Network Operations Project Engineers (NOPEs) establish guidelines pertaining to the configurations and levels of support to be used during the various phases of their mission. Each level of support corresponds to mission phase criticality and the extent to which a station must be

calibrated/tested prior to tracking support. Four levels of support were defined:

(1) Critical redundant support

(2) Critical nonredundant support

(3) Normal support

(4) Minimum (load and go) support.

The time required to perform the pretrack preparation (PTP) depends on the configurations and the level of support required. The time includes time for station reconfiguration, but not time for normal housekeeping chores.

For the 64–m stations with S- and X-band downlinks, high rate telemetry streams and ranging for Voyager, the typical PTP times prior to quick turnaround averaged:

	Level 4	Level 3	Level 2	Level 1
VGR	1.5 h	3.5 h	6-8 h	8-10 h
PNV	1.0 h	3.0 h	6-8 h	8-10 h

NOTE: Radio Science Occultation Data Assembly (ODA) PTP time is not included in the Level 3 and 4 times.

Utilizing the normal PTP times as a starting point, the strategy utilized under the new philosophy essentially provides for: (1) performing the PTP at one time for both spacecraft; (2) reducing the time between tracks to a minimum; and, (3) performing posttrack activities for both spacecraft at one time.

The basis strategy under this new philosophy can then be broken down for the station considering two spacecraft support by specified time periods. (View periods for the 64–m network allow PNV tracking followed by Voyager tracking.)

A	B	C	D	E	F
0 . .	3 . . 6 . . 9	. .	12 . . 15 . . 18	. .	21 . . 24

A = PTP (TLM and CMD for both S/C) = 2 hours, 20 minutes.

B = Tracking PNV S/C ≅ 8 hours

C = Turnaround VGR S/C = 30 minutes

D = Tracking VGR S/C ≅ 8 hours

E = Posttrack activities for PNV and VGR ≅ 1 hour 30 minutes

F = Station maintenance ≅ 4 hours.

(Approximately 30 minutes are added to Period A when Radio Science equipment is required.)

Period A includes the following activities:

(1) RF calibrations on S- and X-bands

(2) Ranging calibrations for Voyager (when required)

(3) Telemetry and command strings for Voyager

(4) Telemetry and command strings for PNV

(5) Command data transfer for PNV.

Period C (quick turnaround) includes:

(1) Loss of signal on PNV

(2) Reconfiguring the front end equipment for Voyager

(3) Mounting of new tapes

(4) Resetting command suppression

(5) Command data transfer on String #2 while antenna moving in AZ

(6) Continue antenna to point for acquisition of signal (AOS) of Voyager

(7) After AOS, reinitialization of telemetry and command string for Voyager.

Period E (posttrack activities) requires:

(1) Ranging calibration for Voyager

(2) Recalls for PNV and Voyager

(3) Playback high rate telemetry for Voyager

(4) Providing data package for both projects.

After some "implementation pains" at the stations, the concept was implemented and successfully used for support during this critical period. Maximum tracking time was provided for each project so that prime requirements were fully satisfied. Additionally, this new philosophy increased the station's overall utilization percentage for flight projects.

B. Level 3 Implementation

The period from January through July 1979 included the encounter periods for both Voyager spacecraft. Concurrently, the implementation of new capabilities had to be considered. Background work which did not affect configuration requirements was authorized throughout the period; however, each item was very carefully scrutinized to assess the impact. The period between mid-March and the latter part of June was authorized for implementations that changed software and/or configurations by adding capabilities or correcting known

anomalies. These Voyager encounter tasks/Phase III were all completed on time and included:

(1) Metric data assembly software patch to correct halts and other anomalies

(2) Meteorological monitor assembly clock and sensor failure alarm corrections

(3) Occultation data assembly POCA truncation problem correction

(4) Installing a remote interim spectral signal indicator for the project and the DSN

(5) Installing a 168 kbit wide-band data line to DSS 43

(6) Installing a real-time signal combiner system at DSS 14 for antenna array between DSS 12 and 14.

(7) Completing a Radio Science open-loop recording wide-band data capability at DSS 14

(8) Providing tracking-radio science

(9) CPA-CMA noise elimination fix

(10) Providing a multi-mission receiver (MMR II B) at DSS 14.

As stated, these implementations were completed after the Voyager 1 closest encounter and were used to support the Voyager 2 closest encounter.

C. Real-Time Video

A request was made by DSS 14 to check the feasibility of providing the station with the processed video data from the real-time data stream during the encounter period. The station received comments from the Operations Chief of Trakon on the quality of the pictures as displayed in the Space Flight Operation Facility (SFOF), but could not make a correlation with only the operational indications of the Telemetry Processing Assembly. In addition, it was felt that it would be a morale factor for the station to observe a portion of the product of its efforts.

Several alternatives were investigated, including amateur radio facilities that produce video tapes. In the course of the process, Robot Corporation of San Diego volunteered to provide slow scan equipment, free of charge, for JPL and Goldstone with display capability at Goldstone. The interface was completed with the Mission Control and Computing Facilities at JPL and implementation of equipment at Goldstone prior to the closest approach of Voyager 2. The system worked very well and station personnel were able to view the encounter pictures along with the Network Operations Control Team in near-real-time. Presently, we are investigating the possibility of providing this capability permanently to Goldstone, as well as to other DSN facilities.

D. Backup to the Real-Time Monitors (RTM)

The number of Modcomp computers provided to the DSN for Real-Time Monitor (RTM) activities at JPL did not allow for backup or spare equipment. Since the test and training (T&T) RTM was not being utilized to any extent during the encounter period, it became a candidate to fill the void.

The different systems' software was successfully tried and tested in the T&T RTM. Likewise, provisions were made whereby the T&T RTM could be interfaced with the real-time systems as a substitute for another RTM.

The T&T RTM was declared capable, but not committed, as backup to any system RTM. This additional capability was maintained during the encounter period, but never required.

E. DSS 12/14 Arraying

The arraying of DSS 12 and DSS 14 antennas provide the real-time combiner with the dual input of the telemetry stream, thus improving the acquired composite signal, which required special procedures as well as equipment.

It was determined during preliminary testing that the desired results could only be obtained from an effective basic telemetry string, operating at maximum efficiency. Therefore, it was determined that the stations should count down both telemetry strings and determine which had the better performance. The high performance string was then used for the data which was, in turn, used as the input to the real-time combiner.

Likewise, DSS 14 was required to count down the real-time combiner along with other station equipment. It was determined that nominal time for station PTP was two hours for DSS 12 and three hours for DSS 14. This type of schedule was maintained for pre-closest encounter day, with the pretrack preparation time being doubled on closest encounter.

DSS 12 and 14 supported the encounter by arraying from 3 July through 9 July and again on 11 July. The average X-band telemetry signal gain from this operation for the period was 1.1 dB + 0.2 dB (which was the anticipated signal gain).

F. Io Torus

An additional radio science experiment was confirmed in May 1979 for the Voyager 2 near encounter. The experiment was to evaluate plasma density during occultation by the Io

torus. The experiment required simultaneous tracking of both Voyager spacecraft from 9 July through 11 July. The basic requirement for the experiment was:

(1) Voyager 1: Continuous track, acquire high-rate Doppler (S/X-band) and range data;

(2) Voyager 2: Continuous track, acquire high-rate Doppler (S/X-band) and open-loop receiver recording.

This experiment represented an added requirement at DSS 43 and 63 as DSS 14 already had occultation requirements that were not changed. To ensure this requirement could be satisfied, a series of tests were conducted during June 1979, to exercise the Radio Science subsystem at DSS 43 and DSS 63. Testing at DSS 14 had already been scheduled as part of the radio science implementation sequence.

Since the observations of the Io torus were characterized by narrow data spectra and small orbit uncertainties, it was advantageous to use very narrow open-loop filter bandwidths. The test plan utilized 1.0 kHz/3.0 kHz filter pairs for S/X-band. Recording periods were scheduled when the spacecraft was two-way noncoherent. Special radio science predicts for Voyager 2 were provided the 64–m stations. The Doppler sample rate was set to one per second for the first 15 minutes of the recording period. The SSI was used to monitor and ensure that signals were within loop receiver bandpass.

After four hours of recording, DSS 43/63 configured for ODA/ODR replay. The replay consisted of sampling the first 15 minutes of the first tape recorded: IDRs were made of the data and evaluated. The ODA/ODR tapes were shipped to JPL for final processing.

G. DSS 43 Multiple Wide-Band Data Line

During the Voyager 1 Jupiter encounter, a multiple wide-band data line (MWBDL), capable of transmitting data rates above 44.8 kbit in real-time was not available from DSS 43. The capability was available from DSS 14 and DSS 63. This configuration required that when DSS 43 was tracking and therefore receiving high-rate data, a special data replay strategy had to be implemented (Ref. 1).

A MWBDL capability was planned from Australia to Goddard for shuttle operation, so it was proposed to advance the date of implementation to cover the Voyager 2 Jupiter encounter beginning with the movie phase on 27 May 1979. The implementation was pushed and the capability made available for operations, on a best-effort basis, on that date for the movie. A considerable number of problems were experienced during the testing phase and a major effort was expended to solve the problems. The last real problems were experienced on 28 June. The period following through closest encounter was practically problem-free with very few error blocks being received.

A limitation of the system was the capability of Goddard and JPL to handle either, but not both, DSS 43 and DSS 63 data at the same time. Since the DSS 43 view period overlapped the DSS 63 view period, it became necessary to specify the time at which Goddard would switch from providing DSS 43 data to providing DSS 63 data. By advising the switch time beforehand and monitoring the switchover, the data interruption was only a matter of seconds.

This capability greatly reduced the data replay requirement for IDR production. Instead of having to recall all high-rate data (above 44.8 kbit/s) from DSS 43, it was only necessary to recall the portion of data lost due to the switchover interruption, assuming no other problems occurred during a pass. Approximately 851 IDRs were produced for the first 15 days of July, all being on time, with no backlog being experienced. For the entire encounter period only 12 minutes of data were actually lost (nonrecoverable from DODR) when the antenna at DSS 43 drove off point due to a hardware problem.

H. Radio Science

About 22 hours after the closest approach to Jupiter by Voyager 2, the spacecraft was occulted by the planet. This occultation lasted about 1.8 hours and data acquisition during this period was a key objective of the Radio Science experiment. A special sequence and acquisition of special data types required by the experiment involved activities at the three DSN stations at Goldstone, as well as special configuration of Voyager 2.

DSS 14 was the prime station for occultation data acquisition with DSS 11. Likewise, during Voyager 2 occultation tracking, DSS 12 tracked Voyager 1 to provide necessary calibrations of the Voyager 1 and 2 ray paths and to allow for solar plasma and the Earth's ionosphere. The data from DSS 12 was applied to the Voyager 2 data to enhance the overall data quality.

Special configuration was required at DSS 14 to support this activity. Real-time bandwidth reduction for the prime open-loop data was performed by the Occultation Data Assembly (ODA) using predicts generated and transmitted from JPL. The detection bandwidths were 5.0 kHz at S-band and 15.0 kHz at X-band and resulted in a recording rate of 320 kbits/s. Also, redundant S-band and X-band receivers were operated without bandwidth reduction to provide backup capability for the instrumentation supporting prime occultation data acquisition. The signals from the backup receiver was digitized in real-time and recorded on the digital recording

assembly (DRA) with effective filter bandwidths of about 300 kHz at S-band and 1.7 MHz at X-band. Nonreal-time processing of the DRA tapes can be performed at CTA-21, if required.

Closed-loop data acquisition was performed using the Block IV multimission receivers at DSS 14 and Block III receivers were used at DSS 11 and 12. High-rate Doppler was required at all three stations with the highest rate being 10 samples per second.

The maximum refractive loss expected as the Voyager 2 signal passed through the Jovian atmosphere was to be about 24 dB. However, it was anticipated that the closed loop receivers would remain in lock throughout the occultation period and that the SSI signal-to-noise ratio (SNR) would be adequate to provide good visibility of the open-loop received signal.

Figure 1 shows the actual and anticipated S-band signal profile.

(1) The refractive signal loss in dB is shown on the left side.

(2) The S-band downlink (ground received) signal is shown on the right side.

(3) The RCVR-3 SNR is given for the W-10 Hz bandwidth.

(4) The ODA SSI SNR is given for the 5 kHz bandwidth.

(5) The DRA SSI SNR is given for the 50 kHz bandwidth. If the 300 kHz bandwidth is used, the SNR will be about 8 dB less.

Figure 2 shows the actual and anticipated X-band signal profile.

(1) The refractive signal loss in dB is shown on the left side.

(2) The S-band downlink (ground received) signal is shown on the right side.

(3) The RCVR-4 SNR is given for the W-30 Hz bandwidth.

(4) The ODA SSI SNR is given for the 20 kHz bandwidth.

(5) The DRA SSI SNR is given for the 50 kHz bandwidth.

The figures show the correct ODA SSI and DRA SSI SNRs for the DSS 14 (800 line) display. The displays at JPL are 400 line; therefore, the SNR will be 3 dB less than shown in the figures. Also, establishing the true noise floor can easily be another 3 dB error source at JPL.

During the geometric Earth occultation it was found that the S/X-band signal profile dropped about 14 dB lower than anticipated and that the closed loop receiver lock was lost during the occult period. The best estimate of the signal during this period is −170 dBm for X-band and −178 dBm for S-band. X-band was out of lock approximately 50 percent of the time due to signal fading and S-band approximately 90 percent of the time due to threshold condition. Figures 1 and 2 reflect the estimated actual signal profile during the occultation period.

A tracking system graphic display capability was provided for the monitoring of pseudo-residual and tracking system noise during the occultation period. Terminal equipment was installed in the Network Analysis Team (NAT) tracking area and in the project Radio Science area. The primary display was a volatile digital TV, with a hard copy capability in NAT track. The system was controlled by NAT track in the fulfillment of their primary task; however, project requirements were accepted and honored whenever possible. Coordination was maintained between project and NAT track over the normal voice circuit.

Likewise, remote Spectral Signal Indicator (SSI) displays were slaved to the DSS 14 open-loop receivers and installed in the NAT track area and in the project Radio Science area. Using this display, both project and NAT could monitor the station activity during occultation and correlate with other data available.

Reference

1. For detail see J. E. Allen and H. E. Nance, "Voyager Support," *The Deep Space Network Progress Report 42-51*, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, March and April, 1979.

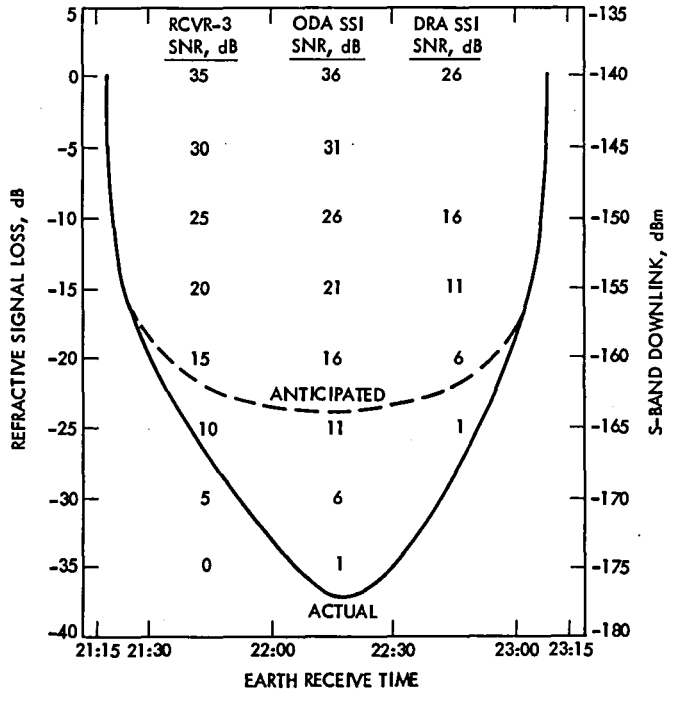


Fig. 1. Actual and anticipated S-band signal profile during Jupiter occultation (July 10, 1979)

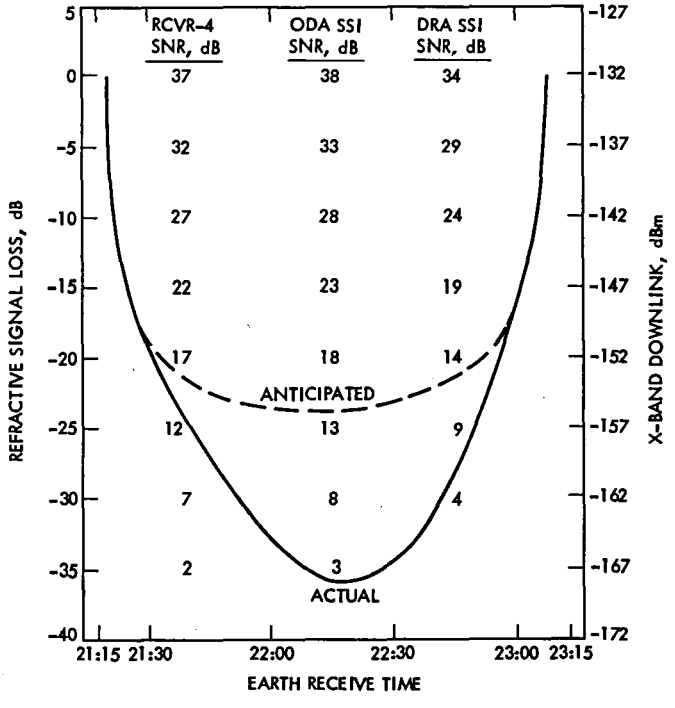


Fig. 2. Actual and anticipated X-band signal profile during Jupiter occultation (July 10, 1979)

Pioneer 11 Saturn Encounter Support

G. M. Rockwell
Deep Space Network Operations Section

This article reports on activities of DSN Operations in support of Saturn Encounter Operations of Pioneer 11.

I. Introduction

This article is mainly concerned with the preparation of the Deep Space Network for support of Pioneer 11 Saturn Encounter Operations. Points that will be discussed include: the encounter's main objectives, DSN testing, special DSN configurations and procedures, Radio Science activities, and any other mission-related activities.

II. Mission Operations and Status

The Pioneer 11 spacecraft was launched on April 6, 1973 and began its long cruise to Jupiter, which it encountered on December 3, 1974. By using Jupiter's strong gravitational field, the Pioneer 11 spacecraft's trajectory was redirected for its flight to Saturn, which it will encounter on September 1, 1979. The total cruise time between Jupiter and Saturn will total 1732 days. Throughout this cruise phase, the spacecraft has remained in excellent condition with very few system malfunctions. During this cruise period, the DSN has routinely tracked Pioneer 11 under the direction of Pioneer 11 operations at the Ames Research Center (ARC) located at Moffett Field, California.

III. Encounter Description and Objectives

As stated above, Pioneer 11's closest approach to Saturn will occur on September 1, 1979. Earth-received time for periapsis will be about 1756 universal time coordinated (UTC), with a round-trip light time of 2 hours and 52 minutes. The actual encounter period began on August 2, 1979 (encounter minus 30 days), and will end on October 1, 1979 (encounter plus 30 days). The spacecraft will pass within 81,426 km (1.36 Rs) of Saturn at periapsis. Inclination of the trajectory will be 6.5 degrees to the Saturn equator, with periapsis occurring in the southern hemisphere. Earth and solar occultations occur shortly after periapsis.

Besides the observation of Saturn itself, the other major objectives of this encounter will be to get a better observation and understanding of its rings and a close observation of Saturn's largest moon Titan. The spacecraft will pass within an estimated 3400 km under the edge of the visible rings, and its closest approach to Titan will be approximately 356,000 km.

During this encounter sequence, an estimated 25,000 real-time commands will be required, most of which are required to operate the Imaging Photo Polarimeter (IPP). The IPP imagery will exceed Earth-based resolution at encounter minus 6 days. Figure 1 shows the Saturn flyby trajectory as viewed from Earth, and Figure 2 shows the heliocentric geometry of the Pioneer 10 and 11 trajectories and positions on September 1, 1979. Table 1 also lists the closest approach events at Saturn and its satellites. The Project is expecting 50 pictures of better-than-Earth-based quality during encounter. Twenty will be taken prior to periapsis and thirty following

encounter. For a further description of the Saturn encounter, see Ref. 1.

Saturn-Earth occultation will occur shortly after encounter and last for about 78 minutes. Data taking for this event and the Occultation Experiment will begin one hour prior to entry and continue until one hour following exit. The entry phase will be supported by both DSS 63 and DSS 14. The exit will be supported by DSS 14 only, since DSS 63 will have set by that time. The spacecraft will be in the two-way mode at entry with DSS 62 providing the uplink. Upon exit, DSS 12 will be two-way noncoherent with the spacecraft and providing the uplink. The uplink sweep will be timed to start a one-way light trip time prior to spacecraft exit.

Superior Conjunction occurs on September 11, 1979. The Sun-Earth-Probe (SEP) angle at encounter will be 8 degrees. This angle will decrease to 1.75 degrees on September 11. The data rate will be decreased as the SEP angles decrease below 3 degrees. The period of angles less than 3 degrees is expected to be September 8 through September 14. Solar flare activity could alter these data degradation dates. Figure 3 shows the Sun-Earth-Saturn angle, which is nearly equal to SEP, in degrees versus the date.

Like Pioneer 10, Pioneer 11 will escape the solar system on its postencounter trajectory. Pioneer 11 will travel in the opposite direction of its sister ship. The escape asymptote is in the general direction of the solar apex, which is in the opposite direction from the Sun's motion through the interstellar medium. The spacecraft may be tracked through 1987.

IV. Special DSN Configurations and Procedures Required for Encounter Support

The 64-meter subnet, consisting of DSS 14 (California), DSS 43 (Australia), and DSS 63 (Spain), will be supporting the encounter phase. During the period of encounter ± 7 days, the 26/34 meter subnet will provide dual-station coverage. During this period, the 64-meter stations will be in a listen-only mode, and the 26/34-meter stations will provide the uplink for commanding the spacecraft. An improvement of 0.8 dB in the downlink is expected in the listen-only mode. Figure 4 shows the 64-meter ground station configuration, and Fig. 5 shows the 26/34-meter configuration.

The configuration for support of occultation by DSS 14 and DSS 63, as described in paragraph III of this article, is shown in Fig. 6.

A. Antenna Arraying

During the encounter ± 7-day period, a Research and Development antenna array technique will be used at Goldstone to gain additional telecommunications performance on a best efforts basis. Arraying is expected to provide an expected SNR improvement above elevations of 15 degrees of 0.4 to 0.5 dB. The signals from DSS 12 and DSS 14 are combined using a Real-Time Combiner (RTC) at DSS 14. This configuration is shown in Fig. 7 at a gross level and in more detail in Fig. 8.

B. Contingency Manual Commanding

In the event of a failure in the ARC Command System or communications between Ames and the DSN that would prevent the transmission of commands in the automatic mode, the Project may require certain safing commands to be transmitted in the manual mode.

Due to the complexity and length of these safing command sequences, a special command procedure has been developed. For proper spacecraft response to the command sequence, the commands must be transmitted contiguously.

C. Exciter and Receiver Ramping

The closest approach will be less than two Saturn radii from the center of the planet and the doppler rates are expected to be extremely high, nearly 80 Hz per second (S-band), prior to entering occultation. This coupled with low signal levels will present a less than optimum tracking situation for ground receivers. In addition, the spacecraft receiver will see more frequency change than it can track.

To help alleviate the problem, the use of exciter and receiver linear ramps using Programmed Oscillator Control Assemblies (POCAs) at both prime encounter complexes will be utilized. Therefore exciter POCAs have been installed at DSS 62, and exciter plus receiver POCAs installed at DSS 12. The prime 64-meter stations (DSS 14 and DSS 63) already had POCAs.

A ramping strategy is being developed and will be used during the encounter tests described later in this article.

V. DSN and Project Preencounter Test and Training

To perform at an optimum level during Saturn encounter, the DSN and the Pioneer Project have pursued an extensive training program prior to encounter to perfect the use and understanding of the procedures and configurations described above. Each phase of this program is discussed below, including all available results.

A. Contingency Manual Commanding Operation Verification Tests (OVTs)

A total of 48 OVTs were conducted with the Deep Space Stations, which began July 2, 1979. The stations tested were the 26/34-meter subnet consisting of DSS 12 (California), DSS 44 (Australia), DSS 62 (Spain) and the 64-meter subnet made up of DSS 14 (California), DSS 43 (Australia), DSS 63 (Spain). Each operational crew was tested twice amounting to eight OVTs per facility. The Pioneer Project at ARC also used these tests to train their operators in the use of the new command procedures. The overall results of this testing effort was very successful and all personnel participating gained an excellent understanding of the procedures and their utilization.

B. Saturn Encounter OVTs

Prior to encounter, the DSN will conduct OVTs simulating actual events that will occur during the encounter period. There will be one OVT conducted with DSS 12, DSS 14, DSS 62, and DSS 63.

The objective of each OVT may vary somewhat, as follows: the DSS 12 test will demonstrate receiver ramping and post occultation uplink reacquisition; at DSS 62, the test will demonstrate uplink exciter ramping capability; the tests at DSS 14 and DSS 63 will demonstrate the operations to be performed at each station during their encounter passes. These will include the use of receiver ramps, utilization of the Occultation Data Assembly (ODA) and the Digital Recording Assembly (DRA), the use of the Pre-Post Recording (PPR) system, and the processing of 1024 bits/s coded data.

At this writing successful encounter tests have been conducted at DSS 62 on August 15, 1979, and at DSS 63 on August 16, 1979. The results from tests at DSS 12 and DSS 14 are not available at this time.

C. Antenna-Array Testing at DSS 12 and DSS 14

To accurately determine whether or not the antenna-array technique will prove advantageous during encounter, testing will be conducted using live Pioneer 11 1024 bits/s coded data. Therefore, testing will be accomplished by utilizing demonstration tracks over DSS 12 and DSS 14 prior to encounter. The stations will utilize the configurations depicted in Figs. 5 and 6 during these training passes. The main objectives of these passes will be to verify precalibration procedures, measure the performance of the RTC, validate RTC operational procedures, measure DSS 12's and 14's operational readiness, verify the performance of the DSS 12/14 microwave communications link, and verify the overall antenna-array configuration.

As in the case of some encounter OVTs mentioned above, these tests have yet to be completed at this writing, so no results are available.

VI. Pioneer 11 Radio Science Activities

During the encounter period, three radio science experiments will be supported by the DSN.

A. Ring Plane Crossing and Occultation

On September 1, 1979 (encounter day), DSS 14 and DSS 63 will support an experiment involving the ring plane crossing and occultation. The ODA will be used to take data during this exercise, with the DRA being used for backup data collection. Due to encounter uncertainties, a wide open-loop receiver bandwidth of 8 kHz will be used.

B. Solar Corona Experiment

Beginning on September 2, 1979, ODA data will be taken continuously until September 24, 1979, in support of the Solar Corona Experiment. The narrow open-loop receiver 100-Hz bandwidth will be used for this support with a quantization of 8 bits.

C. High Latitude Solar Wind Experiment

This experiment, known as the Grand Parade, will begin on September 6, 1979. The 64-meter subnet will again utilize the ODA in support of this experiment. The open-loop receiver bandwidth will be 100 Hz and the quantization will be 12 bits.

VII. Summary

In this article we have discussed the Pioneer 11 mission, the Saturn encounter plan, a preencounter preparation plan, and the special activities that will be associated with the encounter. Because of the effort taken prior to Pioneer 11's arrival at Saturn, all participants are highly confident of a successful outcome.

Reference

1. Miller, R. B., "Pioneer 11 Saturn Encounter Mission Support," in *The Deep Space Network Progress Report 42-52*. Jet Propulsion Laboratory, Pasadena, Calif., August 15, 1979, pp. 4–7.

Table 1. Pioneer 11 closest approach events at Saturn and Saturn satellites

Satellite	Closest approach, km	Encounter time, hours: minutes	Ground received, DOY/UTC
Phoebe	9,453,000	- 128:32	239/0928
Iapetus	1,039,000	- 82:29	241/0731
Hyperion	674,000	- 28:03	243/1357
Dione	291,100	- 00:30	244/1730
Mimas	103,400	- 00:07	244/1753
Saturn	81,426	00:00	244/1800
Tethys	331,700	+ 01:54	244/1954
Enceladus	225,200	+ 01:59	244/1959
Rhea	341,900	+ 06:00	245/0000
Titan	356,000	+ 25:31	245/1931

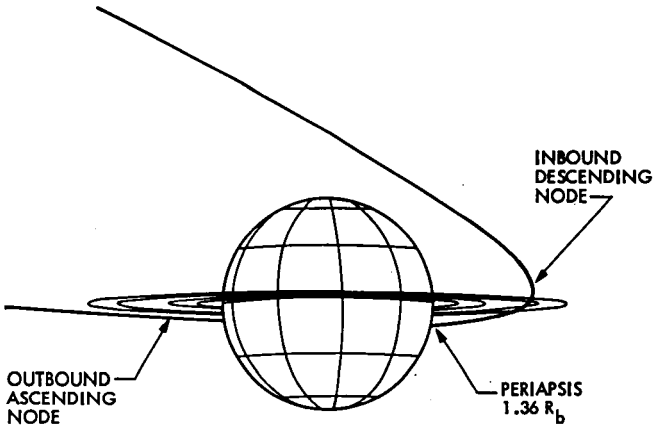


Fig. 1. Earth view of Pioneer 11 Saturn flyby – balanced ring plane crossings at 2.87 R_s

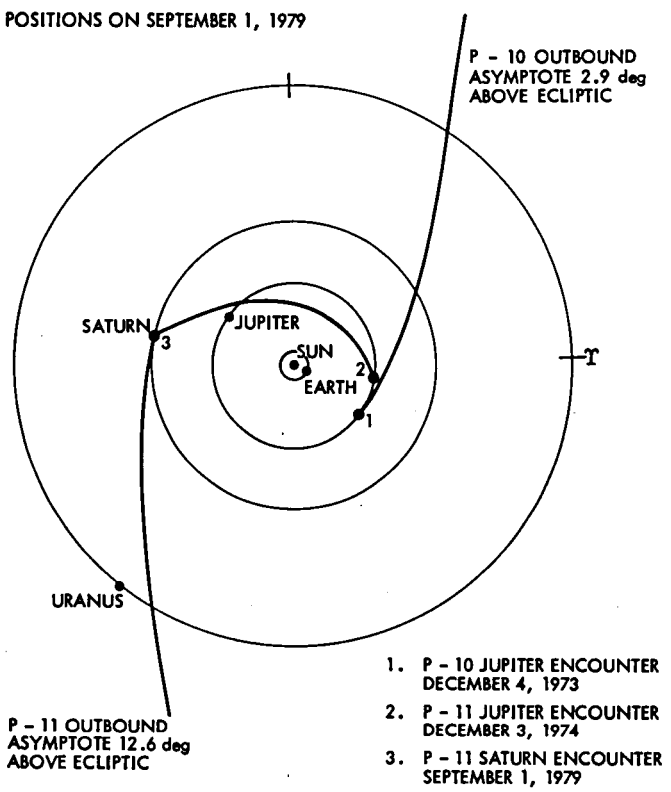


Fig. 2. Heliocentric geometry of Pioneer 10 and 11 trajectories

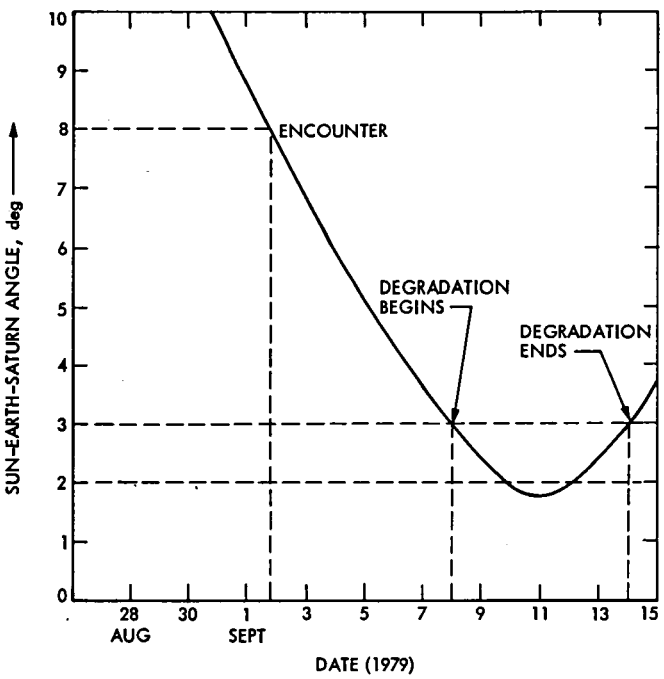


Fig. 3. Angular separation of Sun from Saturn for Pioneer 11 encounter at Saturn

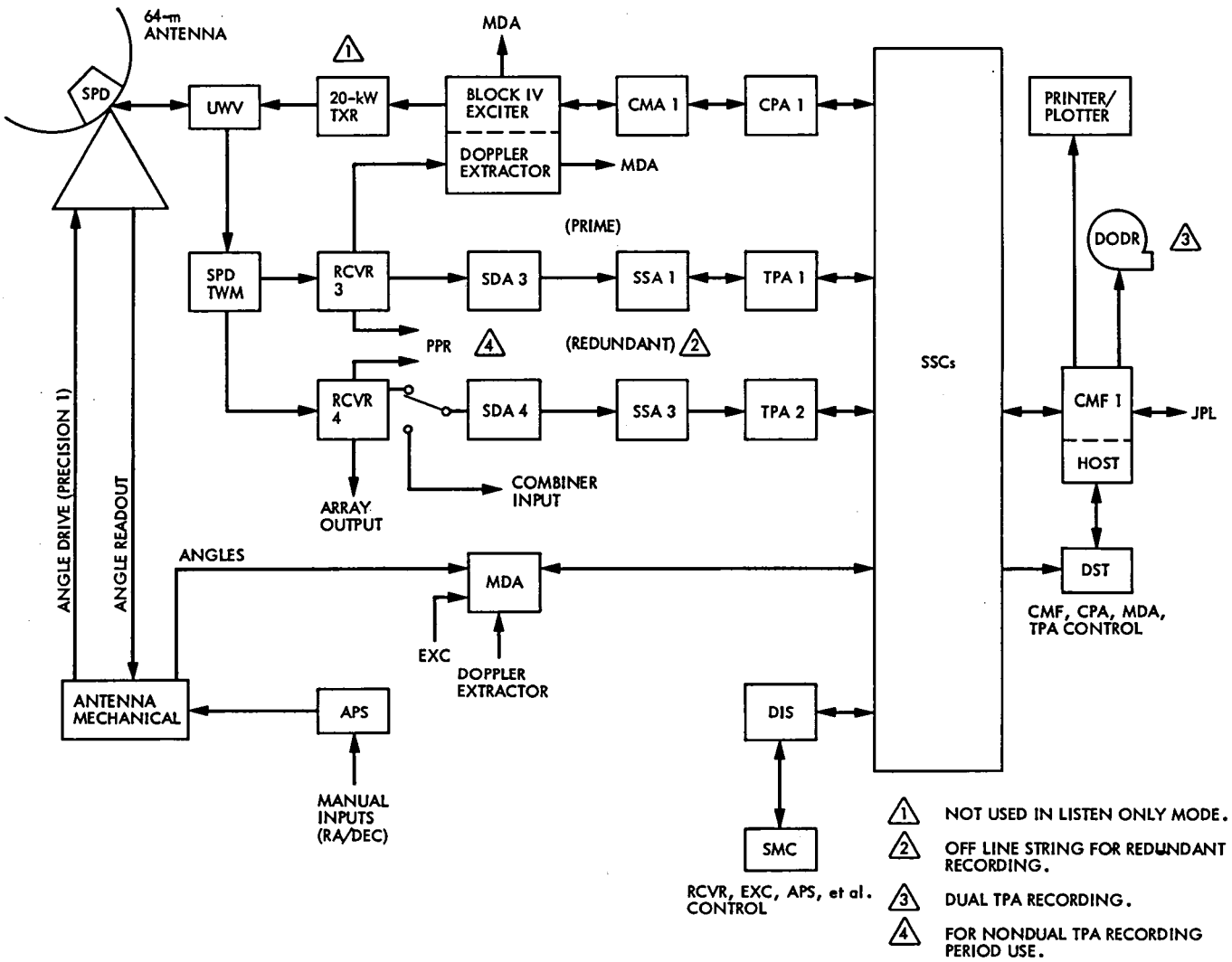
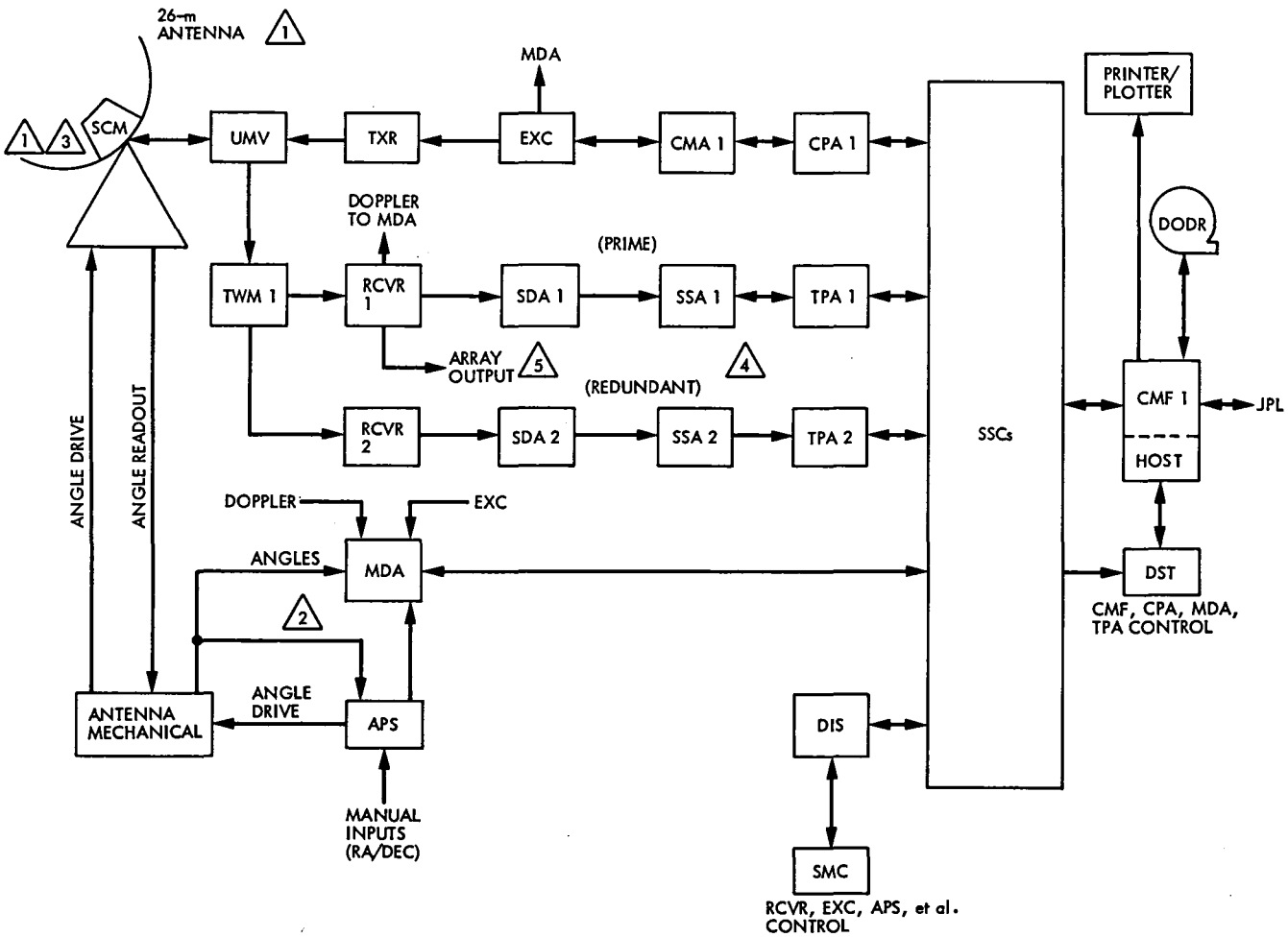


Fig. 4. 64-meter Pioneer encounter configuration



1 34-m AT DSS 12.

2 DSS 44 ONLY. BINARY ANGLE READOUTS CONVERTED TO BCD IN APS, THEN ROUTED TO MDA.

3 SXD AT DSS 12.

4 FOR USE IN CASE OF PRIME STRING FAILURE.

5 DSS 12 ONLY.

6 PRIME TLM STRING WILL BE DETERMINED BY RCVR CONTAINING 3 Hz LOOP FILTER.

Fig. 5. 26/34-meter Pioneer encounter configuration – DSSs 11, 12, 44, and 62

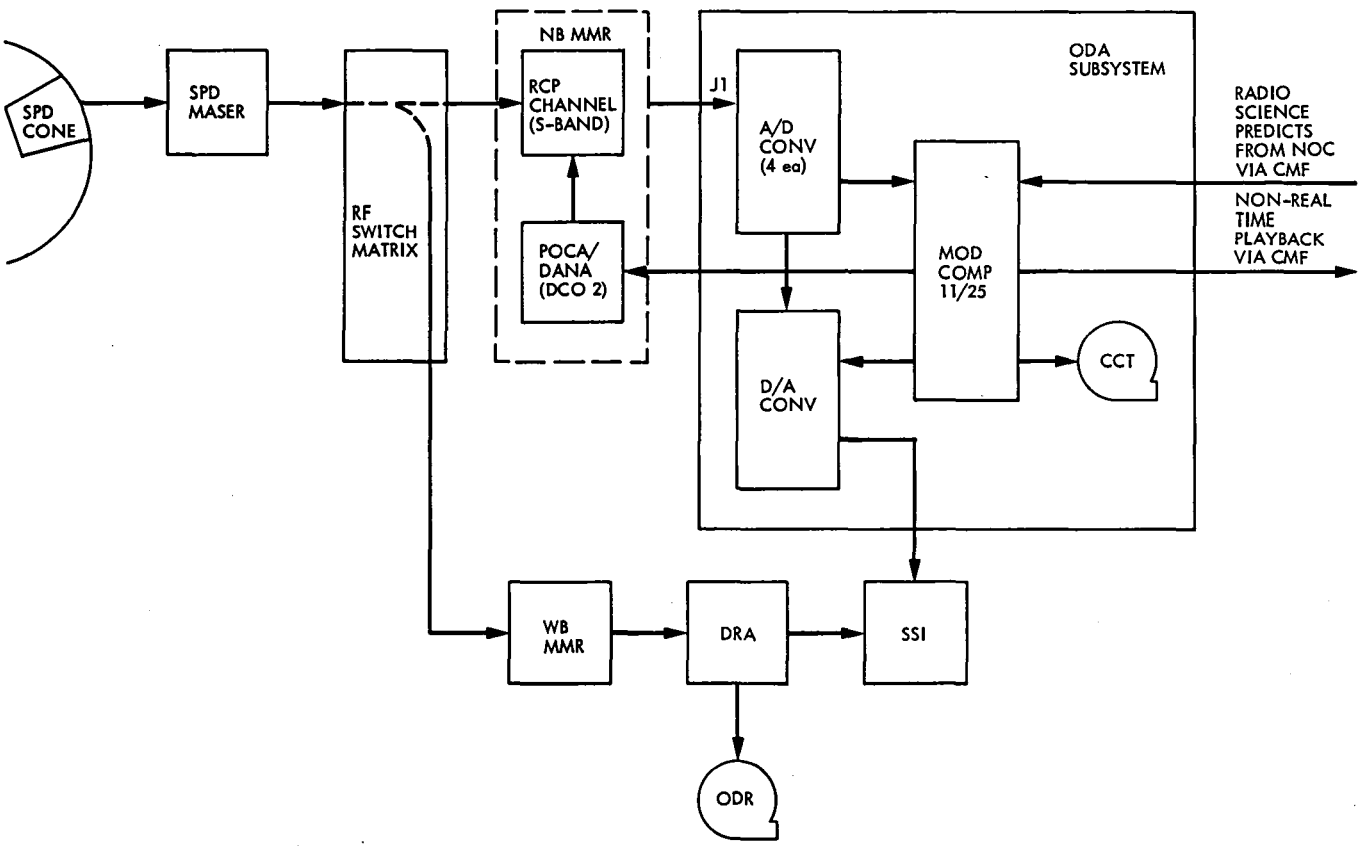


Fig. 6. Occultation experiments (S-band) configuration

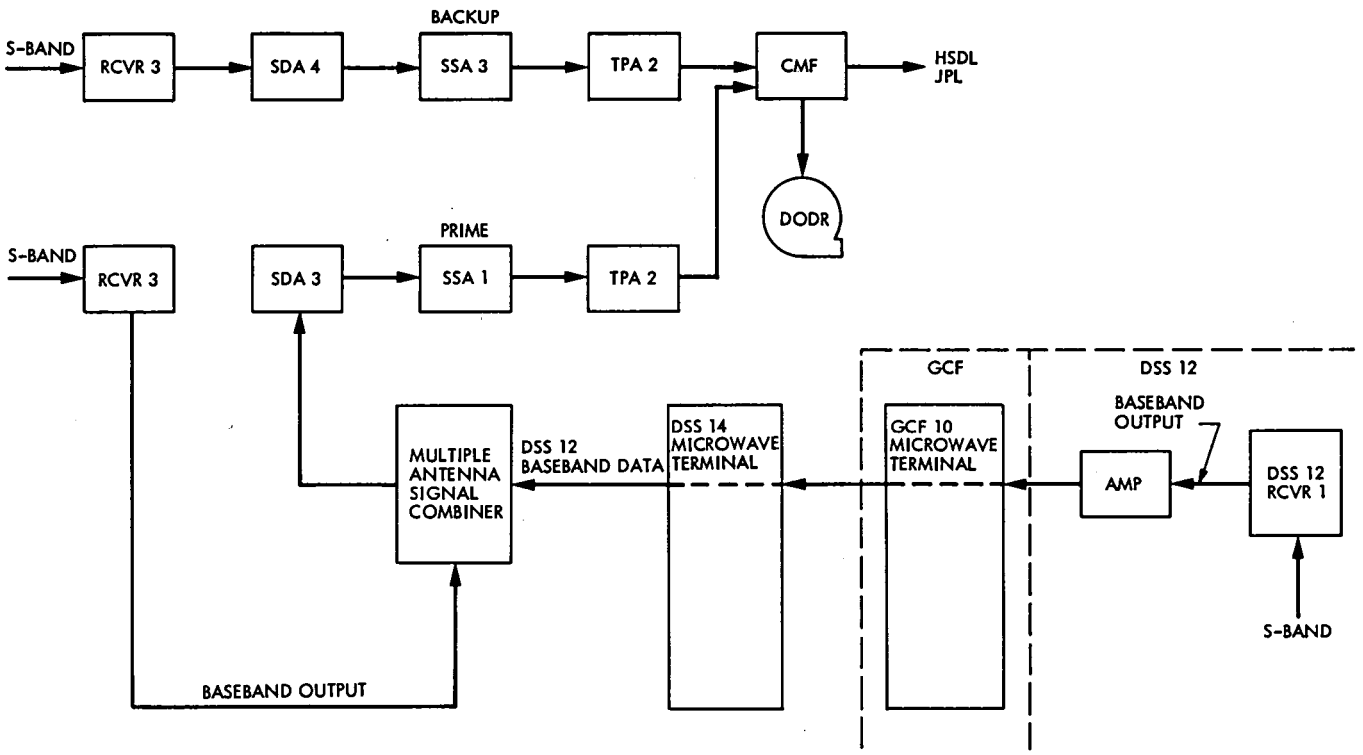


Fig. 7. DSSs 12/14 antenna array configuration

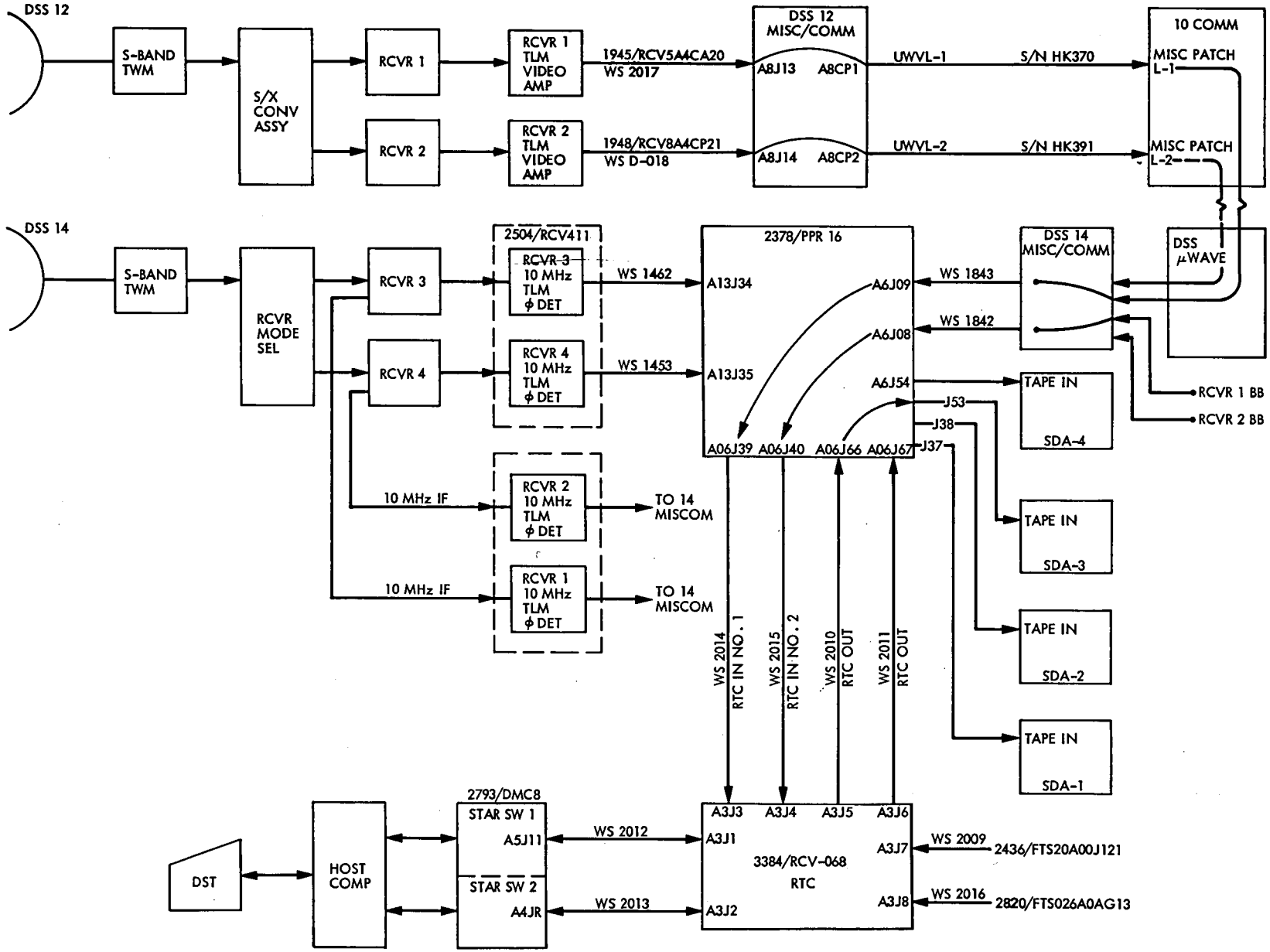


Fig. 8. Real-time combiner interface/configuration

Radio Astronomy

R. J. Fahnestock
TDA Mission Support Office

R. M. Taylor
Deep Space Network Operations Section

Since 1967 radio astronomers have used the Deep Space Network (DSN) 26- and 64-meter diameter antenna stations to investigate the Moon, planets, pulsars, and radio galaxies to study the effects of solar corona on radio signals, and to observe radio emissions from X-ray sources. Very Long Baseline Interferometry (VLBI) techniques have been used for high-resolution studies of quasars, suspected black holes, and other radio sources. The high-power transmitter capability of the Goldstone 64-meter antenna has been used for radar-ranging to the planets in support of various planetary space missions and for the study of comets and asteroids. The overseas facilities in Spain and Australia have been used, on a noninterference basis with space mission operations, by host country radio astronomers. The background for the DSN support, past performance, and current plans for continued support of these activities are discussed. Various experiments of the past year are identified and summarized.

I. Introduction

The first formal recognition and joint endorsement by JPL and NASA for use of DSN facilities to support radio astronomy experiments occurred in 1967. Two years later, as more and more requests were being received, the Radio Astronomy Experiment Selection (RAES) Panel was formed by the JPL Director to evaluate and select the most appropriate and worthy non-NASA proposals to compete for the small amount of station time available for this purpose. In 1971, a NASA Management Instruction (NMI 7100.6) was issued which set forth the policy and responsibilities for "Ground-Based Radio Science." Simultaneously, NMI 1152.44 established a NASA Ground-Based Radio Science Panel "as a standing committee of NASA to provide assistance and advice to the Associate Administrator for Space Science and Applications and the Associate Administrator for Tracking and Data Acquisition on matters pertaining to ground-based radio science." Presently, these are the Office of Space Science (OSS) and the Office of Space Tracking and Data Systems (OSTDS). This panel disbanded in 1973 when it was not extended beyond the designated two-year period.

The latter NMI specified details of a long-range program and annual plan for NASA support of ground-based radio science to be provided to the Associate Administrators of OSS and OSTDS. Annually, JPL submits a fiscal plan (Ref. 1) which serves as a mini-Support Instrumentation Requirements Document for DSN support of radio astronomy.

To minimize confusion between the terms Radio Science and Radio Astronomy as applied to DSN activities, in 1977, the DSN adopted the following definitions:

(1) Radio Science pertains to the acquisition and extraction of information from spacecraft-transmitted signals which have been affected by celestial bodies or have interacted with the propagation media.

(2) Radio Astronomy pertains to the acquisition and extraction of information from signals emitted or reflected by natural sources (i.e., all sources other than spacecraft).

II. Radio Astronomy Operations

During the past year the 26- and 64-meter antenna stations have provided a total of 1972 hours of operation in support of experiments in four different categories: NASA/OSS, NASA/OSTDS, RAES, and host country-sponsored radio astronomy. Also, the Tidbinbilla stations, DSS 42 and DSS 43, provided the support necessary for the development of a Real-Time Short Baseline Interferometer sponsored by the JPL Director's Discretionary Fund.

A. NASA OSS Category

The majority of Network time, 1033 hours, has been given to this category of activity in the past year, with 784 hours being used at DSS 62 in support of the Jupiter Short-Term Variation Experiment. This is a joint investigation involving Sam Gulkis and Mike Klein at JPL, Jose Antonio Turegano at the University of Zaragoza and Eric Gerard of the Nancay Observatory in France. At this time, the data obtained are being processed.

Of the remaining time, Pulsar Rotation Constancy experiments utilized 243 hours. New timing and recording software, developed by Miguel Urech at DSS 62, is presently being tested at Goldstone and should make better and more efficient use of Network facilities when it becomes operational.

Little demand has been made by Martin Slade, the principal investigator of the ALSEP/Quasar VLBI experiment in deference to heavy Flight Project activity, but it is anticipated that work in this area will increase considerably in the near future.

Activity in support of the Southern Hemisphere Radio Source Position investigation is expected to increase in the near future meanwhile the Moon Mapping and Lunation Curve experiment is drawing to a close and publications have been made (Refs. 2, 3 and 4).

B. NASA OSTDS Category

Some 478 hours of Deep Space Network tracking time has been used in this category in support of Planetary Radar Astronomy. The majority of the time, 441 hours, was utilized to map the planet Venus while 37 hours was given over to Saturn rings mapping.

C. Radio Astronomy Experiment Selection (RAES) Category

At this time there are some ten experiments on the RAES panel-approved activities list. All of these experiments request the use of 64-meter antenna time plus the Hydrogen Maser Frequency and Timing Subsystem (Ref. 1). There is one exception to this in that RA137 (Quasar Patrol) will accept 26-meter antenna time when 64-meter antenna time is not available.

Due to the dearth of available 64-meter antenna time, the RAES panel has prioritized these experiments as follows:

RAES Panel Number	Subject	Panel Priority
RA 137	Quasar Patrol	4
RA 164	Compact Radio Source Observation	7
RA 165	Compact Nonthermal Sources	4
RA 168	Galactic Center VLBI	3
RA 169	Compact Nuclei VLBI	1
RA 170	Radio Emisson VLBI	8
RA 171	M87 Interstellar Scintillation VLBI	1
RA 172	Recombination Lines, Nearby Galaxies	10
RA 173	Hotspots in Extra-Galactic Sources	8
RA 174	M82 Compact Source VLBI	4

Only three experiments, RA 137, RA 171 and RA 164, have been supported during the past 12 to 18 months due to the extremely high spacecraft tracking requirements; however, the outlook for next year is brighter.

D. Host Country Category

(1) Australia

The Pulsar observations in Australia have been the prime host country activity receiving support to the tune of about 400 hours in the last year.

(2) Spain

Spanish host country activity has been severely limited in the past year or so, but it is hoped that this situation will improve in the post-Pioneer Saturn Encounter period.

References

1. Fahnestock, R. J., "Ground-Based Radio Astronomy Activities Using DSN Facilities – FY79 Plan," Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, November 30, 1978 (Internal Document).
2. Gary, B., and Keihm, S., "Interpretation of Ground-Based Microwave Measurements of the Moon Using a Detailed Regolith-Properties Model," *9th Proc. Lunar Planet. Science Conference*, 1978, p. 2885.
3. Keihm, S., and Gary, B., "Comparison of Theoretical and Observed 3.55-cm Wavelength Brightness Temperature Maps of the Full Moon," *10th Proc. Lunar Planet. Science Conference*, 1979.
4. Keihm, S., and Gary, B., "Interpretation of High-Resolution Lunar Brightness Temperature Maps at 3.55 cm," American Geophysical Union Meeting, Washington, D.C., 28 May 1979.

A Voltage-Controlled Optical Radio Frequency-Phase Shifter

K. Y. Lau
Communications Systems Research Section

The problem of stabilizing an optical-fiber link, in which an RF-modulated optical wave is used for frequency standard transmission, is investigated in this article. Higher reliability is expected if RF phase-shifting, phase error detection, etc., can be done directly on the modulated optical signal. In the following pages, a variable optical phase shifter for the above purpose is proposed, and its properties are illustrated.

I. Introduction

To transmit time and frequency standards with accuracy of parts in 10^{14}, a feedback system must be used to stabilize the transmission path against external influence. The existing systems at the Deep Space Network use coaxial lines as the path, and the stabilization problem has been investigated by Lutes (Ref. 1). For stabilization, the signal at the receiving end of the path is transmitted, through the *same* path, back to the transmitting end, where its phase is compared with the standard reference source (a hydrogen maser). Any phase fluctuations at the end of the path are corrected by a voltage controlled phase-shifter inserted in the path. Reflections along the transmission line are highly undesirable for such reflections add signals travelling in opposite directions and vary their phases. This can be avoided if the forward and backward signals are of different frequencies, so that they can easily be isolated with filters. However, in this case, the dispersion of the line and the voltage-controlled phase-shifter ($VC\phi S$) produce errors in phase correction, as is evident from the following simple calculations: In Fig. 1, the reference signal of A is transmitted down the line, and the phase at the receiving end B is $\beta_1 L + \phi$, where L is the length of the line, β_1 the forward propagation constant, ϕ the phase shift introduced by the $VC\phi S$. On the return path, the propagation constant is β_2 which may be different from β_1. The phase of the received signal E_3 is thus $(\beta_1 + \beta_2) L + 2\phi$. The phase of E_3 is to be compared with the reference E_1. Now, suppose that the length of the line is changed from L to $L + \Delta L$, due to temperature change or any other means. Then, the phase of E_3 is changed, and the feedback system will produce a signal to change ϕ and compensate for the change in line length. The required change in the $VC\phi S$ is thus $\Delta\phi = -(\beta_1 + \beta_2)\Delta L$. With this change, however, the phase at B is

$$\beta_1(L + \Delta L) + \phi + \Delta\phi = \frac{\beta_1 - \beta_2}{2} \Delta L + \beta_1 L + \phi$$

and thus is a function of ΔL; i.e., phase at B is *not* stabilized unless $\beta_1 = \beta_2$. One can also see from this calculation why the forward and backward signal *must* propagate through the *same* line and *same* phase shifter in order to achieve phase stabilization at B.

In view of the above difficulties, an optical transmission line is proposed using radio frequency modulated optical signals transmitted through optical fibers for frequency and time standard distribution. Line dispersion of the modulated optical wave is virtually zero at all radio frequencies. However, the voltage-controlled phase shift ($VC\phi S$) cannot be performed directly on the optical signal, for a shift in the RF phase (5 MHz signal, say) of a quarter cycle means a change in path length of 16 m and a continuously variable optical delay line covering such a wide range is not available. Thus, one has to perform this operation at radio frequencies, and cannot avoid the dispersion of RF phase-shifters.

In the following paragraphs, a means of performing voltage-controlled phase shift directly on the modulated optical signal is described, and its characteristics are investigated. It is found that with this phase shifter, one can compensate, to a certain extent, any line dispersion, if it exists. Thus, one can use two different optical frequencies, instead of two different radio frequencies for forward and backward signals, even when optical dispersion exists. Shown in Fig. 2 is a proposed system, in which a 5 MHz standard signal is carried forward on 0.85μ and back on 1.06μ optical waves, corresponding to emissions of GaAs semiconductor laser and Nd:YAG laser, respectively.

The advantages of this system are numerous. Separation of forward and backward signals is easy with optical filters. Radio frequency mixers (phase detectors) and frequency dividers used in the present system can be avoided. The usual advantages of fiber optic links such as small size, light weight, insensitivity to electromagnetic interference, and ground loop problems hold for this system also.

II. The Optical Phase-Shifter

Directly shifting the phase of microwave signals on an optical carrier can be achieved with the scheme shown in Fig. 3. The RF signal on an optical carrier is split with either a voltage-controlled optical directional coupler, or beam splitter with variable attenuations in each path. One path goes through a fixed-length delay line with the *same* dispersion characteristic as the transmission line. The signals are recombined with a beam splitter (adder). Here, one assumes that the delay line length is long compared with the coherence length of the optical carrier, so that addition of the signals takes place *incoherently*. Any optical interference effect is thus disregarded. The phase ψ of the recombined signal is controlled by the splitting coefficient A; $A = 0$ corresponds to zero phase shift, and $A = 1$ corresponds to a phase shift of βl, where l is the delay line length and β the propagation constant at that optical frequency. Analytically, one has

$$\psi = \tan^{-1} \frac{A \sin \beta l}{A \cos \beta l + 1 - A} \tag{1}$$

and the recombined signal is thus

$$1 + p \cos \psi \tag{2}$$

where

$$p = [(A \sin \beta l)^2 + (A \cos \beta l + 1 - A)^2]^{1/2}. \tag{3}$$

A plot of ψ and p vs βl, for different A's, are shown in Figs. 4 and 5, where the average optical power is normalized to 1. One observes from Fig. 4 that ψ is fairly linear in βl, but not quite. In Fig. 6, $\psi - A\beta l$ is plotted vs. βl for different A's, and one can see the deviation from linearity for βl near $\pi/2$. For $A = 0, 1/2, 1$, ψ is exactly linear in βl.

One would, however, like the linearity of ψ in βl to hold for all values of A, since linearity is necessary for compensation of the line dispersion. To see this, one goes back to Fig. 1 and assumes that the phase-shifter introduces different phase shifts ψ_1 and ψ_2 on the forward and backward signal, in proportion to their respective β's. Thus

$$\begin{aligned} \psi_1 &= \phi\beta_1 \\ \psi_2 &= \phi\beta_2 \\ \Delta\psi_1 &= \Delta\phi\beta_1 \\ \Delta\psi_2 &= \Delta\phi\beta_2 \end{aligned} \tag{4}$$

To hold the phase of E_3 constant, one has

$$\begin{aligned} -(\beta_1 + \beta_2)\Delta L &= \Delta\psi_1 + \Delta\psi_2 \\ &= \Delta\phi(\beta_1 + \beta_2) \end{aligned} \tag{5}$$

Thus,

$$\Delta\phi = -\Delta L$$

and the phase at B is $\beta_1(L + \Delta L) + \psi_1 + \Delta\psi_1 = \beta_1 L + \psi_1$ independent of ΔL, i.e., stabilized regardless of the dispersion.

III. Compensation of Line Dispersion

One would like to ascertain how much error is introduced due to the slight nonlinearity of ψ vs. βl. To facilitate the analysis, we first develop an approximate formula for ψ near $A = 1/2$. This is done in the Appendix, giving to first order in ϵ

$$\psi = \frac{\beta l}{2} + 2\epsilon \tan \frac{\beta l}{2} \tag{6}$$

where

$$\epsilon = A - \frac{1}{2}, \text{ and}$$

l as before is the length of the delay arm of the phase shifter. The tangent term produces the undesirable nonlinearity. The phase shifts introduced to the forward and backward signal are thus, respectively,

$$\psi_1 = \frac{\beta_1 l}{2} + 2\epsilon \tan -\frac{\beta_1 l}{2}$$

$$\psi_2 = \frac{\beta_2 l}{2} + 2\epsilon \tan \frac{\beta_2 l}{2} \tag{7}$$

Assume that initially A is set to 1/2 (i.e., $\epsilon = 0$). Now, a change in line length ΔL produces an error signal that changes A and thus ψ_1 and ψ_2:

$$\Delta\psi_1 = 2\epsilon \tan \left(\frac{\beta_1 l}{2}\right)$$

$$\Delta\psi_2 = 2\epsilon \tan \left(\frac{\beta_2 l}{2}\right) \tag{8}$$

The amount of change in A necessary for phase stabilization is given by Eq. (5):

$$\Delta\psi_1 + \Delta\psi_2 = -\Delta L\,(\beta_1 + \beta_2)$$

hence giving

$$\epsilon = -\frac{\Delta L}{2}\left(\frac{\beta_1 + \beta_2}{\tan\frac{\beta_1 l}{2} + \tan\frac{\beta_2 l}{2}}\right) \tag{9}$$

With this change, the change in phase at the receiving end is

$$\begin{aligned}\delta &= \beta_1 \Delta L + \Delta\psi_1 \\ &= \beta_1 \Delta L + 2\epsilon \tan \frac{\beta_1 l}{2} \\ &= \left[\beta_1 - \frac{\left(\tan\frac{\beta_1 l}{2}\right)\Delta L\,(\beta_1 + \beta_2)}{\tan\frac{\beta_1 l}{2} + \tan\frac{\beta_2 l}{2}}\right]\Delta L\end{aligned} \tag{10}$$

One would of course like δ to be zero, as would be the case if the tangent functions are replaced by linear functions (i.e., ϕ-shifter linear in β), or $\beta_1 = \beta_2$ (i.e., no dispersion).

To proceed, we approximate

$$\tan \frac{\beta l}{2} = \frac{\beta l}{2} + \frac{1}{3}\left(\frac{\beta l}{2}\right)^3$$

For the range of interest, $\beta l < \pi/2$, the above expression is accurate to within five percent. Substitution in Eq. (10) gives

$$\delta = \left(\frac{\beta_1}{12}\right)\left(\beta_1^2 - \frac{\beta_1^3 + \beta_2^3}{\beta_1 + \beta_2}\right) l^2\ \Delta L. \tag{11}$$

Furthermore, dispersion data (Ref. 2)[1] gives the following: at carrier wavelength of 0.85μ, $\beta = \beta_1 = 2\,\pi/3984$ cm for 5 MHz modulating signal, and at 1.06μ, $\beta = \beta_2 = 2\,\pi/3989$

[1] For multimode fibers far from cutoff, intrinsic material dispersion dominates waveguide dispersion.

cm. Hence, we can write $\beta_2 = \beta_1 + \delta\beta$, where $\delta\beta$ is small. Substitution into Eq. (11) gives, to first order in $\delta\beta$,

$$\delta = \left(\frac{\beta_1}{12}\right)\left(\beta_1^3 - \frac{2\beta_1^3 + 3\beta_1^2\,\delta\beta}{2\beta_1 + \delta\beta}\right) l^2\,\Delta L$$

$$\approx \frac{\beta_1^3}{12}\left[1 - \left(1 + \frac{3}{2}\cdot\frac{\delta\beta}{\beta_1}\right)\left(1 - \frac{\delta\beta}{2\beta_1}\right)\right] l^2\,\Delta L$$

$$\boxed{\delta \approx \frac{\beta_1^2\, l^2}{12}\,\delta\beta\Delta L} \qquad (12)$$

As expected, this phase error is small when βl is small, for then ψ is more linear in βl.

Numerical Estimation: The present coaxial system uses a 300-m line operating at 5 MHz. Under field conditions, a fluctuation in line length of 30 cm is measured without line stabilization, and is reduced to ~0.02 cm when stabilized. Thus, the maximum phase error expected in the stabilized system is $\sim 3 \times 10^{-5}$ radians, limiting factors being line dispersion, phase shifter dispersion, and component stability. Now, *assume* that without stabilization, the optical fiber line experiences the same fluctuation of 30 cm, equivalent to 2π (30/4000) rads phase fluctuation. Then, taking $\beta l = 2 \times 2\,\pi$ (30/4000) rads, $\delta\beta = 2 \times 10^{-6}$ cm^{-1}[2], $\Delta L = 30$ cm, one estimates from Eq. (12) that $\delta \approx 4 \times 10^{-8}$ rads – three orders of magnitude lower than the present value.

Noise: What may hinder the system from approaching the above performance is the stability of the components (laser diodes, detectors) and noise. It is known that when laser diodes are coupled to fibers, reflections from fiber end faces cause the lasers to fluctuate (Ref. 3). This noise, the origins of which are not yet well understood, is large by any standard for a communication link. It could be suppressed by inserting an optical isolator between the fiber and the laser, or by anti-reflection coating the fiber end face which, as a bonus, would also increase the coupling efficiency. The stability of the components shall be investigated experimentally and if found unsatisfactory, feedback stabilization of individual components will be applied.

IV. Conclusion

Higher stability in frequency and time standard distribution is expected by using optical fibers and the above-mentioned optical phase-shifter. The simplicity of an optical system using separate optical wavelengths for forward and backward transmissions should render it more reliable. Further increase in performance can be obtained if we cascade a series of optical phase-shifters each of which has a small tunable range (small βl), for it was shown that the phase error increases as $(\beta l)^2$. Inherent to this phase-shifter, however, is a loss in average optical power (down 3 dB when the two arms are recombined with a beam splitter) and modulation depth (Fig. 5). This, together with the inevitable loss in optical equipment connections, does not recommend the cascade to go beyond two or three stages. One thus has to optimize the signal-power/phase-error tradeoff.

Further advantage can be made of the optical transmission link by processing signals directly in optics. For example, the phase error detector, as shown in Fig. 2, can be performed with a scheme shown in Fig. 7. The mixing process can be readily done with an optical modulator, thus avoiding RF equipment that is vulnerable to electromagnetic interference. Moreover, automatic gain control can be done with something as simple as an optical attenuator without affecting the RF phase, while it would be difficult at RF.

Acknowledgment

The author thanks G. Lutes for many enlightening and fruitful discussions.

References

1. Lutes, G., "A Transmission Line Phase Stabilizer," *The Deep Space Network Progress Report 42-51,* Jet Propulsion Laboratory, Pasadena, California.

2. Didomenico, M. "Material Dispersion in Optical Fiber Waveguides," *Applied Optics,* Vol. 11, pg. 652.

3. Hirota, O. and Suematsn, Y., "Noise Properties of Injection Lasers Due to Reflected Waves," *IEEE J. Quantum Electron.,* Vol. QE-15, pg. 142.

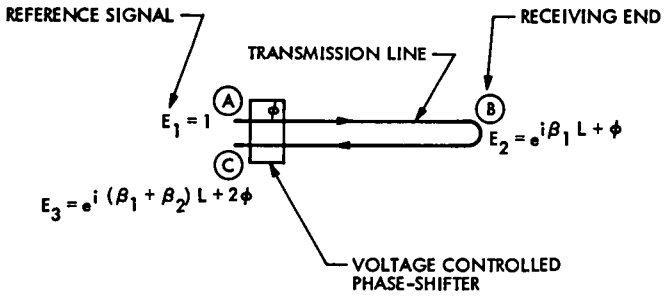


Fig. 1. Simplified diagram of a phase-stabilized transmission line

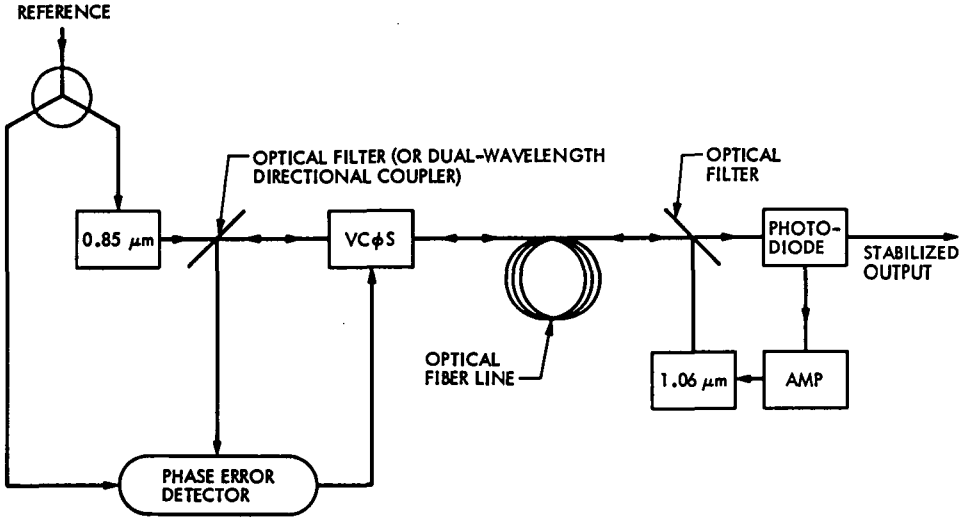


Fig. 2. Schematic diagram of a stabilized line using dual wavelength two-way transmission

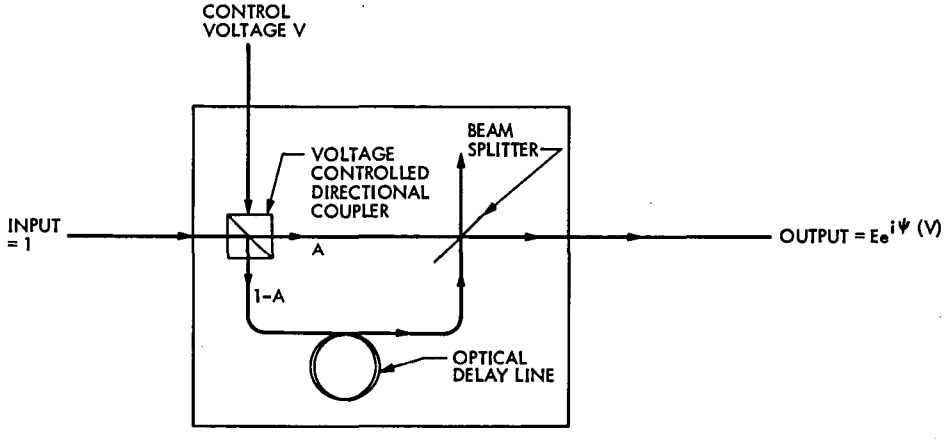


Fig. 3. An optical phase shifter

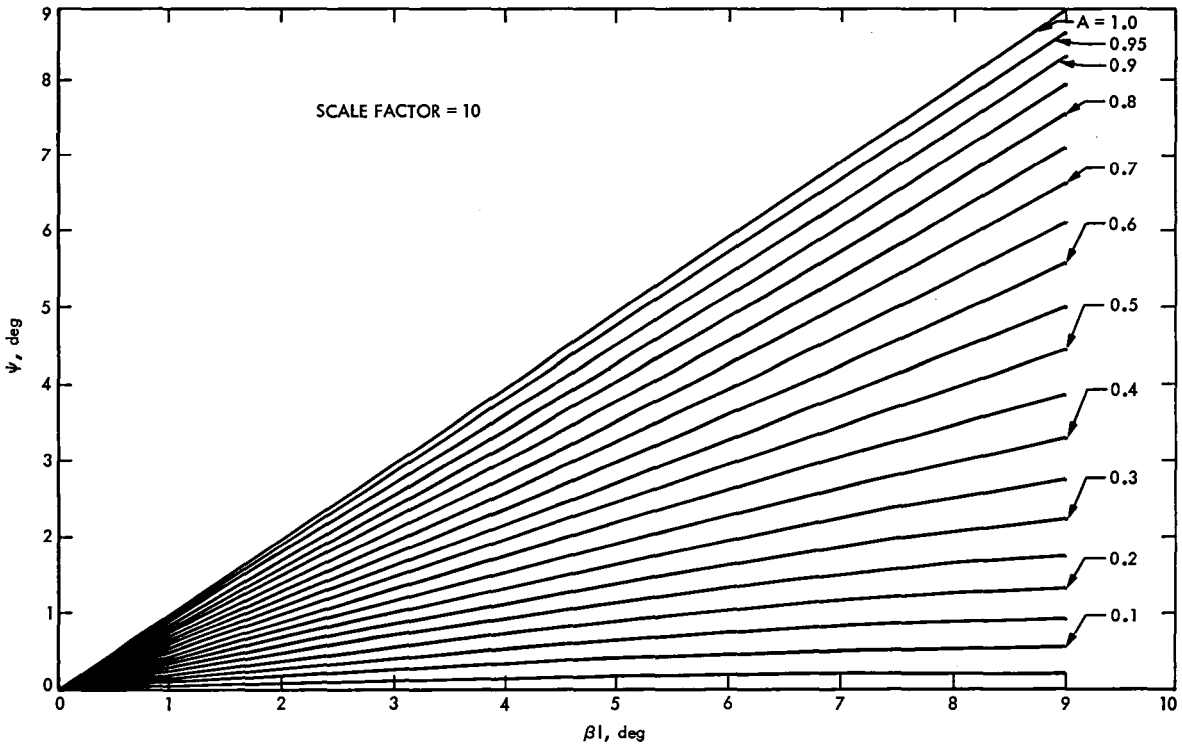


Fig. 4. The RF phase shift ψ vs. $\beta\ell$ for various splitting coefficients A

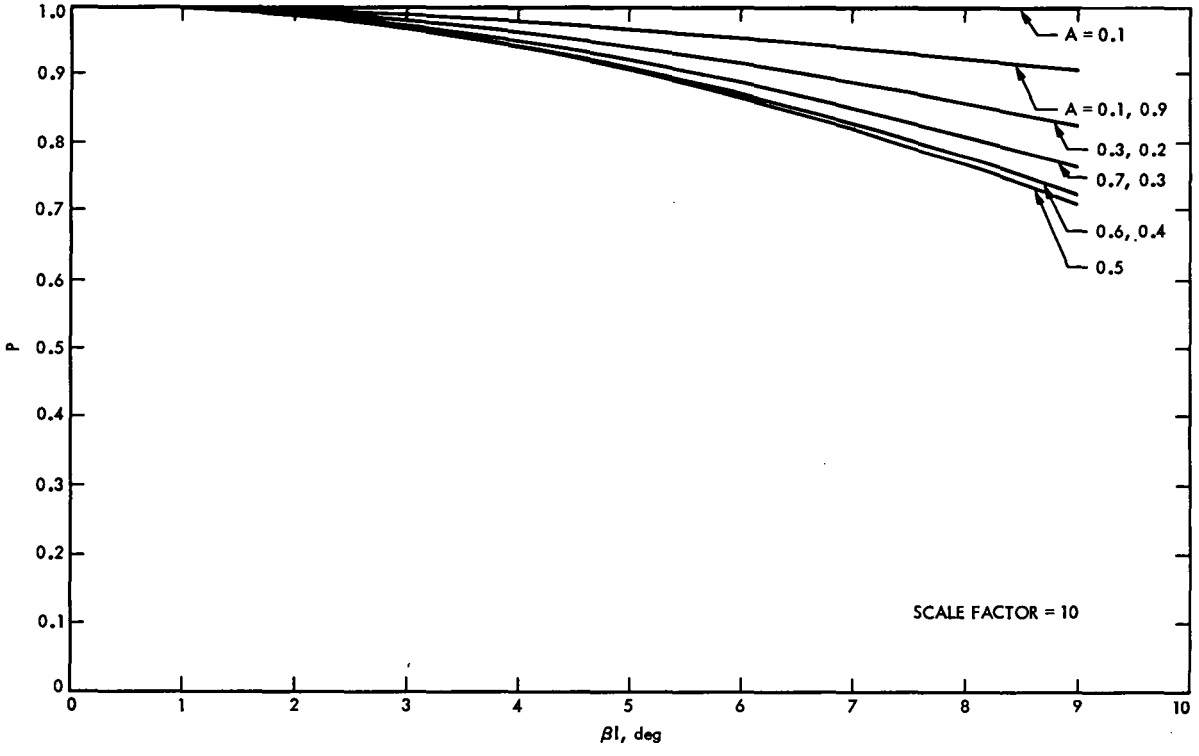


Fig. 5. The RF output power p vs. $\beta\ell$ for various splitting coefficients A

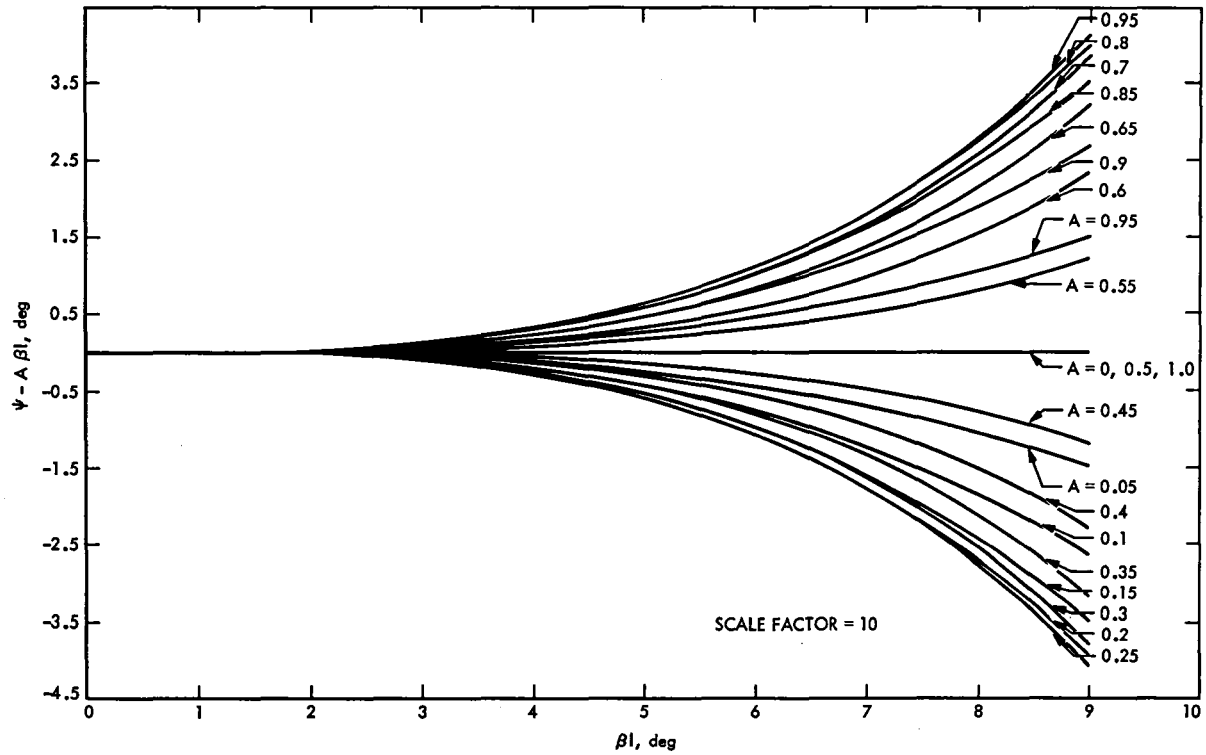


Fig. 6. Deviations from nonlinearity ψ-A$\beta\ell$ vs. $\beta\ell$ for various splitting coefficients A

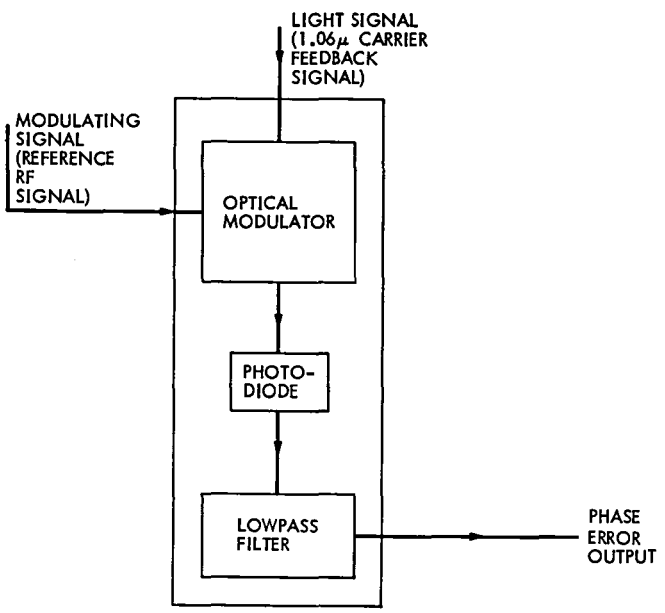


Fig. 7. Phase-error detector

Appendix

Equation 6 is derived below.

We have

$$\psi = \tan^{-1}\left(\frac{A \sin \beta l}{A \cos \beta l + 1 - A}\right).$$

For A near $1/2$, $A = 1/2 + \epsilon$, so

$$\psi = \tan^{-1}\left(\frac{\sin \beta l + 2\epsilon \sin \beta l}{\cos \beta l + 1 + 2\epsilon (\cos \beta l - 1)}\right)$$

$$= \tan^{-1}\left\{\frac{\sin \beta l}{1 + \cos \beta l}\left[1 + 2\epsilon\left(1 - \frac{\cos \beta l - 1}{\cos \beta l + 1}\right)\right]\right\}$$

$$= \tan^{-1}\left[\left(\tan\frac{\beta L}{2}\right)\left(1 + \frac{4\epsilon}{1 + \cos \beta L}\right)\right].$$

Using the expansion

$$\tan^{-1}(x + \epsilon) = (x + \epsilon) - \frac{(x + \epsilon)^3}{3} + \frac{(x + \epsilon)^5}{5} + \cdots$$

$$= \left(x - \frac{x^3}{3} + \frac{x^5}{5} \cdots\right)$$

$$+ \epsilon(1 - x^2 + x^4 - x^6 + \cdots) + 0(\epsilon^2)$$

$$= \tan^{-1} x + \frac{\epsilon}{1 + x^2} + \cdots$$

one obtains

$$\psi = \frac{\beta l}{2} + \frac{4\epsilon \tan \frac{\beta l}{2}}{(1 + \cos \beta l)} \times \frac{1}{1 + \tan^2 \frac{\beta l}{2}}$$

$$= \frac{\beta l}{2} + 2\epsilon \tan \frac{\beta l}{2}$$

This formula is accurate to within seven percent for ϵ as large as ± 0.25.

A Point Design for a Gyrotron Traveling Wave Tube Amplifier

A. Kupiszewski
Radio Frequency and Microwave Subsystems Section

The purpose of this article is to present a point design for a gyro-TWT amplifier. Steps for deriving the growth rate are explained and design results presented. Parameters for some support equipment are given and consequences of potential design parameter changes are examined.

I. Introduction

With current crowding at S- and X-bands and the greater resolution and ranging available at higher frequencies, great interest has been generated in exploring development possibilities for microwave amplifiers in the 33.5- to 35.1-GHz range for future generations of spacecraft uplinks and planetary radar. Power requirements are in the 200- to 400-kW CW range and an examination of presently available microwave tubes capable of attaining these powers at R-band frequencies has yielded little positive return. However, research and development on a new class of amplifiers that will hopefully satisfy the aforementioned needs is currently underway at many institutions throughout the world. These new amplifiers are called gyrotrons and have many advantages over the traditional microwave tubes, including the potential for using overmoded (and hence large) cavities or waveguides to minimize heat transfer problems inherent in any higher frequency scaling application. Also they are not constrained by the usual power-frequency relation, $P_{out} = \text{constant} \cdot 1/f^2$ (Ref. 1) Azimuthal phase bunching due to the dependence of the cyclotron frequency on the relativistic electron mass and magnetic coupling to RF fields in cavities and/or waveguides close to cutoff are responsible for providing the gain characteristic. Relatively high electronic efficiencies are calculated for gyrotron devices utilizing hollow cylindrical electron beams; further research and development will undoubtedly produce new advances in efficiency and output power.

At the present time, to further judge potential usefulness, it is prudent to examine gyrotron devices in view of their above-proposed uses and attempt to determine what kind of ancillary equipment, in terms of power supplies, cooling capacity, magnetic fields, etc., will be required. To do this, it is necessary to specify a particular design. Therefore, in the following sections, the mathematical starting point for the design of a gyrotron traveling wave tube amplifier is presented, and detailed results concerning bandwidth and physical dimensions are listed. It should be noted that this is a point design that is not necessarily completely optimized. Certain parameters such as maximum efficiency, for example, were assumed to be more important than bandwidth or gain per centimeter of interaction space considerations.

II. Dispersion Relation

It is necessary to derive a dispersion relationship so that wave growth rates in cylindrical coordinates can be calculated and hence efficiencies estimated. Therefore, starting with Maxwell's Equations,

$$\nabla \cdot \mathbf{E} = 4\pi\rho \tag{1}$$

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} \tag{2}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \tag{3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{4}$$

one obtains the following:

$$\left(\nabla_t^2 + \frac{\omega^2}{c^2} - k_z^2\right)\mathbf{E} = 4\pi i\left(\rho\mathbf{k} - \frac{\omega}{c^2}\mathbf{J}\right) \tag{5}$$

$$\left(\nabla_t^2 + \frac{\omega^2}{c^2} - k_z^2\right)\mathbf{B} = -\frac{4\pi}{c}\nabla \times \mathbf{J} \tag{6}$$

where c is the speed of light, **E** and **B** are the electric and magnetic fields in the waveguide respectively, **k** is the wave vector, ρ is the charge density, **J** is the current density, and ω is the frequency in radians. Using the TE_{01} vacuum waveguide solutions for **E** and **B** and the techniques outlined by Ott and Manheimer (Ref. 2), substitution for the current density via Vlasov Theory arguments leads to the following in the beam frame (Ref. 3):

$$\omega^2 - k_z^2 - \omega_c^2 = \frac{4\nu(1-\beta_{\|})^{1/2}}{J_0^2(x_1)} \times \left[\frac{-(\omega^2 - k_z^2)\beta_\perp^2 Y(x_1 r_0, x_1 r_L)}{(\omega - \Omega_c)^2} + \frac{\omega Z(x_1 r_0, x_1 r_L)}{\omega - \Omega_c}\right] \tag{7}$$

$$Y(a,b) = (J_1(a)J_1'(b))^2 \tag{8}$$

$$Z(a,b) = 2Y(a,b) + bJ_1'(b)J_1''(b)\cdot(J_1^2(a)(1+a^{-2}) + (J_1'(a))^2) + 2J_1(a)J_1'(a)J_1'(b)(bJ_1'(b) - J_1(b))/ab \tag{9}$$

where $x_1 = 3.8317$, ν is a dimensionless beam density parameter, $J_{0,1}$ represents Bessel functions, r_0 is the radius of the Larmor orbit guiding center in the waveguide (see Fig. 1), r_L is the Larmor radius, ω_c is the waveguide cutoff frequency, Ω_c is the cyclotron frequency, and primes imply derivatives with respect to arguments. Frequencies, velocities, and lengths have all been normalized to c/r_w, c, and r_w, respectively, where r_w is the waveguide radius (unnormalized).

Equation (7) will yield ω_r and ω_i. The latter represents the growth rate for the amplified wave.

The electronic efficiency, η, is given by

$$\eta = \frac{\text{field energy/unit length}}{\text{injected beam energy/unit length}} = \frac{\frac{1}{8\pi}\left(\frac{\omega}{\omega_c}\right)^2 \int_{\text{AREA}} \Psi^* \Psi \, da}{N(\gamma_0 - 1)mc^2} \tag{10}$$

where Ψ is the complex field wave function and * denotes complex conjugates. N is the number of electrons/unit length, m is the mass of the electron and

$$\gamma_0 = \left(1 - \beta_\perp^2 - \beta_{\|}^2\right)^{-1/2} \tag{11}$$

and $\beta_{\|}$ and $\beta_\perp$ are the normalized parallel and perpendicular electron velocities. Bandwidths can be determined from the width of the ω_i vs k_z curve.

The beam current, beam power, and output wave power are defined respectively by the following:

$$I_b = \nu(1-\beta_{\|})^{-1/2} mc^2 \beta_{\|} \tag{12}$$

$$P_b = I_b(\gamma_0 - 1)mc^2 \tag{13}$$

$$P_w = \eta P_b \tag{14}$$

Table 1 lists results that are based on the above equations and extrapolations of curves given in Ref. 3.

III. Discussion

The amplifier (see Fig. 2) can be sized for two different total gain conditions of 40 and 50 dB, respectively. Assuming a linear growth rate over the entire length of the fast-wave structure, the whole interaction-drift space (34.90 cm in the 40-dB case and 43.63 cm in the 50-dB case) must be enclosed in the bore of a superconducting solenoid of at least the same dimension so that the hypercritical control over the strength of the axial magnetic field can be maintained to better than one percent of maximum field strength over the entire waveguide region. As can be easily seen by subtracting the output wave power from the beam power, the collector coolant assembly must be capable of dissipating 318.04 kW of continuous wave power without upsetting the cryostatic stability of the superconducting magnets (which will probably be cooled with liquid helium). Hence, the collector will be the largest and heaviest part of the amplifier just as with conventional klystron tubes. However, due to the nature of the guiding magnetic field, the "spent" electrons will impact on the collector surface in a relatively small area, possibly resulting in local spot melting or flowing of the collector tube material, a cause of failure in some high-power klystron tubes. Peak power densities per collector tube length will be on the order of those for the X-3075 klystron, a 500-kW, CW S-band amplifier (Ref. 4), and optimization of collector design will reduce peak power densities to a more conservative level (around 4 kW per square inch). Stray electrons can be prevented from impacting on the output window by using an extra crossfield magnet farther along the collector tube length. This last idea may be extended to reduce the peak power densities. A deflection coil can be designed to spread the "spent" beam symmetrically about the collector walls and over their length, as suggested by Y. Carmel and J. Nation (Ref. 5).

All of the above-mentioned potential problems appear to be solvable without recourse to completely new technology development. The only major differences between the design for the gyro-TWT amplifier detailed above and conventional klystron tubes are that the guiding magnetic field is of a larger order of magnitude than usually encountered and a magnetron-injection gun is used to form a hollow cylindrical electron beam (see Figs. 3 and 4). This latter structure is quite different from the spherical cathode, solid-beam electron guns typically used in conventional high-power linear beam tubes (Ref. 6). The magnetron-injection gun has an annular emitting surface that gives off electrons having a generally well-specified spread of perpendicular and longitudinal velocities in cylindrical coordinates. Efficiency is enhanced through use of a hollow beam since there is no RF component on the axis of symmetry (Ref. 7). (This is true for most gyrotrom devices operating near cutoff, as with the gyro-TWT being presented in this paper and the gyromonitron tube discussed in Ref. 7.)

It is therefore quite conceivable that a test or prototype gyro-TWT could be built and/or tested at a high-power tube testing site, since the beam power required is 672.40 kW CW (beam perveance is 0.50 micropervs) and there are currently available power supplies capable of delivering even greater amounts of power than this. Should such a project be undertaken at some future time and conditions for bandwidth and gain finalized, the point design presented in this paper could be scaled subsequently to meet the new conditions without any major difficulties. In such an exercise, however, it is relatively safe to assume that a crucial factor will be the cost of the superconducting magnets, whose price scales with bore length. (This is generally true in most plasma experiments using superconducting magnets.) To minimize the magnet costs, it will be necessary to increase the gain per unit length parameter from 1.14 dB/cm to something over 2.2 dB/cm. This can be accomplished at the expense of the efficiency (electronic efficiency can be expected to drop from 52.7 percent to about 20 percent), but there will be a tremendous increase in bandwidth. A −3-dB bandwidth of 3.5 GHz may be possible. Needless to say, the collector will have to dissipate tremendous amounts of power. The weight of the cooling assemblies will undoubtedly have to increase greatly and, perhaps as well, the water-pumping capacity.

However, if a more moderate design is chosen, most of the above-mentioned factors can probably be easily worked out.

IV. Summary

The gyrotron traveling wave tube amplifier presented in this article represents a not necessarily optimized point design, where considerations of maximum efficiency have been given top priority over gain per centimeter of interaction space, bandwidth, and magnet bore length factors. From the dispersion relationship, it is possible to derive the growth rate in the beam frame using convenient normalizations. Lorentz transformations yield results in the laboratory frame that can then be converted to useful design parameters. The present design can be easily scaled for differing bandwidth or gain-per-unit-length requirements, once those factors are specified.

References

1. King, D. D. P., "Millimeter-Wave Prospectus," *Microwave Journal*, Vol. 10, November 1967, pp. 24-29.

2. Ott, E., and W. M. Manheimer, "Theory of Microwave Emission by Velocity-space Instabilities of an Intense Relativistic Electron Beam," *IEEE Transactions on Plasma Science*, PS-3, p. 1-5, 1975.

3. Chu, Drobot, Granatstein, and Seftor, "Characteristics and Optimum Operating Parameters of a Gyrotron Traveling Wave Amplifier," *IEEE Transactions on Microwave Theory and Techniques*, MTT-27, p. 178-187, 1979.

4. Goldfinger, A., *Study Program For Design Improvements of the X-3060 and X-3075, Phase I: Study Definition, Final Report*, JPL Contract No. 954782, Varian Asso., January 1978, pp. 24-25.

5. Carmel, Y., and J. Nation, "High-Power M/W Generation," *Microwave Journal*, pp. 50-51, June 1975.

6. Staprans, McCune, and Ruetz, "High Power Linear-Beam Tubes," *Proceedings of the IEEE*, Vol. 61, No. 3, pp. 299-330, March 1973.

7. Kupiszewski, A., "The Gyrotron: A High Frequency Microwave Amplifier," DSN Progress Report 42-52, in *The Deep Space Network Progress Report 42-52*, May and June 1979. Jet Propulsion Laboratory, Pasadena, Calif., pp. 8-12.

Table 1. Design parameters for a gyrotron traveling wave tube amplifier

Frequency	34.3 GHz	Guiding center radius	2.62 mm
Wave power	354.36 kW CW	Larmor radius	0.597 mm
Beam power	672.40 kW	Waveguide radius	5.46 mm
Current	9.50 A	Gain per unit length	1.146 dB/cm
Beam voltage	70.80 kV	Perpendicular electron velocity	0.401 C
Efficiency	52.7%	Parallel electron velocity	0.268 C
Magnetic field strength	13.07 kG	Beam density parameter	2.076×10^{-3}
	Conditions for maximum gain of 40 dB		
Length	34.90 cm		
Drive power	35.44 W		
−3 dB bandwidth	384 MHz (1.12%)	$c = 2.998 \times 10^{10}$ cm/sec	
−1 dB bandwidth	199 MHz (0.58%)		
	Conditions for maximum gain of 50 dB		
Length	43.63 cm		
Drive power	3.54 W		
−3 dB bandwidth	336 MHz (0.98%)		
−1 dB bandwidth	165 MHz (0.48%)		

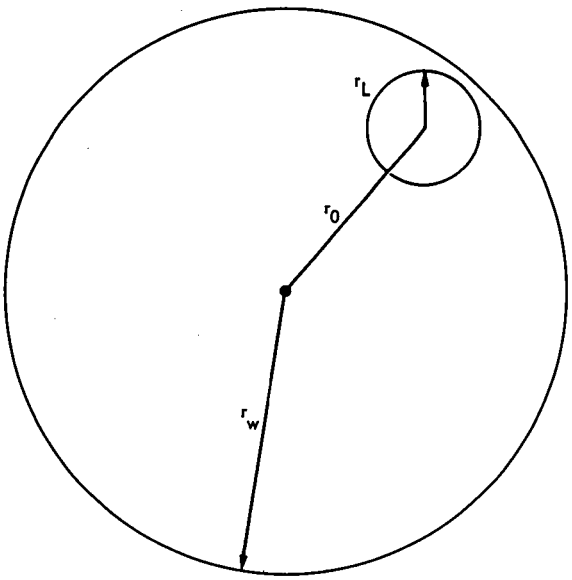


Fig. 1. Radial vectors

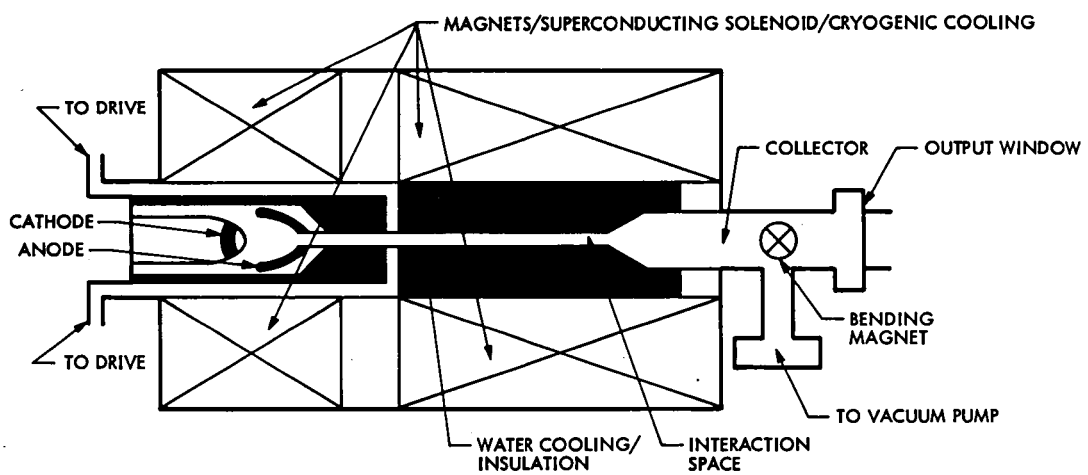


Fig. 2. Gyro-TWT amplifier

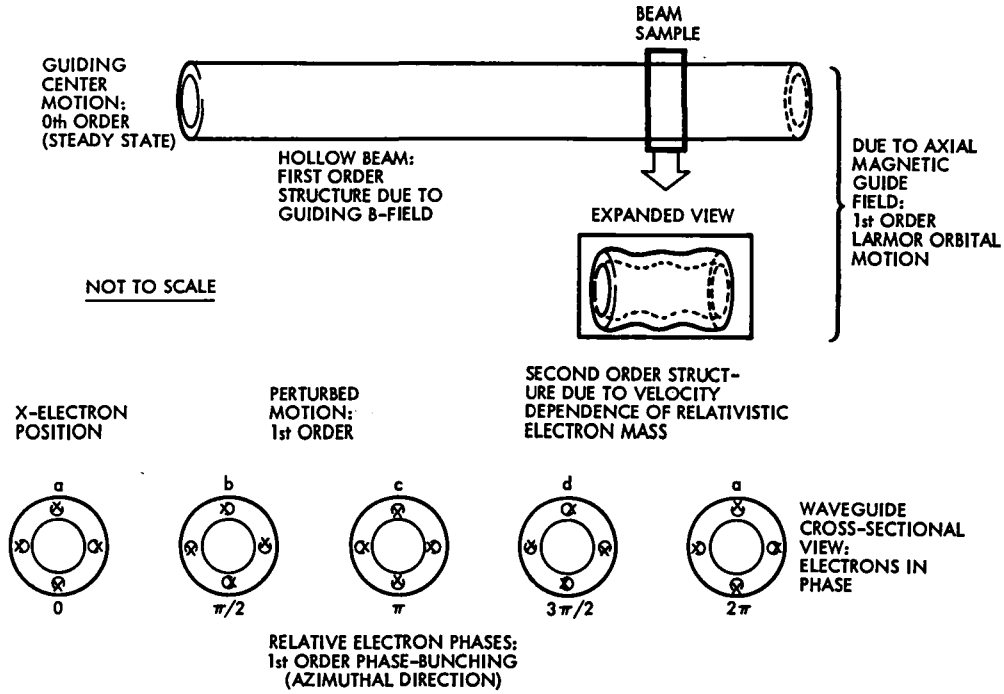


Fig. 3. The gyrotron hollow electron beam

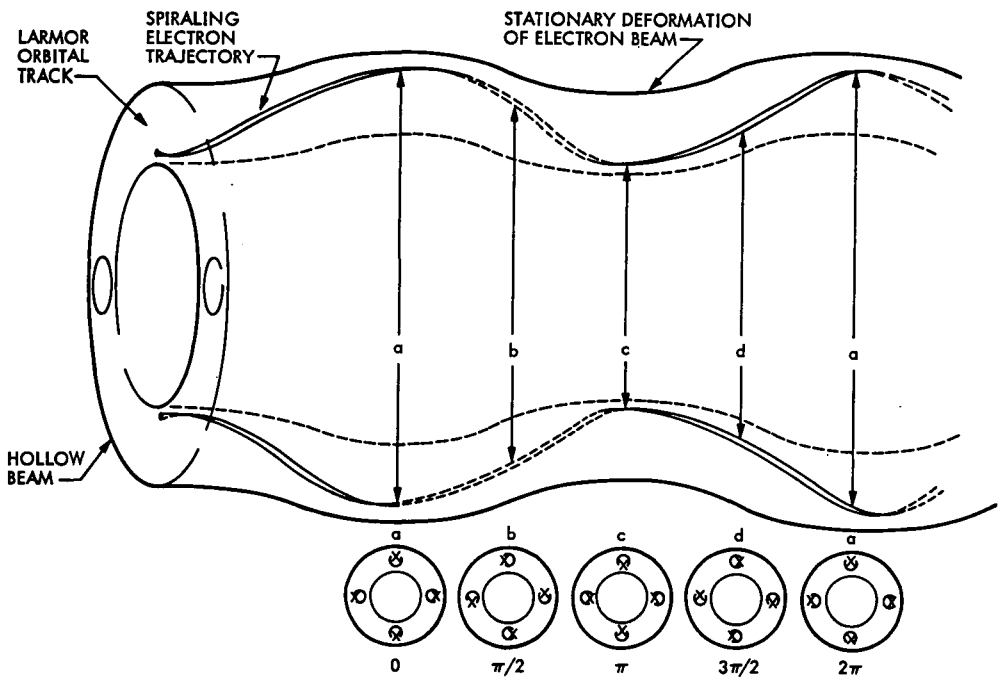


Fig. 4. Gyrotron hollow beam

Demonstration of Remote Clock Monitoring by VLBI, With Three Baseline Closure

C. M. Cheetham, W. J. Hurd, and J. W. Layland[1]
Communications Systems Research Section

G. A. Madrid and T. P. Yunck
Navigation Systems Section

The capability of very long baseline interferometry (VLBI) to monitor the stability of remotely-located hydrogen maser frequency standards has been demonstrated by a series of experiments conducted from September 1978, through January 1979, between Deep Space Stations in Australia, Spain, and California. The measured stabilities of the clock systems, over approximately 10-day intervals, were 1 to 3 parts in 10^{13}, with the instabilities due to the oscillators, the clock distribution systems, the receiving system delays, and the VLBI measurement error.

Experiments were conducted independently using two different systems (BLOCK 0 and WBDAS). Later comparison shows agreement in the order of 1 part in 10^{13}. Closure was demonstrated on three separate occasions to 33, 10, and 13 ns with an error uncertainty of ±42 ns. The results clearly demonstrate the resolution and consistency of VLBI measurements.

I. Introduction

In order to improve the quality of radiometric observables at outer planet distances, the monitoring of time offsets and frequency standards at DSN tracking stations has become a necessity. Very long baseline interferometry (VLBI) presents the most promising technique available to monitor clock epoch and rate offsets to the level required for advanced deep space missions. The feasibility of such an application has been independently demonstrated at JPL by Goldstein in 1967 (Ref. 1), Hurd in 1972 (Refs. 2-5), and Thomas, et al., in 1974 (Ref. 6). Using current technology the technique has the potential of determining frequency standard rates to 1 part in 10^{13} with a few minutes of observing time and to 1 part in 10^{14} over periods of approximately 1 week, by differencing clock offset estimates.

This article reports the results of a series of clock synchronization experiments that were conducted from September 1978 through January 1979, between DSSs in

[1] Now with TDA Planning Office.

Australia, Spain, and California. During this entire time, the cesium clock at DSS 63 in Spain and the hydrogen maser clock at DSS 43 in Australia drifted only a few microseconds with respect to the DSS 14 hydrogen maser clock at Goldstone, California. The measured stabilities of the clock systems were 1 to 3 parts in 10^{13}, with the instabilities due to the oscillators, the clock distribution systems, the receiving system delays, and the VLBI measurement error.

On three separate occasions, measurements were made between the Spain and Australia stations – the first VLBI measurements ever made on this extremely long baseline. Unfortunately, these experiments could not be conducted concurrently with measurements on the other baselines. However, the data were compared to clock offsets on the other two baselines, interpolated or extrapolated to the time of the Spain-Australia measurements. These experiments demonstrated closure to 33, 10, and 13 ns, with an error budget of ±42 ns. This is a powerful demonstration of the consistency of VLBI measurements.

Two different VLBI systems were utilized in parallel in the reported experiments, thus providing both redundancy and a comparison of the two systems. The main distinction between the two systems lies in the manner in which the digitized received energy signals are recorded and correlated. One application records the digitized signal across the entire received bandwidth at distinct intervals of time, the other records selected channels across the bandwidth in a continuous stream of data from which the phase delay across the bandwidth is then reconstructed. The first mode has accordingly been termed the "Wide Band Data Acquisition System" (WBDAS) (Refs. 4 and 5) while the second is termed the "Bandwidth Synthesis" System (BWS) (Refs. 6–9). The WBDAS approach was developed strictly for near-real-time clock and frequency monitoring, requiring a minimum amount of data. The BWS technique was developed to fully utilize the maximum recording rate of available recorders, with the intent of providing a technology base for use in the DSN that could be used not only for clock synchronization, but also for determining the station geodynamic parameters required for improved spacecraft navigation (i.e., earth rotation rate and crustal motion), and for improving the positional estimates of the extragalactic radio sources to be used to conduct VLBI experiments.

The specific BWS system used in these experiments is the Block 0 system, utilizing 4-Mbit/s digital recording on video tape recorders. This is an interim system leading to DSN implementation of a near-real-time Block I BWS System, and a wider bandwidth Block II System. At the time of these experiments, the DSN stations did not have phase calibrators (Ref. 10) and cable stabilizers that will be part of the Block I and Block II Systems, and that are required to do clock synchronization with the BWS technique. Thus clock synchronization with the Block 0 system could not use BWS, but used only the bit alignment of individual channels and therefore was similar in concept to the operation of the WBDAS.

II. Experiment Configuration

The two VLBI systems used in the experiments are briefly described in this section, and the configuration of the two VLBI data acquisition systems (Fig. 1) within the DSSs is discussed. It is argued that instabilities in the station 1-pps signal supplied to the VLBI systems, and in generation of VLBI epoch references from the 1-pps inputs, are probably the dominant cause of discrepancies between results for the two systems.

A. Signal Path

The signal from the radio source passes through the antenna system and a traveling wave maser (TWM) amplifier, and is then translated from the RF center frequency to 55 MHz, using reference frequencies generated from the station frequency standard. Both VLBI systems receive 55 MHz IF signals, but there is one more stage of amplification and filtering for the WBDAS than for the Block 0 System. This restricts the passband to 55 ± 18 MHz, whereas the signal to the Block 0 System is bandwidth limited by the TWM or by a 55 ± 36 MHz filter. The difference in group delays between the two systems is consistent from experiment to experiment at a level considered to be insignificant. The total signal delay does vary significantly from experiment to experiment, however. Variations in tuning of the TWM amplifiers can cause group delay variations of ±10 ns at each station, and the path length difference between the two TWMs at each station is as much as 52 ns, including waveguide lengths. Unfortunately, no accurate record of the TWM configurations was kept for all experiments, although each station was instructed to use the same TWM for each experiment whenever possible.

B. Block 0 System

Block 0 is a bandwidth synthesis VLBI system using 4-Mbit/s digital recording on video tape recorders and sampling sequentially up to eight BWS channels, each 1.8 MHz wide. The system is designed to measure fringe rates directly and group delay by either single channel bit stream alignment or bandwidth synthesis. Because the numerical differentiation inherent in bandwidth synthesis requires precise phase coherence between the channels, the group delays produced by that method are not meaningful without the incorporation of phase calibrators to remove instrumental effects. Since operational phase calibrators were not available for these experiments, the

group delays reported here were measured from the alignment offset of a single 1.8-MHz channel and are referenced to the Block 0 sampling clocks. As described below, the timing system implementation may have resulted in experiment-to-experiment delay changes which would not have occurred if the system had been in its BWS mode with phase calibrators. The system noise effects ranged from 3-20 ns, compared to the 0.1 ns which would be achieved with BWS.

For these experiments the system was configured to sample three S-band channels, spending 0.5 second in each. Although frequently X-band data were recorded as well, they were not included in these results.

C. Wideband System

The wideband VLBI data acquisition system utilizes a high instantaneous sampling rate in order to observe the entire signal bandwidth, as limited only by the receiving system. The receiver output is digitally modulated to baseband by sampling at 50 MHz in each of the two phase-quadrature analog-to-digital (A/D) converters. The time delay observable is the differential group delay to the A/D converter sampling clocks. The A/D converter outputs are low-pass filtered, by summing N consecutive samples in a digital integrate-and-dump filter. These experiments typically used $N = 3$, thus reducing the effective bandwidth to 16-2/3 MHz. The filter outputs are quantized to 1 bit, and stored in a 4096-bit buffer. When the buffer is full, sampling is inhibited and the buffer is emptied onto magnetic tape. Fourteen bursts of data are taken each second, for an average data rate of 57 kbits/s. The WBDAS achieves a lower signal-to-noise ratio than the Block 0 system because of the lower data rate, but achieves a time-delay error due to system noise of 1-5 ns, because of the wider bandwidth.

D. Station Timing

The VLBI systems are referenced to the station frequency standards through the coherent reference generator (CRG) and the time format assembly (TFA). Power to the frequency standards, the CRG, and the TFA is nominally uninterruptible, so phase and timing is in principle continuous except when catastrophic failures occur. The function of the CRG is to generate various frequency references coherently from the station standard. For the purposes of this experiment, the CRG probably does not degrade the station standard. The station 1 pps is generated from 1 MHz in redundant divider chains. Because the divider chains are constructed of obsolete and slow circuits, the 1-pps signal is reclocked by 5 MHz in the TFA. This reclocking results in possible 200-ns glitches, which have been observed at DSS 14 during the course of these experiments. Both VLBI systems initially synchronize their internal 1-pps references to the TFA 1-pps signal, and then allow the internal clocks to free run until synchronization is lost. This loss of synchronization normally occurs only when there is an interruption in power to the TFA or to a VLBI system. Such interruptions did not occur within one day's experiment, but did occur between experiments.

E. WBDAS Timing

The WBDAS 1-pps reference is generated from a 50-MHz signal from the CRG, by dividing this signal to 1 pps using emitter coupled logic (ECL). The internal 1 pps is initially synchronized to the TFA 1 pps, and thereafter the WBDAS monitors the phase difference between the TFA 1 pps and the internal 1 pps, in increments of 10 ns. The 10-ns resolution is achieved by observing the TFA signal both directly, and delayed by 10 ns. Normally the phase relationship does not change by more than 10 ns either within an experiment or between experiments. This 10-ns variation is expected, due to drifts in the WBDAS ECL circuits or in the TFA TTL circuits. Occasionally, jumps of 200 ns were observed at DSS 14; these jumps did not accumulate, but typically changed back and forth within an experiment on some days. We attribute this effect to the TFA. These jumps occurred only at DSS 14, and were always in the same direction. Therefore it is likely that the WBDAS clock was always consistently synchronized to the 50-MHz reference, within one 20-ns count interval, even when it was necessary to resynchronize due to power outages between experiments.

F. Block 0 Timing

The Block 0 VLBI System has a sampling rate of 4 Mbit/s, a frequency which is not available in the DSSs. The 4 Mbit/s is derived from 5 MHz in a phase locked loop synthesizer system. This system generates 1 MHz from the 5-MHz reference and from a 4-MHz voltage controlled oscillator (VCO), using digital dividers. The 1-MHz signals are then phase locked. A problem with this system is that the phase relationship between the 5 MHz and the 4 MHz can change up to 200 ns in increments of 50 ns upon resynchronization. Thus, power outages to the Block 0 System, and consequent resynchronization, may result in timing offsets in increments of 50 ns, in addition to the possible 200-ns TFA offset. This synchronization error is a likely source of discrepancies between the results from the two VLBI systems.

III. Results

A. Experiments

From 3 September 1978 to 21 January 1979, a total of 34 VLBI clock sync passes were scheduled. The pass durations ranged from approximately 2 hours to 25 hours. Each pass consisted of a number of runs, i.e., time spent taking data on a

particular source, separated by antenna move time. Eight of the longer passes were scheduled by the Block 0 experimenters and consisted of 2.5-minute runs. The other passes were scheduled by the WBDAS experimenters and consisted of 9-minute runs. Because the time required for setup was uncertain, runs were scheduled from the start of the pass. Thus data was not always taken on the initial runs or the final runs. Some passes were not successful at all due to equipment failures in one or both VLBI systems, or in the DSS configuration. The results in this section are the estimates of the clock offset for the successful passes.

B. Processing and Results

1. Processing. The WBDAS results were produced in two stages. The first stage correlated the data from each separate run and produced an estimate of the clock offset and its rate of change, as well as estimates of the standard deviation for each parameter. The second stage combined the estimates for each successful run in a pass and produced an estimate for the clock offset and its rate for the pass.

Table 1 contains the results of the second stage. The column labeled Date contains the nominal date of the experiment, Epoch contains the time of the clock estimate, Clock contains the clock offset, Sigma clock contains the formal uncertainty of the clock offset, Residual contains the rms residual of the runs with respect to the clock estimate for the pass, Clock rate contains the rate of change of the clock offset, Sigma clock rate contains the formal uncertainty of the clock rate, No. of observations contains the number of runs or observations in the pass that were used to produce the pass estimates.

2. Closure Results. Three of the passes were performed on the 43/63 baseline, using only the WBDAS System. This provides a consistency check on the clock offsets, since the offsets between pairs of stations should sum to zero. Because the reference time for each pass is different and the clocks are all drifting, it is necessary to make some estimates to reference these clock offsets to the same epoch. Figure 2 shows the 14/43 and 14/63 clock offsets used to make these estimates. The clock offsets are modelled by straight-line, least-square fits to the data, based on the assumption that the clocks at the three stations have constant but different frequencies. On or about 16 November, the clock at DSS 43 apparently had a sudden frequency change and so two straight-line fits are made to the 14/43 clock offset. Table 2 contains the data used to calculate the 14/43 and 14/63 fits. The rms residuals to these fits are on the order of 70 ns. The closure is to about 10 to 30 ns. It should be noted that the 14/63 data has to be extrapolated to 14 October. The earliest 14/63 experiment used here was on 23 October since the preceding experiment of 1 October deviated considerably from the straight-line fit. The hydrogen maser at DSS 63 failed in September and presumably had not settled at its final frequency on 1 October. The fact that the closure is as good as it is suggests that it had settled before 14 October.

C. Block 0 Processing and Results

In the Block 0 System, the digital video tapes are shipped from the stations to JPL, then cross-correlated in quadrature on the hardware processor at Caltech. Postcorrelation analysis is performed on the IBM 3032 at Caltech and begins with a step called "phase-tracking" in which each source observation (typically 3-9 minutes) is divided into segments 20 to 60 seconds long. Each segment is fit by least squares to a complex sinusoid giving solutions for amplitude, phase, and fringe rate. Simultaneous interpolation of fringe amplitude in the lag domain with a sin x/x function yields the single channel group delay, while cross-channel differencing of phase solutions yields the synthesized delay for each segment. A priori values for the first segment solution are taken from an initial Fourier analysis, while those for the other segments are taken from the solutions for the segment preceding.

Segment solutions are then analyzed collectively to yield a solution for the entire observation. Amplitude and alignment delay are obtained by a weighted average of segment solutions, while fringe phase and rate result from a linear fit to segment phases. A linear fit to synthesized delays gives the final synthesized delay and a direct measurement of delay rate.

In the final processing step, solutions for all observations are supplied to a global fitting program which produces, for the entire experiment, single solutions for clock offset, fringe rate, and clock rate. In addition, when the number of observations is sufficient (typically >7), the program redetermines the baseline, thus providing some compensation for errors in the a priori UT1 and PM values. Although the program can also solve for selected source positions, we did not use that feature, electing instead to discard obviously bad data.

The Block 0 data reported here are from this final processing step. The clock rate reported is that derived from fringe rate rather than from delay rate, because this is more accurate. The rate accuracy is currently limited over the short term by systematic and random effects of ionosphere, instrumentation, and modeling errors.

Table 3 contains the Block 0 results and provides the accumulated "Allan Variance" stability estimates derived from the clock offsets. Note that the accumulation has been restarted at points of major breaks or jumps in the clocks.

D. Comparison of Results

Figure 3 shows the DSS 43 minus DSS 14 clock offsets versus epoch for both the WBDAS and Block 0 Systems. The offset is nearly linear from 17 September (Epoch 22.5) to 16 November (Epoch 27.7). The hydrogen maser failed at DSS 43 between the 3 September and 11 September passes and was put back on line just before the 11 September pass. Thus the early clock offsets are not colinear with those following. As mentioned above, the WBDAS data shows a rate change on about 16 November. This rate change is not as precisely located in the Block 0 data since there was no Block 0 result from 16 November. The frequency of the DSS 43 maser was intentionally shifted in late December and thus the offsets from 20 December (Epoch 30.7) to 12 January (Epoch 32.7) have a different rate than those previous.

Figure 4 shows the DSS 63 minus DSS 14 clock offsets versus epoch for the two systems. The DSS 63 hydrogen maser was not on line until January of 1979, thus only the last two points represent a comparison of two masers. However, the data is quite linear from 23 October (Epoch 25.7) to 24 December (Epoch 30.9) while DSS 63 was on the cesium standard.

The scale of Figs. 3 and 4 permits only a coarse comparison of the two systems. However, Fig. 5 shows the 14/43 data with a linear clock estimate removed, $\hat{c} = -26.71 + 1.35 \times$ Epoch. In addition, a constant of 0.4 microseconds has been added to the Block 0 data, which represents an estimate of the difference of the signal and clock path lengths in the two systems. The rate change in the 14/43 offset mentioned above is quite obvious in Fig. 5, however it should be noted that the slope is not really negative past 16 November since the axis of Fig. 5 has a slope of 1.35×10^{-12}. There are 10 passes on the 14/43 baseline for which both systems reported results; the rms difference (after removal of the 0.4-μs offset) is 64 ns.

Figure 6 shows the 14/63 data with a linear clock estimate removed, $\hat{c} = -11.4 + 0.29 \times$ Epoch. In addition, a bias of 56 ns has been added to the Block 0 data. The rate of 0.29×10^{-12} reflects the rate observed from 23 October (Epoch 25.7) to 24 December (Epoch 31.0). Before 23 October, the points are outside the range of Fig. 6, due to station clock adjustments. A large clock jump occurred between 24 December and 16 January, so 1.1 μs was subtracted from the passes on 16 January and 21 January to keep them on Fig. 6. There were 7 passes between 23 October and 24 December for which both systems reported results. The average difference between the two systems was 56 ns and after removal of this constant, the rms difference was 105 ns. Two days, 5 November and 3 December, disturb these calculations and may be the result of clock jumps.

IV. Analysis of Results and Error Sources

The principal objectives of VLBI clock sync experiments are to determine the offset and combined stability of the station frequency standards. In the present case, with near-simultaneous results from two different VLBI systems, we can also form some conclusions about the performance of the VLBI technique itself.

On the 14/63 baseline, over the period during which the cesium standard was on line at 63, the Block 0 results show a frequency offset of 3×10^{-13} and a stability (square root Allan variance) of 2×10^{-13} with a sample standard deviation on the latter figure of 0.9×10^{-13}. The WBDAS results show comparable values of 2.7×10^{-13} for the offset and 1.2×10^{-13} for the stability. On the 14/43 baseline both sets of results show a change in frequency offset sometime in the period from mid-November to early December. A lack of Block 0 results for mid-November prevents a more accurate determination of the time of the change. Block 0 data yield an offset of 1.7×10^{-12} before the change and 8.2×10^{-13} after, with a stability over the whole interval of 3×10^{-13}. The sample standard deviation is 1×10^{-13}. From the WBDAS data, the estimated offsets are 1.7×10^{-12} before the change and 1.2×10^{-12} after, with an overall stability of 1.9×10^{-13}. The average interval between samples is approximately 10 days; however, because the intervals vary, the Allan variances must be considered nonstandard.

Because neither S/X ionosphere calibration nor instrumental phase calibration were employed in these experiments, errors in the measured clock offsets are dominated by transmission media and instrumental effects. The stability values should therefore be regarded only as loose upper bounds on the instability of the clocks themselves.

Figures 5 and 6 show the differences between the WBDAS and Block 0 measured clock offsets, after the removal of a constant bias, on the days for which both systems obtained measurements. With the exception of a few greater discrepancies, the agreement is at about the 50-ns level. In all likelihood, the larger discrepancies are due not to large random errors but rather to real temporary changes in the instrumental delays of one system with respect to the other. For example, it has been observed that reinitializing the Block 0 clock, which is routinely done, can change its epoch with respect to the station clock by several hundred nanoseconds. The phase calibration systems now being installed will remove the effects of those jumps.

The VLBI systems described here have measured the combined instability over ~10-day intervals of frequency standards separated by intercontinental distances to low parts

in 10^{13} with an uncertainty of 1 part in 10^{13}. It is known that well-maintained hydrogen masers will show a stability over such intervals of a part in 10^{14} or better. With VLBI systems now under development using dual-frequency ionosphere calibration, accurate measurement and modeling of the wet and dry components of the troposphere, instrumental phase calibration, and simultaneous solution for UTI and polar motion corrections, stability measurements of a few parts in 10^{14} should be attainable.

Although a thorough analysis of the error sources affecting the interferometer used for these experiments has not yet been completed, a preliminary set of mean value error estimates have been compiled based on prior experience with the instrument. These values are presented in Table 4 mainly to indicate the estimated relative magnitude of the errors at S-band. For the sake of consistency, the geometric effects have been scaled to a hypothetical 10,000 km. baseline and approximate worst case partials are provided as well as the corresponding one sigma error values. In the cases of "System Noise", "Instrumentation", "Ionosphere", and "Bandpass Shape", only those values footnoted by an "a" or a "c" are applicable to the current instruments. The other values presented in these categories represent an estimate of the level to which these error sources will be reduced once phase calibration and dual-frequency charged particle error cancellation have been introduced.

The system noise contribution ranges from 1 ns for strong sources with either system, to 20 ns with the Block 0 System for sources too weak to be determined with the WBDAS. The range of 10-40 ns for instrumentation effects depends on the station configuration integrity. The bandpass shape factor is smaller, about 1 ns, for the WBDAS than the 10 ns for the Block 0, due to system bandwidth utilized (Ref. 11). Since there are normally some strong radio sources in an experiment, the dominant error sources are instrumentation stability and the ionosphere, whose contributions cannot be accurately estimated. Overall, the error of the current measurements is believed to be in the range of 24-51 ns.

For the closure experiment, the expected error is $\sqrt{3}$ times the individual experiment error, or 42-88 ns, neglecting frequency stability induced errors. Thus the closures of 33, 10, and 13 ns were better than anticipated.

Acknowledgments

The authors express their appreciation to B. Bronwein, E. Cohen, R. Henderson, R. Shaffer, D. Spitzmesser, and the Deep Space Station personnel for their assistance in scheduling, configuring, and operating the stations, and processing the data.

References

1. Goldstein, R., "Clock Calibration via Quasar," in the *Space Programs Summary 34-48.* Vol. II, pp. 79-82, Jet Propulsion Laboratory, Pasadena, Calif., November 30, 1967.

2. Hurd, W. J., "A Demonstration of DSN Clock Synchronization by VLBI," in *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. XII, pp. 149-160, Jet Propulsion Laboratory, Pasadena, Calif., December 15, 1972.

3. Hurd, W. J., "An Analysis and Demonstration of Clock Synchronization by VLBI," *IEEE Transactions on Instrumentation and Measurement* Vol. IM-23, No. 1, pp. 80-89, March 1974.

4. Hurd, W. J., "Preliminary Demonstration of Precision DSN Clock Synchronization by Radio Interferometry," in the *Deep Space Network Progress Report* 42-37, Jet Propulsion Laboratory, Pasadena, Calif., February 15, 1977, pp. 57-68.

5. Hurd, W. J., et al., "Submicrosecond Comparison of Intercontinental Clock Synchronization by VLBI and the NTS Satellite," *Proceedings of the Tenth Annual Precise Time and Time Interval Applications and Planning Meeting*, NASA TM 80250, Goddard Space Flight Center, Greenbelt, MD., Nov 28-30, 1978, pp. 629-642. Also in the *Deep Space Network Progress Report*, Technical Report 32-1526, Vol. 49, pp. 64-69, Jet Propulsion Laboratory, Pasadena, Calif., February 15, 1979.

6. Thomas, J. B., et al., "Radio Interferometry Measurements of a 16 km Baseline with 4 cm Precision," in *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. XIX, pp. 36-54, Jet Propulsion Laboratory, Pasadena, Calif., February 15, 1974.

7. Thomas, J. B., "An Analysis of Long Baseline Radio Interferometry," in *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. VIII, p. 37, Jet Propulsion Laboratory, Pasadena, Calif., February 1972.

8. Thomas, J. B., "An Analysis of Long Baseline Ratio Interferometry, Part II," in *The Deep Space Network Progress Report.* Technical Report 32-1526, Vol. VIII, p. 29, Jet Propulsion Laboratory, Pasadena, Calif., May 1972.

9. Thomas, J. B., "An Analysis of Long Baseline Radio Interferometry, Part III," in *The Deep Space Network Progress Report*, Technical Report 32-1526, Vol. XVI, p. 47, Jet Propulsion Laboratory, Pasadena, Calif., August 15, 1973.

10. Thomas, J. B., "The Tone Generator and Phase Calibration in VLBI Measurements," in *The Deep Space Network Progress Report*, 42-44, Jet Propulsion Laboratory, Pasadena, Calif., February 1978, pp. 63-74.

11. Layland, J. W. and W. J. Hurd, "VLBI Instrumental Effects, Part I," in *The Deep Space Network Progress Report*, 42-42, Jet Propulsion Laboratory, Pasadena, Calif., December 1977, pp. 54-80.

Table 1. WBDAS results

Date	Epoch, secs past 1 Jan 1978 × 10^6	Clock, ns	Sigma clock, ns	Residual, ns	Clock rate, × 10^{-12}	Sigma clock rate, × 10^{-12}	No. of observations
			14/43				
11 Sep	21.9816	3.551	0.4	5.3	0.78	0.08	4
17 Sep	22.518	3.744	23.9	62.8	3.0	0.7	7
23 Sep	23.0292	4.379	18.0	–	5.0	2.0	1
30 Sep	23.6304	5.292	0.4	5.4	1.41	0.06	6
14 Oct	24.8436	7.320	0.6	2.1	3.4	0.15	2
23 Oct	25.6176	8.591	0.7	7.9	1.8	0.1	3
27 Oct	26.0064	9.316	0.2	10.9	2.28	0.007	23
4 Nov	26.6724	10.659	0.4	11.5	0.98	0.04	9
16 Nov	27.6912	12.337	0.5	5.4	1.6	0.07	4
29 Nov	28.8072	13.688	0.6	11.2	2.1	0.2	3
12 Dec	29.9808	15.125	1.1	4.3	−2.8	0.9	2
31 Dec	31.5864	16.407	0.2	8.2	0.18	0.005	14
12 Jan	32.6628	17.423	0.5	14.2	0	0.15	6
			14/63				
9 Sep	21.8484	−8.768	0.8	6.8	−1.8	0.2	5
1 Oct	23.7348	−2.443	0.15	14.8	0.7	0.4	4
23 Oct	25.6536	−3.845	0.8	7.6	−1.4	0.2	5
30 Oct	26.2008	−3.705	0.7	18.9	4.09	0.08	9
5 Nov	26.7984	−3.451	0.3	32.4	0.66	0.01	10
20 Nov	28.0044	−3.303	0.5	5.9	1.0	0.1	6
27 Nov	28.6092	−3.126	0.5	6.4	0.3	0.1	5
3 Dec	29.1996	−2.955	1.4	4.3	−0.7	0.3	5
24 Dec	30.9492	−2.368	0.3	6.0	−0.28	0.04	10
16 Jan	32.9472	−0.664	0.4	4.2	−0.5	0.07	6
21 Jan	33.3756	−0.564	0.8	4.5	−0.2	0.1	4
			43/63				
14 Oct	24.858	−11.400	1.0	18.2	−3.4	0.3	5
3 Nov	26.604	−14.033	1.0	1.4	−4.0	0.2	3
28 Nov	28.7532	−16.643	2.0	6.3	−4.0	0.6	2

Table 2. WBDAS closure results

Date	Epoch	Clock	Residual to Fit	
			14/43	
30 Sep	23.6304	5.292	0.070	$\hat{T}_{43-14}$ = −36.1159
14 Oct	24.8436	7.320	−0.025	+ 1.74937 X Epoch
23 Oct	25.6176	8.591	−0.108	rms residual = 0.076
27 Oct	26.0064	9.316	−0.063	
4 Nov	26.6724	10.659	0.115	
16 Nov	27.6912	12.337	0.011	
16 Nov	27.6912	12.337	0.003	$\hat{T}'_{43-14}$ = −21.386
29 Nov	28.8072	13.688	−0.005	+ 1.2177 X Epoch
12 Dec	29.9808	15.125	0.003	rms residual = 0.004
			14/63	
23 Oct	25.6536	−3.845	0.007	$\hat{T}_{64-14}$ = −10.669
30 Oct	26.2008	−3.705	0.001	+ 0.26574 X Epoch
5 Nov	26.7984	−3.451	0.097	rms residual = 0.062
20 Nov	28.0044	−3.303	−0.076	
27 Nov	28.6092	−3.126	−0.060	
3 Dec	29.1996	−2.955	−0.046	
24 Dec	30.9492	−2.368	0.077	
			43/63	
			$\hat{T}_{43-14}$	$\hat{T}_{63-14}$ RESID = $T_{63-43} + T_{43-14} - T_{63-14}$
14 Oct	24.858	−11.400	7.370	−4.063 0.033
3 Nov	26.604	−14.033	10.424	−3.599 −0.010
28 Nov	28.7532	−16.643	13.628	−3.028 0.013

Table 3. Block 0 results

Date	Epoch	Clock, μs	Sigma clock, ns	Rate, $\times 10^{-12}$	No. of observations	Square root Allan variance, $\times 10^{-13}$
			14/43			
3 Sep	21.322339	4.368	5.2	−2.38	32	–
17 Sep	22.518534	3.377	9.7	2.22	7	–
30 Sep	23.629294	4.844	1.7	1.93	5	–
14 Oct	24.844771	6.910	1.0	3.82	3	2.68
23 Oct	25.618483	8.156	4.7	3.26	4	1.95
27 Oct	25.975236	8.744	2.4	2.31	117	1.60
4 Nov	26.670332	10.317	5.0	5.96	24	2.58
29 Nov	28.806020	13.397	1.6	2.91	4	3.47
13 Dec	29.981315	14.776	6.7	1.08	3	3.26
20 Dec	30.671408	15.361	2.0	2.23	3	3.14
31 Dec	31.543594	15.955	1.7	1.68	33	2.97
13 Jan	32.662017	16.984	14.6	2.32	6	2.86
			14/63			
4 Sep	21.385637	−7.668	3.8	0.68	56	–
16 Sep	22.410587	−10.059	7.3	−5.15	59	–
23 Oct	25.655368	−3.965	8.0	−0.62	5	–
30 Oct	26.180608	−3.828	6.2	−1.82	57	–
5 Nov	26.754849	−3.672	3.9	0.24	127	0.08
20 Nov	28.005401	−3.329	9.3	0.13	7	0.06
27 Nov	28.611230	−3.184	18.9	0.09	5	0.15
3 Dec	29.202867	−2.800	7.5	1.05	3	1.45
16 Dec	30.249901	−2.604	8.0	0.16	17	1.96
24 Dec	30.942036	−2.385	5.3	0.16	15	1.82
16 Jan	32.947914	−0.582	2.1	−2.08	4	
21 Jan	33.374313	−0.469	15.5	−0.84	5	

Table 4. Estimated magnitudes of error sources at S-band for a 10,000-km baseline

	Delay			Delay rate, $x\omega$		
Error Source	Partial	1σ	Final value	Partial	1σ	Final value
Source position	130 ns/"	0.015"	2 ns	22 mHz/"	0.015"	0.33 mHz
Baseline	3.33 ns/m	1.0 m	3.33 ns	0.56 mHz/m	1.0 m	0.56 mHz
UT1	2.0 μs/sec	0.003 sec	6 ns	0.32 Hz/sec	0.003 sec	1.0 mHz
PM (X)	3.33 ns/m	0.7 m	2.3 ns	0.56 mHz/m	0.7 m	0.4 mHz
PM (Y)	3.33 ns/m	0.7 m	2.3 ns	0.45 mHz/m	0.7 m	0.4 mHz
System noise			1–20 ns[a]			2.3 mHz[a]
	n/a[b]	n/a	3.5 ns	n/a	n/a	0.4 mHz
Instrumentation			20–40 ns[a]			1.0 mHz[a]
	n/a	n/a	3 ns	n/a	n/a	0.1 mHz
Ionosphere			10–20 ns[c]			4.0 mHz[b]
	n/a	n/a	~ 0 ns	n/a	n/a	~ 0 mHz
Troposphere	n/a	n/a	1 ns	n/a	n/a	0.2 mHz
Bandpass shape			1–10 ns[a]			
	n/a	n/a	~ 0 ns	n/a	n/a	~ 0 mHz
Root sum square			24–51 ns[a,c]			5.0 mHz[a,c]
			9.2 ns			1.4 mHz

[a] Without phase calibration.

[b] Not applicable.

[c] Without S/X calibration of the ionosphere.

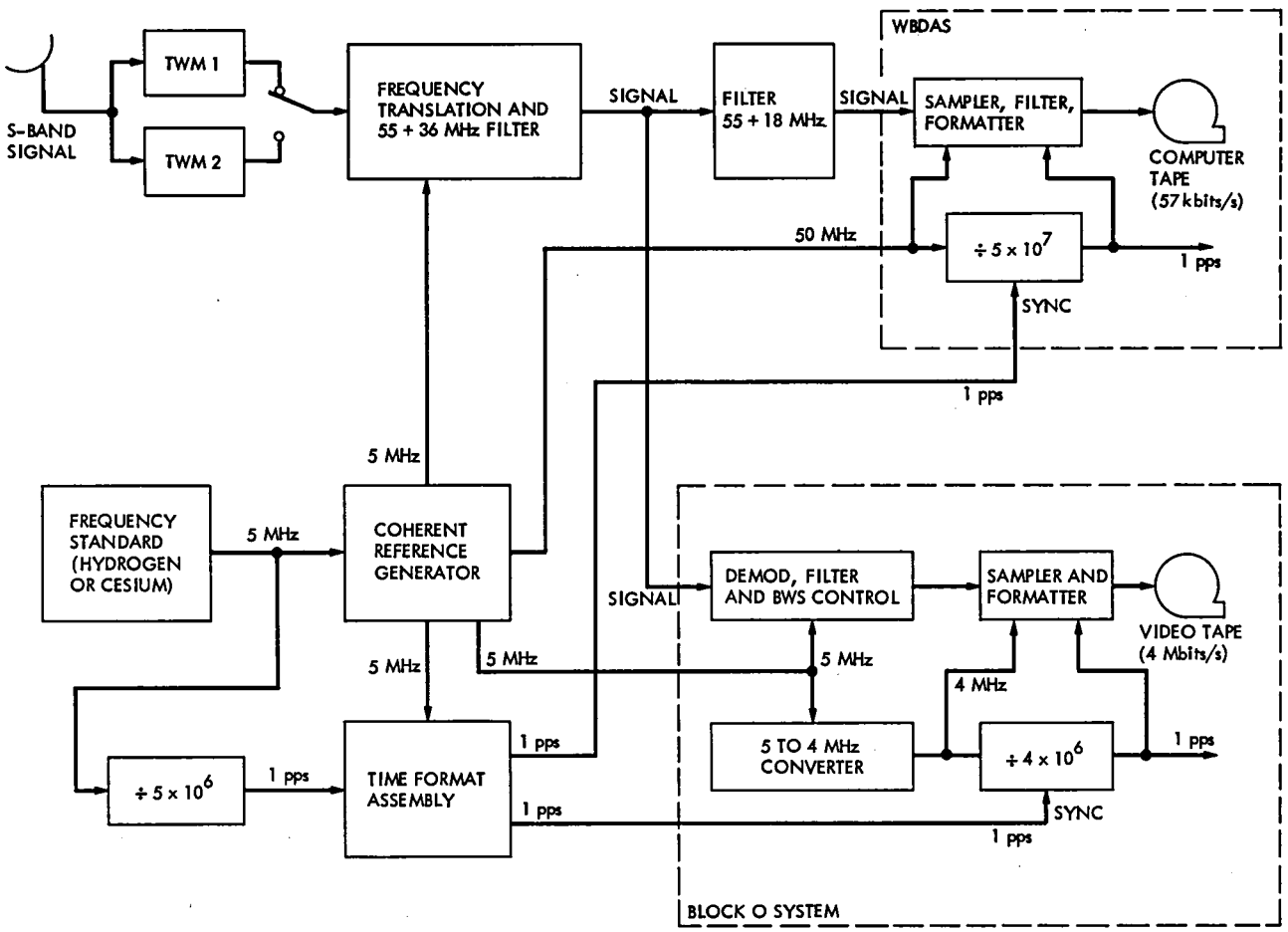


Fig. 1. Configuration of VLBI systems in Deep Space Station

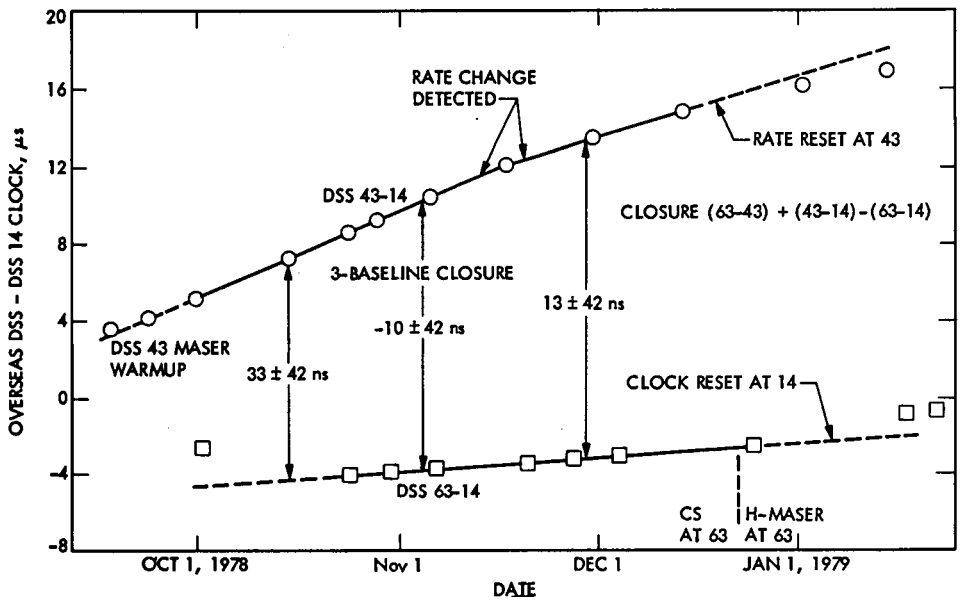


Fig. 2. Clock offsets and closure

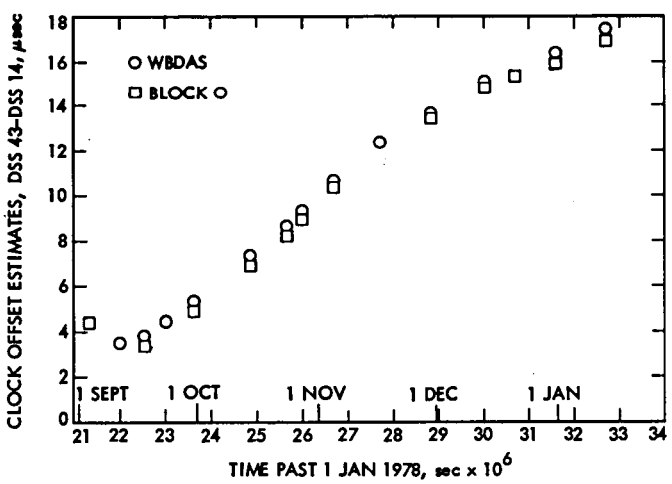


Fig. 3. Clock offset estimates, DSS 43 minus DSS 14

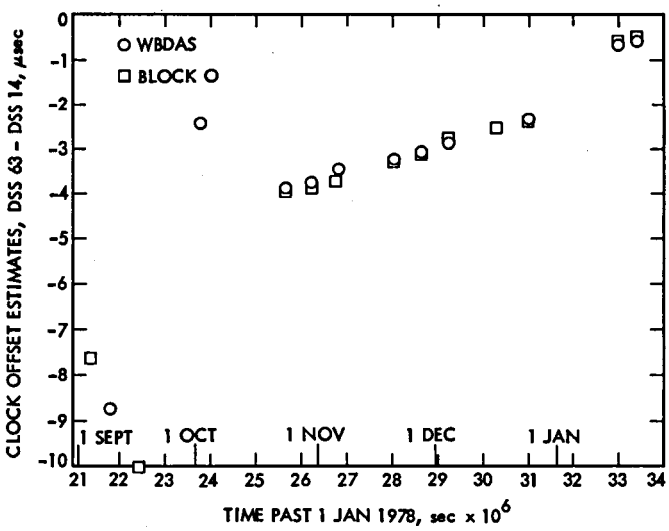


Fig. 4. Clock offset estimates, DSS 63 minus DSS 14

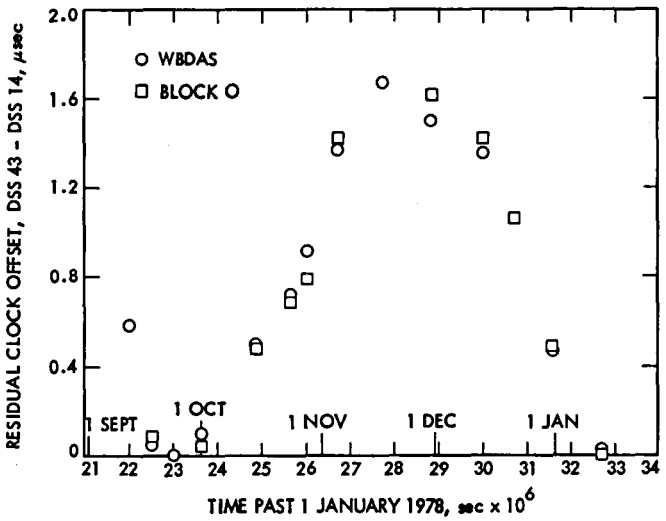


Fig. 5. Residual clock offset estimates, DSS 43 minus DSS 14

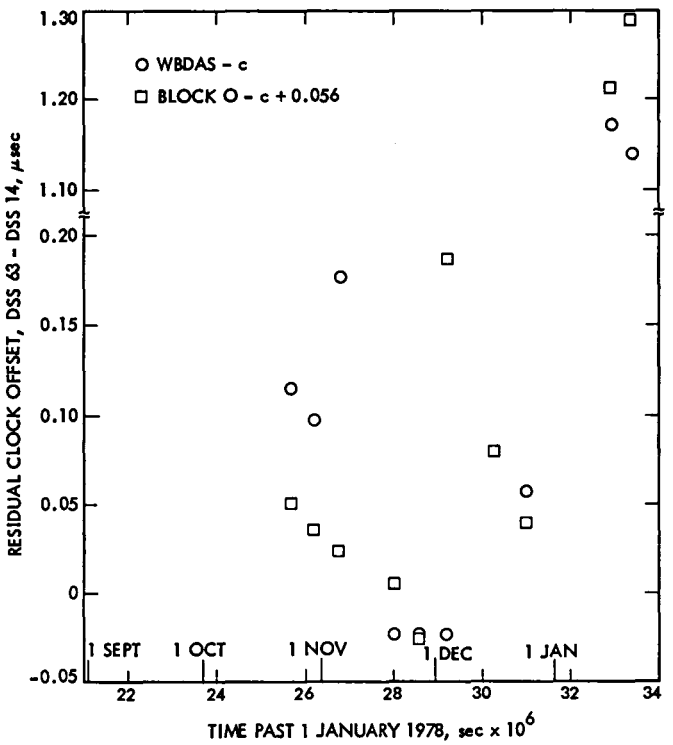


Fig. 6. Residual clock offset estimates, DSS 63 minus DSS 14

Projected State-of-the-Art for Frequency Standards for the DSN in the 1982–1990 Time Frame

W. Higa
Communications Systems Research Section

R. Sydnor
Radio Frequency and Microwave Subsystems Section

A study has been made of frequency standards that are potentially useful to the DSN in the 1980s. A combination of several devices that have inherently good performance over different averaging time regimes is recommended. This combination consists of: a quartz oscillator for short-term, either a quartz oscillator or active hydrogen maser for medium-term, and a passive hydrogen maser or cesium standard for long-term stability.

I. Introduction

Progress in the development of frequency standards over the last three decades has been truly impressive. Over this time the performance has improved by five orders of magnitude from 10^{-9} in 1950 to 10^{-14} today in 1979. It was not easy to foresee this progress in 1950, just as it is not easy to foretell the changes that will occur in the next ten years. However, better knowledge of the physics of atoms, molecules, and crystals is now available, and techniques for utilization of frequency standards are better understood.

For the present and near future, atomic and molecular frequency standards provide the best overall performance. Quartz crystal oscillators and superconducting cavities or a combination of both may be a cost effective means to achieve the required stability over short to medium averaging times.

The purpose of this article is to present an elementary discussion of the physical bases for various kinds of frequency standards.

II. Physical Basis for Modern Frequency Standards

Two recent publications (Refs. 1 and 2) on frequency standards written by H. Hellwig, who was chief of the Time and Frequency Division, NBS, provide excellent summaries of the present state of the art.

Here the similarities and differences between different standards will be explained in a qualitative way so that the nonspecialist can, with the help of Hellwig's publications, become acquainted with the rationale for various devices and techniques.

A. Electromechanical Oscillators

Historically, the quartz crystal oscillator was the first precision oscillator used as a frequency standard. The piezoelectric properties of quartz are used in various ways to sustain stable oscillations. The frequency range is from a few kilohertz to several tens of megahertz and Qs of the order of 10^6 are

readily achieved. A great advantage of the crystal oscillator is the relatively high power level at which it oscillates; this provides the high signal-to-noise ratio needed to obtain stability at short averaging times. The disadvantage of this type of oscillator is its long term drift, which is attributed to aging phenomena in the crystal.

However, recent studies have shown that the aging process can be reduced by operating newly designed crystals at low power levels. Thus, it is possible to use a high-power crystal for good short-term stability, and a low-power crystal for good medium-term stability.

A new kind of electromechanical oscillator has recently been developed that uses the fundamental longitudinal mode of vibration in large single-crystal cylinders of sapphire or silicon. These crystals are not piezoelectric, but are readily excited by electric fields from capacitor plates located near the ends of the cylinder. By cooling the crystal cylinder to cryogenic temperatures (1.8 K), Qs of the order of 8×10^9 have been achieved for resonators at 20 kilohertz (Ref. 3). These resonators were developed for gravity wave detection and have not yet been used in frequency standards. The high Q obtainable is impressive, but unfortunately cannot be achieved at higher frequencies since Q is proportional to the ratio of the length of the cylinder to the surface finish of the end faces, i.e., to the parallelness of the two ends.

B. Superconducting Cavity Stabilized Oscillators

Electromagnetic resonators with high Q can be obtained by the use of superconducting cavities at microwave frequencies. Recent advances in fabrication techniques have resulted in Qs of the order of 10^{10} at X-band and temperatures of 1 to 1.2 K. By using a high-Q cavity as a frequency discriminator, it is possible to phase lock a crystal oscillator to the line center of the cavity. Stabilities of the order of $< 10^{-15}$ have been achieved for averaging times of the order of minutes. They have been measured by JPL on a superconducting cavity stabilized oscillator built by Stanford University. For this kind of fractional frequency stability, a 3-cm-diameter (D) cavity (X-band) must be held to $\Delta D = 3 \times 10^{-15}$ cm. This is two orders of magnitude smaller than the classical electron radius! Such dimensional stabilities require temperature control to within a few micro-Kelvin of the operating temperature. Equally important would be the control of vibration and of the VSWR of all transmission lines connected to the cavity. These challenging problems must all be solved before the superconducting cavity stabilized oscillator can be regarded as more than a laboratory curiosity, except for very short averaging times ($\tau < 300$ sec).

C. Atomic Frequency Standards

1. Devices employing the hyperfine transitions. All the currently available atomic frequency standards employ the hyperfine atomic transitions. Hydrogen, rubidium, and cesium have nonvanishing nuclear moments that give rise to magnetic moments. Each of the atoms has a single valence electron exposed to the weak but precise nuclear magnetic field. The spinning electron in conjunction with the nuclear spin gives rise to energy levels as shown in Fig. 1. Although the energy levels look different for the three atoms, it is noted that they all have in common a pair of levels that, for small external fields, is independent of the field. Transitions between these levels are called sigma transitions and the others are called pi transitions.

It is a simple matter to select the sigma transition over the field dependent pi transitions by a proper orientation of the RF magnetic field relative to a weak reference dc magnetic field.

For high magnetic fields, all energy states are strongly dependent on the magnitude of the fields. This dependence may be used to separate those with positive slopes from those with negative slopes.

In principle, all three may be operated as active standards (masers) or as passive devices. So far, only hydrogen and rubidium have been used in the two modes; cesium has been used only in the passive mode.

Table 1 summarizes the frequencies and line-widths in typical hydrogen masers, cesium standards, and rubidium standards.

In all these standards, neutral atoms in the ground state (the state in which the orbital angular momentum of the valence electron is zero) are employed. This makes it easier to approach the ideal condition for spectroscopic observation of an atom: one which is isolated in space and free from all perturbations. However, because the magnetic moments due to nuclei are weak, the hyperfine transition frequencies are in the lower microwave range as shown in Table 1. It turns out that higher Qs are achievable in atomic transitions for higher energies or higher frequencies. This can be realized, for example, by employing fine structure transitions in atoms as discussed later.

It is important to point out the differences among the standards employing hyperfine transitions. In the hydrogen maser, the transitions are used directly to produce oscillations at 1420 MHz. In the rubidium standard, an optical transition that is coupled to the hyperfine transition is used as a mechanism in an optical frequency discriminator.

Finally, in the cesium standard, a particle detector discriminates between atoms that have made transitions from those that have not.

Active standards are desirable when good short-term stability is necessary. In the hydrogen maser, a flux produces the high signal-to-noise ratio needed for short-term stability at the expense of a lowered Q. However, a high Q is desirable for long-term stability. The recent interest in passive hydrogen standards is due to the higher Q that may be realized by the use of a low hydrogen flux. The National Bureau of Standards employs a cesium standard of several meters in length to achieve a high Q in its standard.

The cesium standard is regarded as a primary standard because the atoms are observed spectroscopically in free flight in a vacuum chamber. Thus, any number of standards could be expected to have the same line center frequency. This is different than the case of the hydrogen maser. In the hydrogen maser, the atoms are stored in a teflon-coated quartz bulb for approximately one second during the emission of microwave energy. Since each storage bulb will affect the atoms in a slightly different way, it will produce a small and different frequency shift (wall shift) in the line center frequency. This is why the hydrogen maser is regarded as a secondary standard.

Similarly, the atoms in a rubidium standard are contained in a cell with a buffer gas, such as argon, when spectroscopic measurements are made to determine the center frequency. Again, no two cells are expected to have identical characteristics, and the rubidium standard is referred to as a secondary standard.

2. Active versus passive frequency standards employing the hyperfine transitions. Many parameters determine the usefulness of a particular atomic species in a frequency standard, not the least important of which is the vapor pressure of the substance at room temperature. It is the differing vapor pressures in the alkali atoms (H, Rb, and Cs) that led to the different configurations in frequency standards. Hydrogen and rubidium have been employed both in masers and in passive standards. To date, cesium has been operated only in the passive mode because of its low vapor pressure.

The passive hydrogen maser shows promise as a frequency standard with good long-term stability. The passive hydrogen maser operates as a frequency discriminator to phase lock a crystal oscillator to the hydrogen resonance at 1420 MHz. The flux of atoms entering the maser cavity is reduced below the level at which oscillation occurs. The device then operates as a frequency discriminator and standard techniques are then used to stabilize a crystal oscillator.

The principal advantage of the passive standard is that reducing the flux reduces the width of the hydrogen line significantly since collision (spin-spin exchange) broadening is reduced. The atomic line Q is thus increased and the cavity pulling effect reduced correspondingly. Since the principal cause for long-term instability is due to cavity-pulling, the long-term stability is improved.

A serious disadvantage of the passive hydrogen standard is the poor performance for short averaging times since the signal-to-noise ratio is reduced appreciably below that in the active maser. The only change made in the maser was a reduction in the atomic flux; hence, the maser still operates as a negative temperature device. However, the noise generated is proportional to the magnitude of the negative temperature, and this may be quite large when the maser stops oscillating. However, the crystal oscillator that is locked to the maser determines the short-term stability that may be made very good by proper design.

The passive hydrogen standard differs from an absorption cell in a subtle way. Since no change except beam intensity was made, the atoms are still emitting microwave power. The emission, however, is far below the level required to overcome cavity losses, and the device may appear to operate like an absorption cell. Figure 2 shows the actual amplitude and phase characteristics of the passive standard as compared with an absorption cell, and demonstrates that the term "passive hydrogen maser" is not a misnomer.

The Q of the cavity may be reduced to reduce the cavity pulling effect even more in the passive maser. A small-sized frequency standard has been made in which a dielectrically loaded cavity is fabricated such that the inner wall of the cavity, when coated with teflon, forms the storage bulb. The practical problem here is to be able to find a low-loss dielectric with the desired dielectric constant. Alumina, fused quartz, and sapphire are being evaluated.

In the hydrogen maser, a fractional frequency stability of the order of 10^{-15} requires that the cavity dimensions be held constant to a fraction of an Angstrom, which is 10^{-10} meter or roughly the diameter of a hydrogen atom. The passive maser seeks to relax this requirement by an order of magnitude or more.

3. Standards employing atomic fine-structure transitions. Historically, the fine structure in atomic spectra was observed long before hyperfine transitions. The observation was that optical radiation from atoms often consisted of a multiplicity of regularly-spaced spectral lines. Quantum mechanics explained the spectra in terms of the magnetic moment of the electron due to its orbital motion interacting with the

magnetic moment due to electron spin. Since magnetic moments of electrons and protons are inversely proportional to the mass of the respective particles, the magnetic moment for electrons is roughly 1800 times stronger than for protons. Thus the fine structure splitting of spectral lines is in the submillimeter range, and high Qs are achievable.

A working standard employing the fine structure has yet to be realized for several reasons:

(1) The atom selected must have no nuclear magnetic moment lest the sharp resonance line be smeared by hyperfine transitions.

(2) The orbitally excited state must be sustained during the spectroscopic observation to realize the benefits of the high Q.

(3) An atomic beam type of device is difficult to implement at the high frequencies involved because the waveguides and cavities are extremely small and necessitate quasi-optical techniques.

(4) The orbitally excited state must have a sufficiently long lifetime so as to appear to be in a stationary state during the spectroscopic measurement. Such metastable states in atoms are found only through difficult laboratory investigations. Magnesium and calcium are being studied for this application.

4. Trapped ions. As mentioned in the discussion of hyperfine transitions, the resonant frequencies are in the microwave region due to the weak nuclear magnetic moments of the light nuclei in which nuclear spins cancel in pairs, and it is necessary to find atoms with an odd number of nuclei so that one unit of spin will be uncancelled. As the nucleus gets heavier, however, it is possible for more than one unit of spin to remain uncancelled if such a configuration turns out to have a lower energy. Such an atom with a single valence electron would behave like hydrogen, rubidium, and cesium, and might give a good atomic frequency standard because of the high Q.

The trapped ion technique creates an atom to meet these requirements. To illustrate with a specific example, mercury (Hg 199) has two valence electrons with a ground state of spin zero. If one of the electrons is removed, there results an ion with a single valence electron and a nucleus with spin 1/2; it has the hyperfine characteristics of atomic hydrogen, except that the zero-field splitting is 40 GHz rather than 1.4 GHz. An electrostatic field of the proper size and shape is used to trap the ion in a microwave spectroscope. The trapping of an ion is difficult and considerable effort will be required before a frequency standard can be realized (Ref. 4).

III. Conclusion

From the preceding, one may determine that within the next five years short-term stability will be most readily achieved with a quartz oscillator; medium-term by a rubidium vapor passive standard, a quartz oscillator, or a hydrogen maser; and long-term stability by a passive hydrogen maser, or a cesium standard. There are, of course, unknowns in this evaluation. An active hydrogen maser at high flux still out-performs in short term the best quartz oscillator; it may, in its cryogenic form, be better in the long term than the passive hydrogen maser.

The cryogenic systems (i.e., the superconducting cavity oscillator, the cryogenic hydrogen maser) are as yet unproved, untested, and uncertain. The known best short-term stability is available from the superconducting cavity stabilized oscillator, the hydrogen maser, and the quartz oscillator. Medium-term stability, in most oscillator systems, is at best a compromise between long- and short-term optimized oscillators. The most likely candidates at the moment appear to be the hydrogen maser and (with some uncertainty) the quartz oscillator and the superconducting cavity.

Long-term stability is an ephemeral, difficult achievement reached by selection of oscillating systems with very little aging and environmental influences and careful engineering to isolate them from the outside world. Because of the times required to test and develop such standards, very little can be forecast about future performance. The best presently available unit is the cesium standard; the most promising candidate among units being developed is the passive hydrogen maser; a possible future candidate may be the cryogenic hydrogen maser.

Combining several oscillators with different regimes of optimum performance is a straightforward systems problem. One example, which would produce the best possible frequency standard system using presently available standards, would combine a superconducting cavity in the $\tau < 300$-sec area, an active hydrogen maser for the $300 < \tau < 30000$-sec regime, and a cesium standard for the $\tau > 30000$-sec regime. Such a system is shown in Fig. 3a. The performance (hypothetical) for the three standards in Fig. 3b is combined in this system to produce the heavy curve, the best of all possible frequency standards, for the nonce.

References

1. Hellwig, H., "Frequency Standards and Clocks: A Tutorial Introduction" *NBS Technical Note 616 (2nd Revision)*, June 1977. U.S. Government Printing Office, Washington, D. C. 20402.

2. Hellwig, H., "Atomic Frequency Standards: A Survey," *Proceedings of IEEE*, Vol. 63, No. 2, February 1975, pp. 212-229.

3. McGuigan, D. F., and D. H. Douglass, "Clocks Based Upon High Mechanical Q Single Crystals," in *Proc. 31st Annual Symposium on Frequency Control*, 1977, pp. 616-619.

4. McGuire, M. D., "The Trapped Mercury Ion Frequency Standard," *In Proc. 31st Annual Symposium on Frequency Control*, 1977, pp. 612-615.

Table 1. Characteristics of hydrogen, cesium, and rubidium atomic frequency standards[a]

Element	Nuclear spin	Electron spin	Hyperfine separation resonance frequency, Hz	Second-order magnetic-field dependence, Hz, where H is in oersteds	Typical Qs in frequency standards
Hydrogen 1	1/2	1/2	1,420,405,751.768	2750 H^2	10^9
Rubidium 87	3/2	1/2	6,834,682,605	574 H^2	10^7
Cesium 133	7/2	1/2	9,192,631,770	427 H^2	10^7 to 3×10^{8}[b]

[a]Data was extracted from references (2) and (5).

[b]The Q for cesium standards is for portable devices (10^7) and large laboratory (e.g., NBS, standards (3×10^8)).

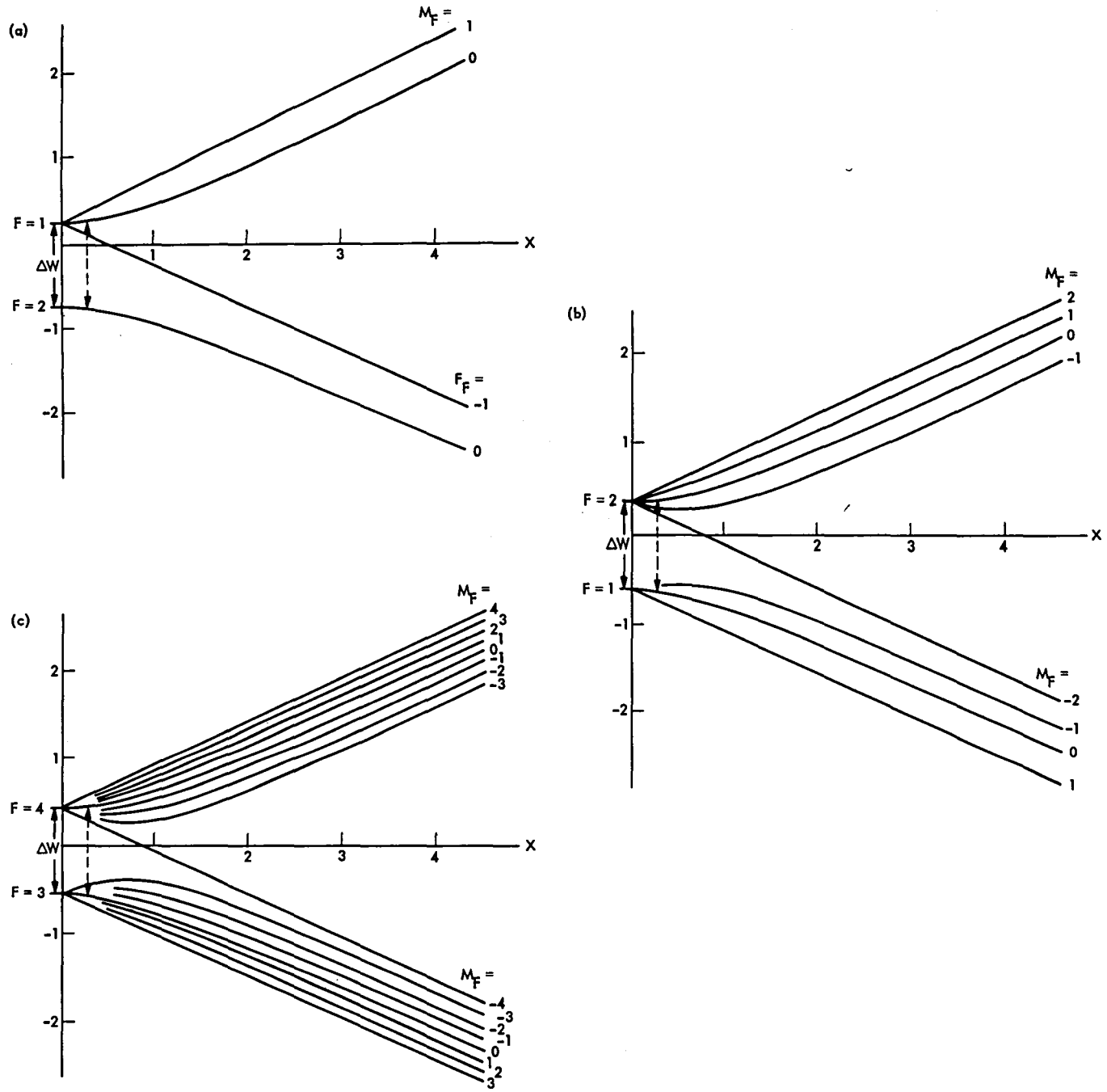


Fig. 1. Hyperfine structure: (a) atomic hydrogen, with nuclear spin $I = 1/2$, $\nu_0 = \Delta W/h$ = 1,420,405,751 Hz; (b) ^{87}Rb, with nuclear spin $I = 3/2$, $\nu_0 = \Delta W/h$ = 6,834,682,605 Hz; (c) ^{133}Cs, with nuclear spin $I = 7/2$, $\nu_0 = \Delta W/h$ = 9,192, 631, 770 Hz

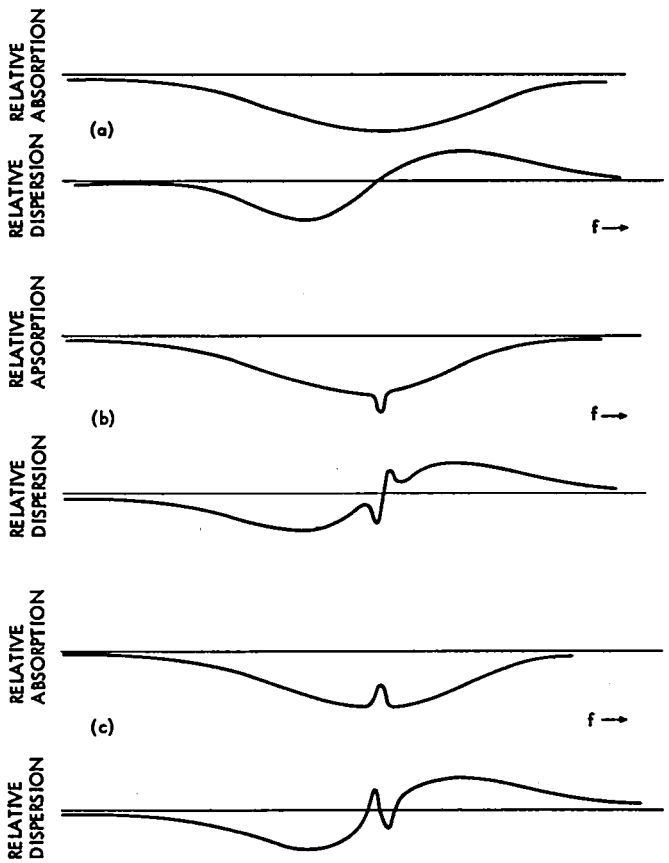


Fig. 2. Absorption and dispersion: (a) cavity only; (b) cavity with absorbing atoms; (c) cavity with emitting atoms

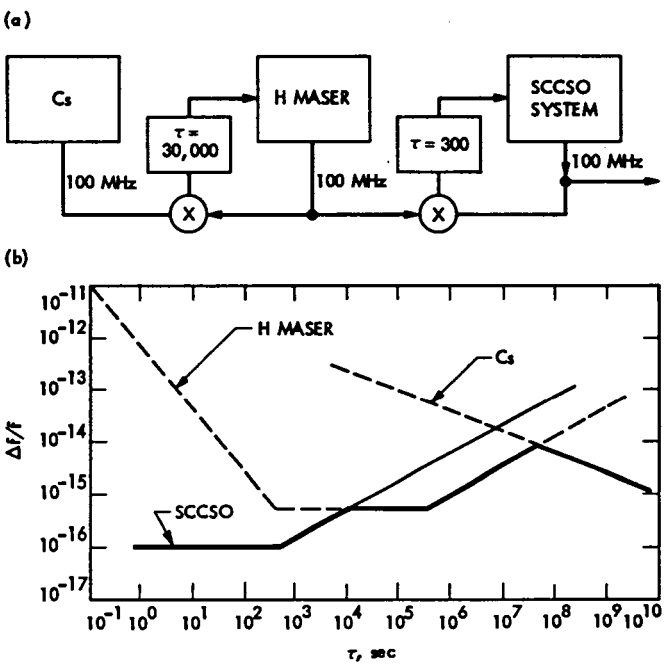


Fig. 3. Combining system for frequency standards: (a) block diagram; (b) performance of individual standards and of combining system

The R_0-Parameter for Optical Communication Using Photon Counting

R. J. McEliece
Communications Systems Research Section

We show that even under ideal conditions (no thermal noise or dark current, continuously-variable intensity lasers, perfectly accurate photon counters), R_o is, at most, one nat per photon in optical systems which use photon counting techniques. Since even under less ideal circumstances channel capacity is infinite, this is a surprising and in some ways disappointing result.

I. Introduction

In a recent report (Ref. 1), it was shown that the R_0-parameter for the noiseless optical channel with pulse/no pulse modulation is exactly one nat per photon. Since R_0 is widely believed to measure the maximum rate at which "practical" reliable communication is possible, and since 1 nat/photon is disappointingly small,[1] it is natural to ask what R_o would be if more general modulation schemes were used. In this article we will show that even with infinitely variable amplitude-modulation and perfectly accurate photon counters, R_0 remains equal to one nat per photon. This result supports Pierce's (Ref. 2) judgement that, "practically, the rates we can attain by photon counting will be limited by how elaborate codes we can instrument rather than by thermal photons."

In Section II, we will give a definition of R_0 which applies to any memoryless time-discrete channel whose output alphabet is finite or countable. Also, we will give a simple upper bound on R_0 which applies to any such channel. Then, in Section III, we will use this bound as applied to the specific channel model appropriate for noiseless optical communication to show that $R_0 \leqslant 1$ nat per photon. Finally we will argue that, in fact, $R_0 = 1$ nat per photon.

[1] Disappointing (and very surprising) since the channel capacity, i.e., the maximum rate at which reliable communication is possible (questions of practicality aside) is infinite! See (Ref. 1) or (Ref. 2) for a proof of this fact.

II. The R_0-Parameter for a General Channel

Consider a time-discrete memoryless channel with input alphabet A and output alphabet B. We assume B is finite or countable. For $x \epsilon A, y \epsilon B$, we denote the probability that y will be received, given that x is transmitted, by $p(y|x)$.

For each pair of input letters x_1, x_2 we define the Bhattacharyya distance between them as

$$d_B(x_1,x_2) = -\log \sum_{y \epsilon B} \sqrt{p(y|x_1)p(y|x_2)}\,. \qquad (1)$$

If now X is a random variable taking values in the set A, and if X_1, X_2 are independent random variables, both with the same distribution as X, we define

$$R_0(X) = -\log E(\exp - d_B(X_1,X_2))\,. \qquad (2)$$

Finally, the quantity R_0 is defined as:

$$R_o = \sup_X R_0(X), \tag{3}$$

the supremum in Eq. (3) being taken over all possible probability distributions on the set A (see Ref. 1, p. 68 and Ref. 3).

The quantity R_0 has dimensions nats[1] per channel use, and is second only to channel capacity itself as a measure of the channel's capabilities. In particular, it is widely believed to be the rate beyond which the implementation of reliable communication systems become extremely difficult (Ref. 4).

We conclude this section with a simple and useful upper bound on R_0. Since the function $f(t) = e^{-t}$ is convex, it follows from Jensen's inequality (Ref. 1, Appendix B) that $E(\exp - d) \geqslant \exp - E(d)$, and hence from Eq. (2) that

$$R_0(X) \leqslant E(d_B(X_1, X_2)) \tag{4}$$

$$R_0 \leqslant \sup_X E(d_B(X_1, X_2)) \tag{5}$$

III. R_0 = 1 Nat/Photon for Optical Channels

We assume that our optical communication system works as follows. The time interval during which communication takes place is divided into many small intervals ("slots") of duration t_0 seconds each. The transmitter is a semiconductor laser which is pulsed during each slot. The intensity of the pulse in the i-th slot is denoted by x_i; this means that the expected number of photons emitted is x_i. The intensity can be any nonnegative real number, but the actual number of photons emitted is, of course, an integer. Because of the Poisson statistics governing photon emissions, the probability that exactly k photons will be emitted by the laser during the i-th slot is $e^{-x_i} x_i^k / k!$ The receiver is a photon counter, which we assume correctly reports the exact number of photons emitted during each slot.

Thus described, the optical channel fits the model of the previous section. The input alphabet A is the set of nonnegative real numbers; the output alphabet B is the set of nonnegative integers; and if $x \epsilon A$ is transmitted, the probability that $k \epsilon B$ is received is

$$p(k|x) = e^{-x} \frac{x^k}{k!} \tag{6}$$

[1] Throughout, all logarithms will be natural.

The first step in computing R_0 for this channel is the computation of the Bhattacharyya distance $d_B(X_1, X_2)$. According to Eqs. (1) and (6)

$$e^{-d_B(x_1,x_2)} = \sum_{k=0}^{\infty} \sqrt{p(k|x_1)p(k|_2)}$$

$$= e^{-(x_1 + x_2)/2} \sum_{x=0}^{\infty} \frac{1}{k!} \sqrt{x_1 x_2}^{\,k}$$

$$= e^{-(x_1 + x_2)/2} \, e^{\sqrt{x_1 x_2}}$$

$$= e^{-(\sqrt{x_1} - \sqrt{x_2})^2/2} \tag{7}$$

Hence

$$d_B(x_1, x_2) = (\sqrt{x_1} - \sqrt{x_2})^2/2 \tag{8}$$

Note also that if we take only the term $k = 0$ in the sum in Eq. (7), we get the estimate

$$d_B(x_1, x_2) \leqslant (x_1 + x_2)/2 \tag{9}$$

It thus follows immediately from the bound Eq. (4) that

$$R_0(X) \leqslant E(X) \tag{10}$$

In words, Eq. (10) says that if the average laser intensity is β photons per slot, then the R_0-parameter is at most β nats per slot. In units of nats per photon, then, we have

$$R_0 \leqslant 1 \text{ nat/photon} \tag{11}$$

Note that the bound Eq. (11) was derived under very generous assumptions about the kind of signalling equipment available (infinitely variable laser intensity, perfectly accurate photon counters). We shall now show that, in fact, R_0 = 1 nat/photon, even if only two laser intensities ("on" and "off") are available, and if we replace the ultrasensitive photon counter with a simple photon detector, which emits a 1 if it is struck by one or more photons in a given time slot, and a 0 if it is not. An indirect proof that R_0 = 1 in this situation was given in Ref. 1. Here we will give a different proof, using q-ary pulse position modulation.

The idea is to select a fixed positive integer q, and to divide the transmission interval into consecutive blocks of q slots each. In each such block, the laser is pulsed exactly once, so that there are exactly q basic patterns in the signalling alphabet. For example, with $q = 4$, if we denote "no pulse" by 0 and "pulse" by 1, these patterns are 1000,0100,0010,0001. There are, however, $q + 1$ possibilities for the received pattern, because the laser may emit no photons when it is pulsed. The probability that a given transmitted pattern will be received in error is just the probability that the laser will emit no photons during a single pulse: $e^{-\lambda}$, if the laser's intensity is λ.

Thus, the appropriate channel model for this situation has input alphabet (illustrated for $q = 4$) $A = \{1000,0100,0010, 0001\}$ and output alphabet $B = \{0000,1000,0100,0010,0001\}$. The transition probabilities are

$$p(y|x) = 1 - e^{-\lambda} \quad \text{if } x = y$$

$$= e^{-\lambda} \quad \text{if } y = 0000$$

$$= 0 \quad \text{otherwise}$$

From this it easily follows that the Bhattacharyya distances are given by

$$d_B(x_1, x_2) = 0 \text{ if } x_1 = x_2$$

$$= \lambda \text{ if } x_1 \neq x_2 \tag{12}$$

Hence by Eq. (4), we have

$$R_0(X) \leqslant \lambda \cdot P\{X_1 \neq X_2\} \tag{13}$$

the units in Eq. (13) being nats per channel use. It is easy to see that $P\{X_1 \neq X_2\} \leqslant (q - 1)/q$, with equality if and only if X is equally likely to be any of the q channel input symbols. Thus,

$$R_0(X) \leqslant \lambda \frac{q-1}{q} \quad \text{(nats per channel use)} \tag{14}$$

or, since each channel use requires λ photons on the average,

$$R_0 \leqslant \frac{q-1}{q} \text{ nats per photon.} \tag{15}$$

On the other hand, if X is uniformly distributed on the input alphabet, a simple calculation gives

$$R_0(X) = -\frac{1}{\lambda}\log\left(\frac{1}{q} + \frac{q-1}{q}(e^{-\lambda})\right) \text{ nats/photon.} \tag{16}$$

The limit of Eq. (16) as $\lambda \to 0$ is easily seen to be $(q - 1)/q$, and so we conclude that for q-ary pulse-position modulation,

$$R_0 = \frac{q-1}{q} \text{ nats/photon.} \tag{17}$$

(For completeness, we remark that a similar calculation shows that the capacity for q-*PPM* is

$$R_0 = \log q \text{ nats/photon.}) \tag{18}$$

Equation (17) shows that by taking q sufficiently large, R_0 can be made as close to 1 as desired. This fact, combined with Eq. (11), shows that $R_0 = 1$, as claimed.

We conclude with two remarks. First, a close examination of our calculations shows that the only possible input distributions that approach $R_0 = 1$ have both average and peak intensity very close to zero. This suggests that efficient coding schemes will have the same property.

Second, note that the bound Eq. (9) applies even if we allow pulsing at different frequencies, since Eq. (9) merely reflects the ambiguity at the receiver if no photons are received. Thus, the bound $R_0 \leqslant 1$ holds even for frequency-modulated direct-detection systems. Of course if we use different frequencies the number of nats per photon is no longer proportional to the number of nats per joule, which is of course the basic unit here. However, a multifrequency system operating at R nats per photon would consume more energy per transmitted nat than a monochromatic system using the lowest frequency of the multifrequency system operating at R nats per photon. And since $R_0 = 1$ for both systems, we would expect monochromatic systems to be more efficient.

References

1. McEliece, R. J. and Welch, L. R., "Coding for Optical Channels with Photon-Counting" *The Deep Space Network Progress Report* Vol. 42-52, Jet Propulsion Laboratory, Pasadena, Calif., July-August, 1979.

2. Pierce, J. R., "Optical Channels: Practical Limits with Photon Counting," *IEEE Trans. Communications*, COM-26(1978) pp. 1819-1821.

3. Omura, J. and Viterbi, A., *Digital Communications and Coding*, New York: McGraw-Hill, 1979.

4. Massey, J. L., "Coding and Modulation in Digital Communications." Proc. 1974 Int'l. Zurich Seminar on Digital Communications.

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

Density and Reliability Predictions for a General Logic Structure for Custom LSI

M. Sievers
Communications Systems Research Section

A general logic structure (GLS) for implementing arbitrary functions in integrated circuits has been described in a previous report. Density and reliability predictions for the GLS will be presented in this article. The GLS has been found to be more dense than programmed logic arrays (PLA) and certain configurations of "optimized" macros. Macro is used here to mean a predefined function that may be inserted into a design.

A reliability model is presented that includes the possibility of undetected manufacturing flaws. This model is more accurate than models that consider only so-called wear-out failures. It may be used to indicate how much preinstallation test coverage is necessary to guarantee a given installed reliability.

I. Introduction

A general logic structure (GLS) for constructing integrated circuits has been reported on in Ref. 1. The GLS is a two-dimensional array of wires into which logic functions are mask programmed. A stick diagram of an unprogrammed NMOS GLS is shown in Fig. 1. The GLS consists of metal columns (dot-dashed lines), grounded diffusion rows (dashed lines), and polysilicon implicant rows (solid lines). Metal column density in Fig. 1 is twice that reported in Ref. 1. Alternate columns are used to build gates, the other columns are used to carry signals or power.

Figure 2 shows the stick diagram of a programmed NMOS GLS. Enhancement mode transistors are created at the intersection of polysilicon and a diffused region. The polysilicon becomes the transistor gate and the diffusion forms the contacts to the channel. A NOR logic gate is created by connecting one or more transistor channels between a gate column and ground. When any of the transistors conduct, the gate output is zero. Pull-up resistors pull a gate column to the positive supply voltage when no pull-down transistors are conducting. The pull-up transistors are made from depletion mode transistors. Depletion mode transistors are formed by ion implanting the transistor channel prior to the polysilicon and diffusion steps. The GLS is programmed by creation of pull-up and pull-down transistors, cuts, and contact points.

The philosophy behind the GLS structure is to optimize wiring regularity while sacrificing logic regularity if necessary. This approach is different from the usual in which logic is optimized at the expense of wiring. It turns out that wiring area consumes the larger portion of chip area because logic structures can easily be made compact. It can be expected, therefore, that a structure that optimizes wiring area will often

occupy less area than a structure that optimizes logic area. Density estimates for the programmed logic array (PLA), GLS, and various configurations of "optimized" macros are given in Section II.

Section III will deal with reliability estimates for the GLS. A model is presented that includes the possibility of undetected circuit flaws occurring in an installed chip.

It is common to model integrated circuit failure only in terms of the arrival of random wear-out failures such as metal separation due to electromigration. A Poisson arrival is assumed for the failures. Using this model, the probability of k failures occurring in a time interval t is

$$P_k = (\lambda t)^k e^{-\lambda t}/k! \quad (1)$$

where λ is the failure rate. From Eq. (1), the probability that no failures occur in the interval is

$$P_0 = e^{-\lambda t} \quad (2)$$

This model is accurate when chips are 100-percent tested prior to installation. This is not always possible and may not be necessary as long as the consequences of reduced testing can be determined. The model in Section III predicts the effects of less than perfect testing.

II. Density Estimates

A good measure of the efficiency of a chip design is its gate density. In this section, the gate density for three design methods are estimated and compared.

An estimation of the gate area required to place and wire a given number of gates on a chip can be made by slightly recasting the question posed by the wireability analysis in Ref. 2. That article considers the probability of successfully wiring a given number of gates in a given area. The question considered here is: what area is necessary to guarantee the wireability of a given number of gates? The analysis in Ref. 2 is pessimistic and the area estimates computed here based on that paper will be pessimistic.

Before beginning the analysis, a few definitions are required. Define pitch as the separation between gates. The average pitch is denoted by $\overline{R}$. Demand is defined as the length of wire, measured in pitches, required by a given circuit. Finally, define capacity as the total wire length available in a circuit.

Demand, D, is computed by the equation

$$D = \text{No. of gates} * \overline{R} * \text{fan-in/gate} \quad (3)$$

Wireability can be assured if the capacity is twice the demand.

Let the number of gates to be placed on a chip be m. Assume these are arranged in a square pattern with $\sqrt{m}$ gates on a side. Consider the GLS in which f is the fraction of rows and gate columns available for wiring. Vertically there will be

$$V_w = \sqrt{m} + \sqrt{m}f + w_v \quad (4)$$

wires per row where w_v is the number of wires needed vertically to guarantee wireability. If there are r pitches per column, the total number of vertical pitches is

$$V_p = r(m(1+f) + w_v)/2 \quad (5)$$

since there are two vertical wires per pitch. Similarly, the number of horizontal pitches is

$$H_p = c(2\sqrt{m}f + w_h)/4 \quad (6)$$

where w_h is the number of horizontal wires needed for wireability and there are c vertical pitches per row. There are four horizontal wires per horizontal pitch, so the number of rows r is

$$r = \sqrt{m} + w_h/4 \quad (7)$$

and the number of columns c is

$$c = \sqrt{m} + w_v/2 \quad (8)$$

Summing vertical and horizontal capacity, letting $w_v = w_h = w$, and doing some algebra, the number of wires needed in each direction is

$$w = ((8D - 2m(1+f) + (\sqrt{m}/4(7+3f))^2)^{1/2} - \sqrt{m}/4(7+3f) \quad (9)$$

For an NMOS GLS based on the Caltech design rules (Ref. 3), the horizontal pitch is $14\lambda_d$ and the vertical pitch is $36\lambda_d$

where is λ_d the minimum defineable distance. Total area is therefore

$$A = (14(\sqrt{n} + w/2)\,(36(\sqrt{m} + w/4)) \qquad (10)$$

Table 1 tabulates the demand, w, and GLS area for several numbers of gates. The value for $\bar{R}$ was taken from Ref. 2, fan-in is assumed to be 2.5, and f was set to 1/4.

To compute the area requirements for a programmed logic array (PLA) of m gates, consider Fig. 3. The area labeled PU is the pull-up resistor area. Half of the m gates are located in the upper AND plane, and the remaining half are located in the bottom OR plane.

Using the Caltech NMOS design rules, PLA implicant rows are an average of $7\lambda_d$ apart, as are the columns. A depletion mode transistor is $18\lambda_d$ long. All of the wires needed by the PLA are internal to it in the form of large fan-in AND and OR gates. Let the number of input wires be w where

$$w = n(m/2) \qquad (11)$$

and $n > 0$. From Fig. 3, the area is

$$A_{\text{PLA}} = (n(m/2)\,(7) + 18)\,(m/2{*}7) + (m/2{*}7)\,(m/2{*}7 + 18) \qquad (12)$$

Simplifying

$$A_{\text{PLA}} = 12.25nm^2 + 12.25m^2 + 126m \qquad (13)$$

Table 2 compiles PLA areas for several values of n and m.

The final area estimates are based on the structure shown in Fig. 4 and shows a collection of optimized modules wired together in a grid pattern. These modules are very dense and do not permit wires to run through them. The number of input-output terminals, T, in each cluster can be predicted via Rent's rule (Ref. 2):

$$T = Am^{\text{p}} \qquad (14)$$

where A is the fan-in of the internal gates, m is the number of internal gates, and p is a fraction, $1/2 \leqslant p < 1$, that is related to the relative function per pin. For $p = 1/2$, the number of pins per function is small and most of the wiring is internal to the module. As p approaches 1, the number of terminals becomes the sum of the internal gate fan-in and most of the wiring is external to the module. Assume a "tight" design where $p = 1/2$ and there are 2.5 inputs per gate. The required number of terminals is from Eq. (14):

$$T = 2.5\,m^{1/2} \qquad (15)$$

which is tabulated in Table 3.

Let the area occupied by m gates and connecting wires in an optimized module be half of what would be required by the equivalent number of GLS gates excluding wires. Also, assume the optimized modules are square. Table 3 lists the optimized module area for several values of m.

Tables 4 through 9 list demand, and area for various configurations of optimized modules, and gates per module. Area computations are made following the procedure outlined for the GLS.

Area estimates tabulated in Tables 1, 2, 4, and 7 are plotted in Fig. 5. From Fig. 5, it can be seen that the GLS does better than the PLA or clusters of optimized design for a sufficiently large number of gates. Another conclusion that can be made is that modules should be designed so that some number of wires can pass through them. Reducing the number of blocked wire paths reduces the overall chip area.

III. Reliability Estimation

Reliability estimates based on Eq. (1) assume that chips are 100 percent tested prior to installation. This level of testability is usually very costly and is becoming decreasingly possible as the number of devices per chip increases. A reliability model should include the effects of less than perfect chip testing.

Assume that clustered flaws are of such serious nature that they are always 100 percent detectable. The remaining flaws are thus randomly distributed across the surface of the wafer. If the flaw density is N flaws per unit area, then the probability of no flaws in an area A is

$$P_0 = e^{-NA} \qquad (16)$$

The probability of at least one flaw is

$$P(k \geqslant 1) = 1 - e^{-NA} \qquad (17)$$

Let the ratio of active circuit area to total area be denoted r. Then the probability of a random circuit flaw is

$$P_{cf} = r(1 - e^{-NA}) \tag{18}$$

Denote by $\overline{TC}$ the probability that a flaw goes undetected where $\overline{TC}$ is one minus the test coverage. Then the probability of an undetected circuit flaw is

$$P_{ucf} = \overline{TC} P_{cf} \tag{19}$$

and the probability that no such flaw exists is

$$\overline{P_{ucf}} = 1 - P_{ucf} \tag{20}$$

The occurrence of a random undetected circuit flaw may be considered independent of the occurrence of wear-out failures. Therefore, reliability may be computed by the product of the probability of no wear out and the probability of no undetected circuit flaws.

$$R = e^{-\lambda \cdot} \overline{P_{ucf}} \tag{21}$$

Equation (21) may be used to balance chip area, initial testing, r, and various redundancy schemes to achieve the desired reliability.

The reliability of a triple modular redundant (TMR) system built on a single GLS chip is plotted in Fig. 6 for several variations of area and testing. The voter and associated wires are assumed to occupy the area of one module. The ratio r is 0.7 for the GLS. TMR reliability is computed by

$$R_{TMR} = R_v R^3 + \binom{3}{2} R_v R^2 (1 - R) \tag{22}$$

where R_v is the voter plus wire reliability and R is computed by Eq. (21). Voter reliability is assumed equal to module reliability since these occupy the same area.

Testing of a very large chip will probably be done via self-testing techniques. Self-testing need not have 100 percent coverage to be of value. Equation (21) may be used to predict chip reliability for self-testing chips. Totally self-checking schemes such as (Ref. 4) occupy a considerable area. The result is that manufacturing flaws are more likely. A partially self-checking scheme may therefore outperform a totally self-checking scheme.

References

1. Sievers, M. W., "A General Logic Structure for Custom LSI," *DSN Progress Report 42-50*, Jet Propulsion Laboratory, Pasadena, Calif., January and February 1979, pp. 97-105.
2. Heller, W. R., W. F. Mikhail, and W. E. Donath, "Prediction of Wiring Space Requirements for LSI," *Proc. 14th Design Automation Conf. IEEE/ACM*, New Orleans, June 1977.
3. Mead, C., and L. Conway, *Introduction to VLSI Systems*, 1978, ch. 1, 2, text in preparation.
4. Carter, W. C., A. B. Wadia, and D. C. Jessup, "Implementation of Checkable Acyclic Automata by Morphic Boolean Functions," *Symp. Comp. and Automata*, Polytechnic Inst. of Brooklyn, April 1971, pp. 465-482.

Table 1. Characteristics for several numbers of gates

m	$\overline{R}$	D	W	Area, λ_d^2
36	1.387	250	34	168,084
64	1.590	509	49	331,695
100	1.771	886	66	574,308
225	2.117	2382	111	1,518,993
400	2.410	4820	159	2,996,343
900	2.889	13,001	266	7,927,668
1600	3.276	26,208	382	15,775,452

Table 2. PLA area requirements in λ_d^2

m	Area (n = 0.5)	Area (n = 1)	Area (n = 2.5)
36	28,300	36,288	60,102
64	83,328	108,416	183,680
100	196,350	257,600	441,350
225	958,584	1,268,663	2,173,382
400	2,990,400	3,970,400	6,910,400
900	14,997,150	19,958,400	34,842,150

Table 3. Terminal and area requirements of optimized modules

m	T	Area λ_d^2
6	7	1,512
12	9	3,024
18	11	4,536
24	13	6,048
30	14	7,560
36	15	9,072
42	17	10,584
48	18	12,096
60	20	15,120

Table 4. Gates/module = 6

No. modules	Total No. gates	Demand	Area, λ_d^2
4	24	34	60,348
6	36	51	119,286
8	48	68	197,400
12	72	105	422,082
16	96	140	713,136
36	216	350	3,813,756
64	384	713	14,522,760
100	600	1,240	41,890,500

Table 5. Gates/module = 12

No. modules	Total No. gates	Demand	Area, λ_d^2
4	48	44	106,964
6	72	65	205,350
8	96	87	339,836
12	144	135	726,150
16	192	188	1,310,256
36	432	450	6,453,900
64	768	916	24,361,344

Table 6. Gates/module = 18

No. modules	Total No. gates	Demand	Area, λ_d^2
4	72	53	156,392
6	108	80	309,174
8	144	106	504,432
12	216	165	1,083,138
16	288	220	1,823,024
36	648	550	9,632,604
64	1,152	1,120	36,396,096

Table 7. Gates/module = 24

No. modules	Total No. gates	Demand	Area, λ_d^2
4	96	63	217,872
6	144	94	424,536
8	192	130	738,192
12	288	260	1,502,748
16	384	650	2,530,944
36	864	1,323	13,001,024

Table 8. Gates/module = 30

No. modules	Total No. gates	Demand	Area, λ_d^2
4	120	68	259,572
6	180	101	501,126
8	240	135	826,812
12	360	210	1,770,408
16	480	280	2,977,104
36	1,080	700	15,686,484

Table 9. Gates/module = 60

No. modules	Total No. gates	Demand	Area, λ_d^2
4	240	96	517,860
6	360	144	1,013,526
8	480	192	1,667,016
12	720	300	3,596,148
16	960	400	6,050,064

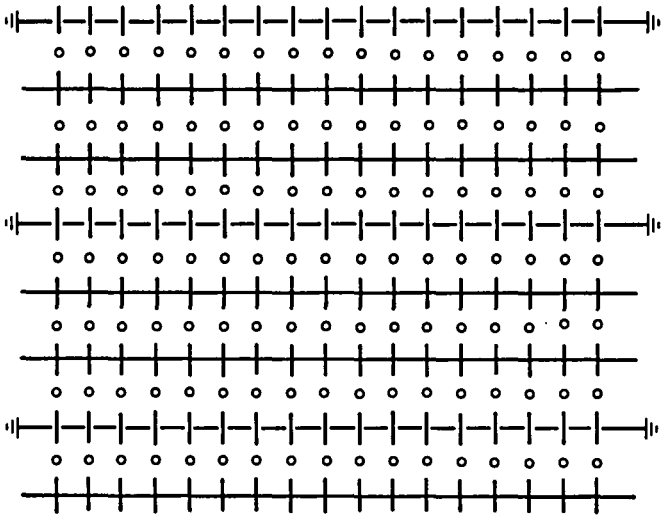


Fig. 1. Unprogrammed NMOS GLS

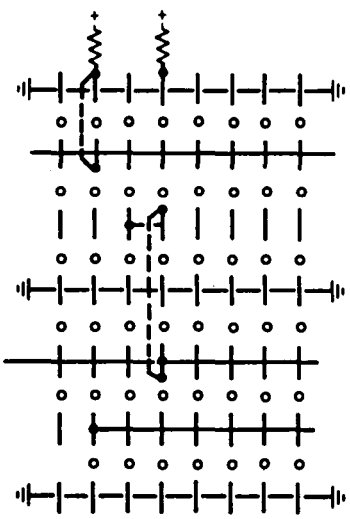


Fig. 2. Programmed NMOS GLS

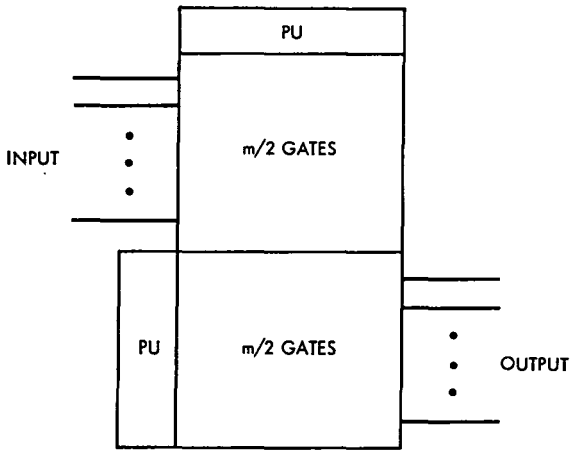


Fig. 3. PLA architecture

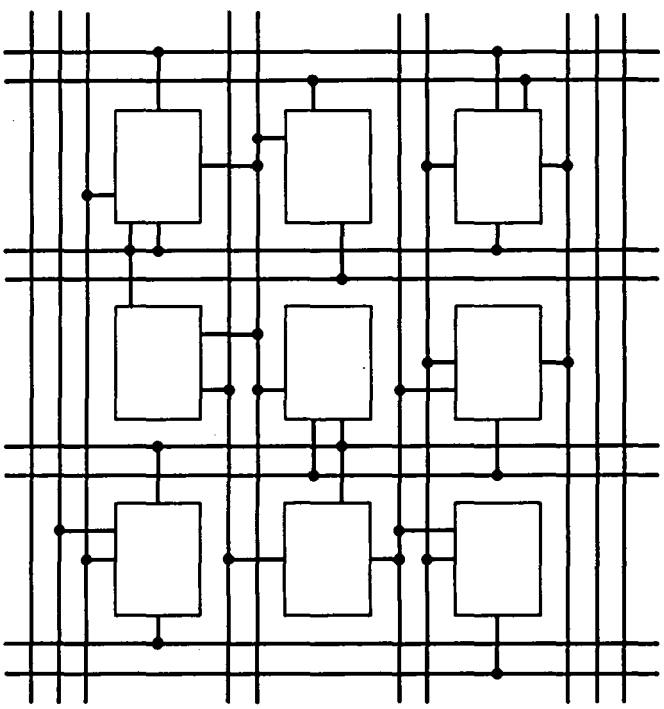


Fig. 4. Optimized macros

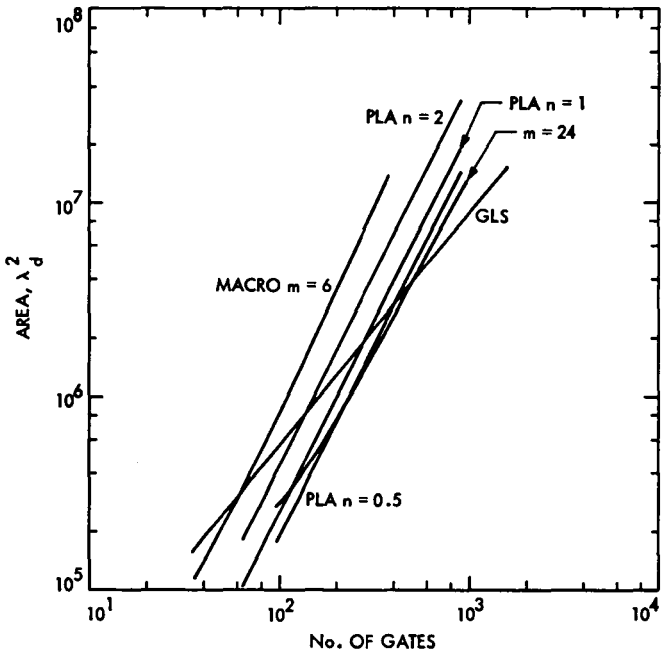


Fig. 5. Area estimates: GLS, PLA, optimized macros

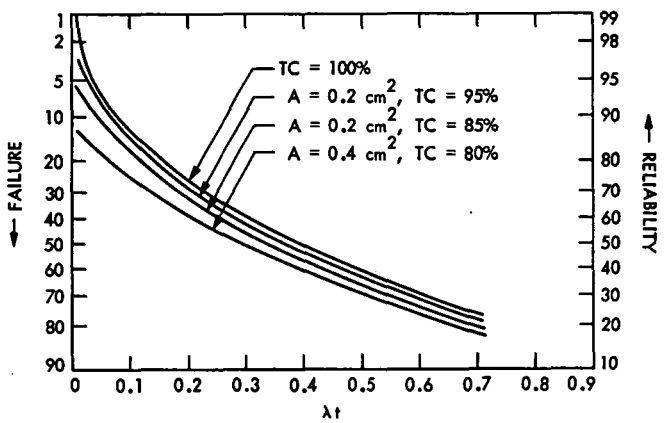


Fig. 6. TMR behavior (N = 2.5 flaws/cm^2)

Frequency Down-Converters as Applied to VLBI

N. C. Ham
Telecommunications Science and Engineering Division

The frequency conversion of a received radio frequency spectrum, from an extra galactic radio source down to video-band frequencies to facilitate data recording or transmission, is quite difficult. A special converter permits this down-conversion while rejecting image-noise "fold over"; however, careful design is required to minimize error contribution to the received information. One possible solution is through the use of combination analog and digital circuitry.

I. Introduction

The frequency down-converters used in very long baseline interferometry (VLBI) systems are necessary to permit recording a segment bandwidth (BW), presently 2 MHz wide or less, from an observed extra-galactic radio source (EGRS) spectrum (Ref. 1). This down-conversion overcomes the incompatibility that exists in obtaining the signal spectrum in a usable form, since a receiving system with adequate sensitivity for detecting extreme weak power levels of the radio sources is at the higher frequency portion of the total spectrum, e.g., S- and X-band, and the magnetic tape recorders, which are available for recording, are in the video band frequencies of approximately 10 kHz to 4 MHz.

The tape recording is a means of obtaining, simultaneously, the radio star signal data at two remotely located receiving stations, which are the terminal ends of a VLBI, and to later bring the two recordings together physically at a processing facility for cross-correlation where the interferometry information is extracted from the radio star observation. An alternate means of getting the data to a common location (to perform the cross-correlation), is to transmit the data via transmission circuits of NASCOM or directly via a satellite. The down-conversion, at the remote locations, into a narrow band of video frequencies is still a necessity, because the data rate or bandwidth limitations of these transmission circuits places a cost constraint on the operations of a VLBI system.

II. The Problem of Image Spectrum

A fundamental problem in down-converting an RF spectrum to the lower frequency bands is that of image or noise "fold over" into the final desired band. For example, Fig. 1 illustrates a typical super-heterodyne receiver converting a spectrum of 2260 to 2300 MHz to an IF band of 260 to 300 MHz. As shown in the figure, the same IF band can be heterodyned by the same local oscillator from an RF band 1700 to 1740 MHz (called the image band). This is undesirable since any signal from the image band causes a signal-to-noise ratio (SNR) degradation. The particular chosen local oscillator and IF has separated the image-band away from the desired band (the relation: image-band is displaced 2 X IF center frequency, e.g., 2280 -(2 X 280) = 1720 MHz) permitting realizable filters to attenuate this image-band.

However, to down-convert a portion of this IF band to video frequencies approaching dc, the above relation of the image-band still exists as shown in the example of Fig. 1b (where a 1-MHz segment BW from 281 to 282 MHz is desired to be downconverted to 3 to 4 MHz) and the resulting "image" band of 274 to 275 MHz cannot be attenuated because of the difficulty of a realizable filter. If the video-band desired is shifted further towards dc, the image band becomes closer to the desired band where finally the image-band is adjacent to the desired band; as shown in Fig. 1c, the image is "folded" into the down-converted dc to video-band range with the second local oscillator situated between these two bands at the IF. This places a burden on realizable filters for rejecting the image band and, hence, other techniques must be utilized.

One technique utilizes the principle of signal cancellation, used for generating single-sideband (SSB) modulation in voice communication application in the reverse order; e.g., a SSB modulator normally takes a baseband signal (in the voice or video range) and modulates a higher RF, and with proper phasing and summation, only the upper sideband is selected (Ref. 2). For the VLBI down-converter application, the "image-band" is cancelled and only the desired band is used. This application has created some semantic problems and expressions such as, "single-sideband demodulator is used to create a base-band signal from the observed radio star spectrum . . ."; such an expression is confusing to a data communication engineer, since the radio star is considered a wide-band noise-like source.

What is really desired is to translate a segment of the noise-like, RF spectrum (of the radio star) down to video-band frequencies without degradation to the received SNR due to "image-noise fold over."

This confusion is extended further by another VLBI application, termed ΔVLBI, which observes alternately the EGRS signal and a spacecraft transmitted radio frequency signal for navigation purposes (Ref. 3). The spacecraft spectrum is either a carrier only (line spectra) signal for narrowband ΔVLBI, or a sinewave modulated carrier signal for wideband ΔVLBI, where the sidebands (also line spectra) are widely spaced about the carrier frequency (typically ±20 MHz).

During the spacecraft signal observation, the requirement is to measure the phase information contained in the extreme sideband frequencies to determine the group delay of the transmitted signal. The method used is to down-convert each sideband and, through low pass filtering, obtain essentially a narrowband width signal, near dc, without image noise fold over.

The procedure is not demodulation to extract the baseband signal (since the baseband signal would be 20 MHz), but the translation of a particular segment of the spacecraft RF spectrum similar to that achieved with the EGRS spectrum.

III. Sideband Spectrum Down Conversion

Figure 2 is a schematic of an analog circuit that achieves the desired down-conversion and functions as follows:

The top multiplier multiplies the input signal spectrum $V_{sig} = V_S \sin(\omega_S \to \omega_{LS})\,t + \sin(\omega_S \to \omega_{US})\,t$ with the local oscillator $V_{LO} \sin \omega_{LO}\,t$. Where $\sin(\omega_S \to \omega_{LS})\,t$ is designated a spectrum band of frequencies from the spectrum center frequency, ω_S, to a lower sideband range ω_{LS} (here sideband is meant the spectrum band on the lower side of the spectrum center frequency), and the second term of the signal as the upper sideband range of frequencies.

The top multiplier output is:

$$V_{TO} = V_S\,[\sin(\omega_S \to \omega_{US})\,t + \sin(\omega_S \to \omega_{LS})\,t]\,[V_{LO} \sin \omega_{LO}\,t]$$

$$= \frac{V_{LO}V_S}{2}\Big\{\cos[(\omega_S \to \omega_{US}) - (\omega_{LO})]\,t - \cos[(\omega_S \to \omega_{US}) + (\omega_{LO})]\,t + \cos[(\omega_S \to \omega_{LS}) - (\omega_{LO})]\,t - \cos[(\omega_S \to \omega_{LS}) + (\omega_{LO})]\,t\Big\} \quad (1)$$

Let $\omega_{LO} = \omega_S$; then $(\omega_S \to \omega_{US}) > \omega_{LO}$ and $(\omega_S \to \omega_{LS}) < \omega_{LO}$. Therefore,

$$V_{TO} = \frac{V_{LO}V_S}{2}\Big\{\cos(0 \to \omega_{US})\,t - \cos(2\omega_S \to \omega_{LS})\,t + \cos(0 \to \omega_{LS})\,t - \cos(2\omega_S \to \omega_{LS})\,t\Big\} \quad (2)$$

It is simple enough to filter out the higher frequency spectrum while preserving the down-converted spectrum, thus:

$$V_{TO} = \frac{kV_S}{2}[\cos(0 \to \omega_{US})\,t + \cos(0 \to \omega_{LS})\,t] \quad (3)$$

The amplitude term is $kV_S/2$ where $V_{LO} > V_S$ by design so that the incremental conductance of the mixer is controlled

as a definite function of time by the local oscillator voltage and has a parameter that varies as a fixed function of time relative to the signal; in other words, the resultant by-product amplitude is proportional to the signal amplitude with a fixed conversion loss accountable by k.

The bottom multiplier output becomes,

$$V_{BO} = V_S [\sin (\omega_S \rightarrow \omega_{US}) t + \sin (\omega_S \rightarrow \omega_{LS}) t]$$

$$[V_{LO} \cos (\omega_{LO}) t]$$

$$= \frac{V_{LO} V_S}{2} \left\{ \sin [(\omega_S \rightarrow \omega_{US}) - (\omega_{LO})] t + \sin \right.$$

$$[(\omega_S \rightarrow \omega_{US}) + (\omega_{LO})] t + \sin [(\omega_S \rightarrow \omega_{LS})$$

$$\left. - (\omega_{LO})] t + \sin [(\omega_S \rightarrow \omega_{LS}) + (\omega_{LO})] t \right\}$$

or

$$V_{BO} = \frac{kV_S}{2} \left\{ \sin (0 \rightarrow \omega_{US}) t + \sin [-(0 \rightarrow \omega_{LS})] t \right\}$$

$$= \frac{kV_S}{2} [\sin (0 \rightarrow \omega_{US}) t - \sin (0 \rightarrow \omega_{LS}) t] \qquad (4)$$

The resultant outputs would represent two sideband spectra "folded-over" each other and converted down to dc, where the edge of the sideband spectrum components are equal or $\omega_{US} = \omega_{LS}$, if they were to be combined now.

But, the top multiplier output signal is instead shifted -90 deg before either being summed or subtracted with the bottom multiplier output signal; hence, Eq. (3) becomes

$$V'_{TO} = \frac{kV_S}{2} [\sin (0 \rightarrow \omega_{US}) t + \sin (0 \rightarrow \omega_{LS}) t] \qquad (5)$$

and Eq. (4) becomes

$$V'_{BO} = \frac{kV_S}{2} [\sin (0 \rightarrow \omega_{US}) t - \sin (0 \rightarrow \omega_{LS}) t] \qquad (6)$$

after passing through the two respective phase shift and delay networks.

Adding the two outputs, $V_{SUM} = kV_S \sin (0 \rightarrow \omega_{US}) t$ is the upper sideband spectrum only. This is followed by a low-pass filter (LPF) whose "break" frequency is equal to ω_{US} and actually sets the resultant spectrum band, or channel bandwidth, that is down-converted down to video band.

In similar fashion the subtracted resultant is

$$V_{SUBTRACTED} = kV_S \sin (0 \rightarrow \omega_{LS}) t$$

and is the lower sideband without the "image fold-over" from the upper sideband spectrum.

An "all-pass" network to achieve a constant -90-deg phase shift with constant amplitude of the down-converted frequencies over the video band is a nonrealizable network, so iterative network sections have been used to synthesize the requirement (Ref. 4).

The resultant amplitude and phase-versus-frequency characteristics of a typical complete down-converter circuit are shown in Fig. 2 and it is seen that the phase is nonlinear versus frequency. When the phase is magnified, and differenced, from an average curve, it results in ripples across the band as shown in Fig. 2d.

The resulting phase ripple of a segment BW (of the total span bandwidth) as it is down-converted through the various channels (minimum of three) to determine the group delay of the radio star signal, is an error contributor to this determination. A wider span bandwidth tends to reduce this effect, and the nature of the signal's continuous type spectrum produces an average phase delay value through each individual channel.

For example, Fig. 3 illustrates how segment bandwidths are obtained from the RF span bandwidth (with an intermediate channel used, not shown, to resolve the ambiguity of the 2π radians/cycle phase information between the extreme end channels) and used in determining the group delay. This technique is called bandwidth synthesis (Ref. 5). The low-end channel results in a net phase delay of -1680 deg comprised of the total accumulated phase shift from the microwave, RF, IF, and down-converter circuits and similarly, the upper end channel results in a phase shift of -320 deg.

The group delay through the receiving subsystem is determined by the difference in the end channel phase shifts divided by the frequency separation of the channels, or

$$\tau_{group} = \frac{\phi_2 - \phi_1}{f_2 - f_1} \times \frac{1}{360} \frac{\text{deg}}{\frac{\text{cycle}}{\text{sec}} \frac{\text{deg}}{\text{cycle}}}, \text{ seconds}$$

For the example shown the group delay is 94 ns. Now if each channel contains only 1-deg peak-to-peak phase ripple (most present down-converter phase shift networks are in the range of 5 to 10 deg peak-peak) and if the upper channel frequency phase terminates on the upper +1 deg and lower end on the lower −1 deg, then the second derivative of the slope results in

$$\sigma_{\tau_g} = \frac{d^2\phi}{dt^2} = \frac{1}{360 \times 40 \times 10^6}$$

and produces 69.4 ps or ≃2 cm equivalent error as a first-order approximation. This is illustrated in Fig. 4.

The continuous spectrum nature of the EGRS signal tends to average out this error and it becomes apparent how a wider span BW and the nature of the signal obviate this error source; however, as the precision of VLBI observation of radio stars is desired or when a line spectra spacecraft signal with doppler effect is observed for ΔVLBI, then it can be seen that the ripple becomes a large error contributor.

This is compounded by the fact that the various channels both at a given station and the complementary station would be hard to match or even to predict or control under environmental and operational conditions.

Various schemes to reduce the phase ripple can be applied, such as using resistance-capacitance bridge circuits and optimum pole-zero pair locations; however, the need for component precision values, matching, and resulting instability are undesirable.

A method that appears feasible in reducing the phase-ripple is to utilize digital techniques to accomplish the second −90-deg phase delay of the preceding circuit. For example, accomplish the first frequency multiplication of the input signal with quadrature components of the local oscillator, as shown in Fig. 2; however, in place of the delay and −90-deg phase shift delay network, following the mixing substitute low-pass filters with a "break-frequency" matched to the sampling frequency, which is used in two following analog-to-digital converters in place of the sum and subtraction circuit.

The LPF is used to reduce "aliasing" by complementing the sampling rate to satisfy the Nyquist criterion, and is chosen to be a higher value than the final LPF segment bandwidth channel value and thus introduce negligible phase error contribution. Also, it filters the higher frequency by products.

The digital signals can be processed to produce the desired −90-deg phase delay digitally where it is more feasible to optimize the pole-zero location and ensure repeatability and stability.

Similarly, after the unwanted sideband image "foldover" is eliminated, a following LPF bandwidth using digital technique determines the final channel BW.

It is expected that the phase ripple resulting from this analog-digital combination can achieve values approaching 1/2 deg peak-to-peak.

IV. Down Conversion of a Monochromatic Source

Consider a typical observation of a monochromatic source such as a spacecraft signal as performed during ΔVLBI, and ignore noise (receiver T_{op}, and galactic background); the received signals at the two stations can be represented as

$$V_1(t) = A_1 \cos 2\pi f_{S_1} t \tag{7}$$

and

$$V_2(t) = A_2 \cos 2\pi f_{S_2}(t - \tau_{geo}) \tag{8}$$

where $A_1 A_2$ is the amplitude of the received signals f_{S_1} and f_{S_2} at antennas 1 and 2, respectively; τ_{geo} is the geometric time delay of the signal wave front to antenna 2 after passing antenna 1, and is related to the orientation of the baseline and source signal direction.

The signal is first heterodyned to an intermediate frequency by the first local oscillator, as shown in Fig. 1:

$$V_{if_1}(t) = [A_1 \cos 2\pi f_{S_1} t]\ [L_1 \cos (2\pi f_{1LO_1} t - \phi_1)]$$

$$= \frac{A_1 L_1}{2} \cos [2\pi (f_{S_1} - f_{1LO_1}) t + \phi_1] \tag{9}$$

where

f_{1LO_1} is the first local oscillator frequency at station 1.

L_1 is the amplitude of the first local oscillator at station 1 and is made larger than the signal amplitude, $L_1 > A_1$.

ϕ_1 represents the phase shift due to the receiver local oscillator and mixer-amplifier.

The band pass characteristics of the IF amplifier rejects the sum term and amplifies the difference term, while the a RF band-pass filter rejected the image frequency, f_{image} (i.e., cos $2\pi f_{image}t \times \cos 2\pi f_{1LO_1} t = f_{if}$) as shown in Fig. 1.

With the local oscillator amplitude being greater than the signal amplitude, the first IF signal becomes

$$V_{if_1}(t) = k_1 A_1 \cos [2\pi f_{if_1} t + \phi_1] \tag{10}$$

where k_1 accounts for the mixer conversion loss and $2\pi f_{if_1} t = 2\pi(f_{S_1} - f_{1LO_1})t$.

For the second station, the first IF signal becomes

$$\begin{aligned} V_{if_2}(t) &= [A_2 \cos 2\pi f_{S_2} (t - \tau_{geo})] \\ &\quad \times [L_2 \cos (2\pi f_{1LO_2} t - \phi_2)] \\ &= \frac{A_2 L_2}{2} \cos 2\pi [f_{S_2} (t - \tau_{geo}) - (f_{1LO_2} t - \phi_2)] \\ &\quad + \frac{A_2 L_2}{2} \cos 2\pi [f_{S_2} (t - \tau_{geo}) + (f_{1LO_2} t - \phi_2)] \\ &= k_2 A_2 \left\{ \cos 2\pi [(f_{S_2} - f_{1LO_2}) t - f_{s_2} \tau_{geo} + \phi_2] \right. \\ &\quad \left. + \cos 2\pi [(f_{S_2} + f_{1LO_2}) t - f_{s_2} \tau_{geo} - \phi_2] \right\} \end{aligned} \tag{11}$$

Following the IF amplifier and bandpass filter it becomes,

$$V_{if_2}(t) = kA_2 \cos 2\pi [f_{if_2} t - f_{S_2} \tau_{geo} + \phi_2] \tag{12}$$

where $2\pi f_S \tau_{geo}$ is the geometric time delay associated with the received radio frequency as shown by Eq. (8) and is typically a slowly varying function of time due to the earth's rotation, or instantaneously is a constant phase in conjunction with the IF signal in Eq. (12).

The final down-conversion to video band is achieved by setting the second local oscillator (see Fig. 5) slightly below the intermediate frequency to avoid shifting the signal to dc and to accommodate doppler shifts. Adopting the notation of Section III, Eq. (1), the first station IF signal, e.g., (10) multiplied by the second LO

$$\begin{aligned} V_O(t) &= kA_1 \left\{ \cos [2\pi (f_{if_1} \rightarrow f_{US}) t + \phi_1] \right. \\ &\quad \left. + [\cos 2\pi (f_{if_1} \rightarrow f_{LS}) t + \phi_1] \right\} \\ &\quad \times [V_{2LO_1} \sin 2\pi f_{2LO_1} t] \end{aligned} \tag{13}$$

where the second local oscillator frequency is less than the IF, $f_{2LO_1} < f_{if_1}$, and the upper sideband edge is greater than the IF, $f_{if_1} < f_{US}$ with the sideband range containing the IF signal, i.e., $(f_{2LO_1} \rightarrow f_{US})$ brackets cos $(2\pi f_{if_1} t + \phi_1)$. See Fig. 5a.

Continuing through the down-conversion process, the summed output then rejects the lower sideband image spectrum and passes the upper spectrum with the resultant video band as

$$V_{v_1}(t) = [\sin 2\pi (o \rightarrow f_{US}) t] \tag{14}$$

which contains the RF signal as,

$$V_{v_1}(t) = A_{v_1} \sin (2\pi f_{v_1} t + \phi_1') \tag{15}$$

where ϕ_1' is the phase delay modified by the down converter circuits (see Fig. 5b).

Similarly, for the down-converted signal from the second station,

$$V_{v_2}(t) = A_{v_2} \sin 2\pi [f_{v_2} t - f_{S_2} \tau_{geo} + \phi_2'] \tag{16}$$

For the EGRS spectrum the same process results except the final down-converted signal is a continuous spectrum from $o \rightarrow f_{US}$, or equivalently a segment of the radio frequency from $f_S \rightarrow f_{US}$ where $f_s' = f_{1LO} + f_{2LO}$ or the sum of the first and second local oscillator frequencies.

These signals are later cross-correlated together (Ref. 5) and processed further with the delay τ_{geo}, and delay rate $\dot{\tau}_{geo}$, the basic VLBI information data.

References

1. J. G. Williams, "Very Long Baseline Interferometry and Its Sensitivity to Geophysical and Astronomical Effects," *Space Summary 37-62*, Vol II, Jet Propulsion Laboratory, Pasadena, Calif., March 31, 1970, pp. 49-55.

2. H. S. Black, *Modulation Theory*, D. Van Nostrand Company, Inc., New York.

3. D. L. Brunn, et al., "ΔVLBI Spacecraft Tracking System Demonstration: Part I. Design and Planning. *DSN Progress Report 42-45*, Jet Propulsion Laboratory, Pasadena, Calif., March-April 1978, pp. 111-132.

4. A. E. E. Rogers, "Broad-band Passive 90° RC Hybrid with Low Component Sensitivity for use in the Video Range of Frequencies," *Proc. IEEE (Letters)*, Vol 59, pp. 1617-1618, November 1971.

5. J. I. Molinder, "A Tutorial Introduction to Very Long Baseline Interferometry (VLBI) Using Bandwidth Synthesis," *DSN Progress Report 42-46,* Jet Propulsion Laboratory, Pasadena, Calif., May-June 1978, pp. 16-28.

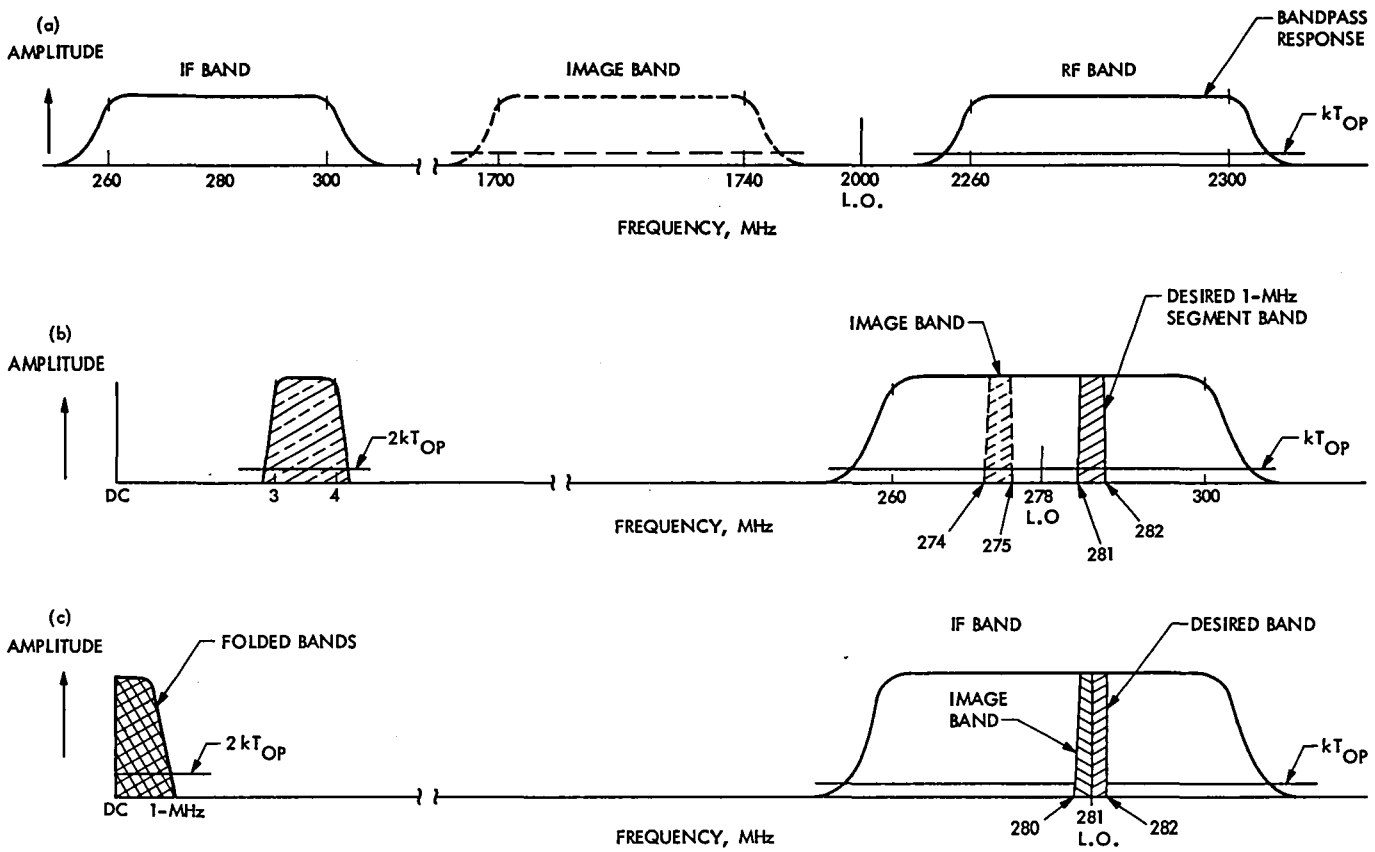


Fig. 1. Image band frequencies resulting from frequency heterodyning, down-conversion: (a) first conversion with rejected image band; (b) second conversion with nonrejected image because of nonrealizable filter; (c) limiting condition where band is to dc when L.O. frequency equals RF band center

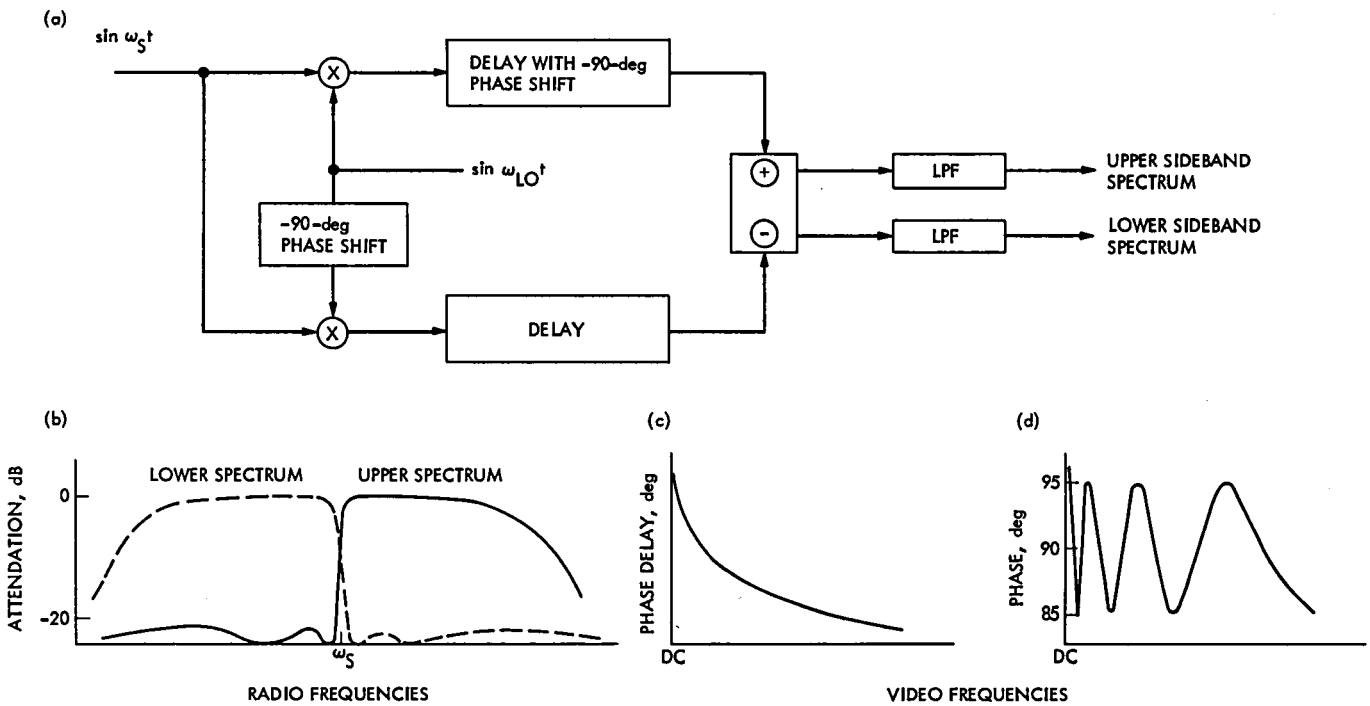


Fig. 2. Frequency down-converter and characteristics: (a) analog circuit; (b) equivalent amplitude response at RF; (c) phase response at video frequency; (d) phase ripple response

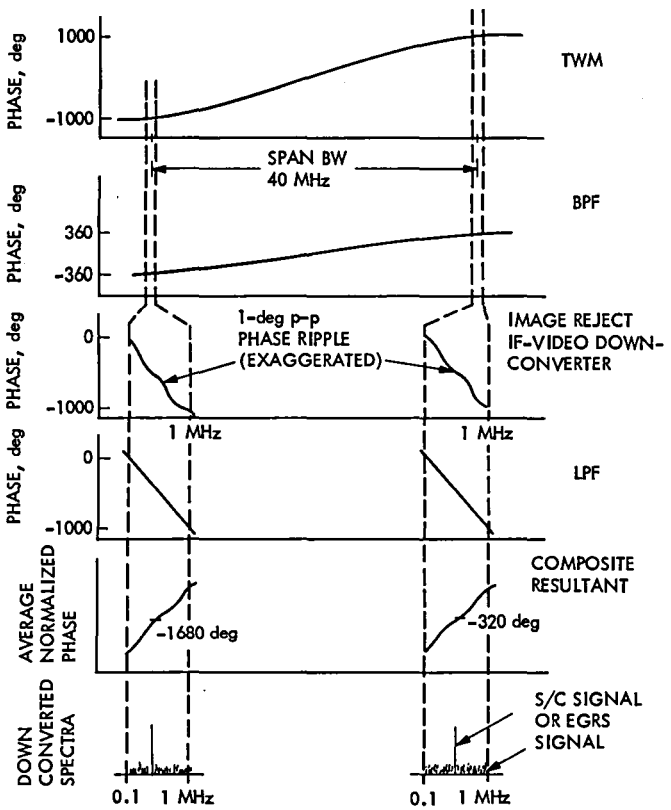
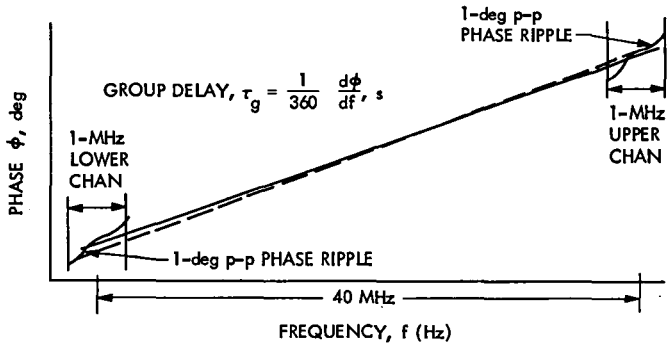


Fig. 3. Resultant down-converted channels



ESTIMATED ERROR IN TIME DELAY MEASUREMENT:

$$\sigma_{\tau_g} = \frac{d^2\phi}{df^2} = \frac{1}{360 \times 40 \times 10^6} = 69.4 \text{ ps} \simeq 2 \text{ cm}$$

= GROUP DELAY ERROR PER STATION (USING ASSUMED VALUE EXAMPLE)

Fig. 4. Simplified group delay error evaluation

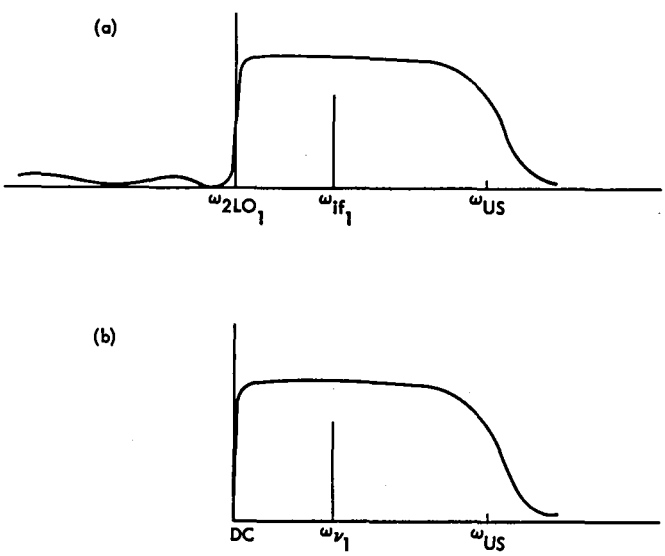


Fig. 5. Down-conversion of a spacecraft line spectra signal: (a) Equivalent amplitude bandpass response at IF where second L.O. is set to less than signal frequency; (b) down-converted signal to video band with resultant RF signal greater than dc and low-pass filter response

The NOCC Radio Science Subsystem – System Performance Validation for the DSN Radio Science and Very Long Baseline Interferometry (VLBI) Systems

A. L. Berman
TDA Engineering Office

The DSN Radio Science and Very Long Baseline Interferometry (VLBI) Systems are functionally similar in that both are concerned with the recording of the baseband output signals of open-loop receivers on either computer-driven magnetic tape units or high-rate digital magnetic tape recorders. In recognition of these similarities, a single Network Operations Control Center (NOCC) subsystem – the NOCC Radio Science Subsystem (NRS) – is being implemented to provide the system performance validation capabilities for both the Radio Science and VLBI Systems.

This article provides a functional description of the key characteristics, requirements, and operation of the NOCC Radio Science Subsystem.

I. Introduction

In 1977, the Deep Space Network (DSN) created the DSN Radio Science System, and in 1978, the DSN Very Long Baseline Interferometry (VLBI) System. These two DSN Data Systems are quite similar in that both are primarily concerned with the recording of the baseband output signals of open-loop receivers on either computer-driven magnetic tape units or high-rate digital magnetic tape recorders. In fact, much of the equipment at the Deep Space Stations (DSS) is shared between the two Systems (Ref. 1).

In general, all DSN Data Systems have a dedicated subsystem in the Network Operations Control Center (NOCC) which provides real-time system performance validation. Besides the already mentioned similarities between the Radio Science and VLBI Systems, both also operate at irregular and somewhat infrequent intervals, in contrast to the nearly continuous operation of other DSN Data Systems. Given the basic similarities between, and the irregular and relatively infrequent operation of, the Radio Science and VLBI Systems, it seemed both convenient and economical to implement a single NOCC subsystem which could provide system performance validation for both Systems. Such a subsystem is in the process of being implemented, and is referred to as the NOCC Radio Science Subsystem (NRS). Some capabilities now assigned to the NRS have been described in previous issues of *The DSN Progress Report*; these capabilities include radio metric graphics (Ref. 2), open-loop spectrum displays (Ref. 2), and planetary atmosphere-corrected predictions (Ref. 3).

This article provides a description of the key characteristics, functional requirements, and operation of the NOCC Radio Science Subsystem.

II. Functional Description of the NOCC Radio Science Subsystem

A. Definition

The NOCC Radio Science Subsystem, an integral element of both the DSN Radio Science and VLBI Systems, performs the following major functions:

(1) Radio Science Functions

(a) Generates graphical radio metric displays

(b) Generates open-loop spectrum displays from Spectral Signal Indicator (SSI) digital data

(c) Generates DSS Radio Science Subsystem (DRS) status, configuration, and control data displays.

(2) VLBI Functions

(a) Generates VLBI calibration and ancillary data displays

(b) Validates VLBI calibration and ancillary data via operator alarms

(c) Generates DRS status, configuration, and control data/displays.

NOCC Radio Science Subsystem functions and interfaces are illustrated in Fig. 1.

B. Key Characteristics

The key characteristics of the NOCC Radio Science Subsystem are listed below for Radio Science and VLBI applications:

(1) Radio Science Key Characteristics

(a) Generates graphical displays of radio metric data for use in validating and controlling the DSN Radio Science and Tracking Systems during critical phase operations.

(b) Generates open-loop spectrum displays for use in validating and controlling the DSN Radio Science System

(c) Provides displays in both NOCC and project Radio Science areas

(d) Provides up to 12 parameter radio metric (graphical) displays simultaneously

(e) Updates open-loop spectrum displays at five-second intervals

(f) Simultaneously processes both VLBI and Radio Science data.

(2) VLBI Key Characteristics

(a) Provides DRS status, configuration, and control data displays

(b) Provides displays of VLBI calibration and ancillary data in two-station comparative format

(c) Validates VLBI operations by comparing status, configuration, control data, and VLBI calibration and ancillary data to standards and limits

(d) Provides operator alarms for configuration errors, out-of-tolerance data, and NRS malfunctions.

C. Functional Operation

Functions and data flow of the NOCC Radio Science Subsystem are presented in Fig. 2. This figure can be referred to in conjunction with the following descriptions of functional operation.

1. Graphical Display of Radio Metric Data. Various radio metric data are accumulated, processed, and formatted in the Metric Data Assembly (MDA) of the DSS Tracking Subsystem (DTK). These data are then provided to the Communications Monitor and Formatter (CMF) for subsequent high-speed data line (HSDL) transmission to the NOCC Tracking Subsystem (NTK). After additional processing by the NTK, the radio metric data are provided to the NRS. The data are processed into graphical displays and are provided to the NOCC Display Subsystem (NDS) for display in the Network Operations Control Area (NOCA) and project Radio Science area. NOCA operators construct and control the graphical display through operator control instruction (OCI) inputs to the NDS.

2. Graphical Display of Open-Loop Spectral Data. Digital spectral data from the SSI of the DSS Receiver-Exciter Subsystem are transferred to the Occultation Data Assembly (ODA) of the DRS. The ODA formats and transmits the data to the NRS via wideband data line (WBDL). The NRS reconstructs the open-loop spectra and provides such replicated spectra to the NDS for display in the NOCA and project Radio Science areas. NOCA operators control the open-loop spectra displays through OCI inputs to the NDS.

3. Display of DRS Status, Configuration, and Control Data. DSS Radio Science Subsystem status, configuration, and control data during both Radio Science and VLBI operations are formatted by the ODA and passed to the CMF for subsequent HSDL transmission to the NRS. The NRS formats this data into displays and provides the data to NDS for display in the NOCA. NOCA operators control the DRS status, configuration, and control data displays through OCI inputs to the NDS.

4. Display of VLBI Calibration and Ancillary Data. VLBI calibration and ancillary data are accumulated in the MDA. The MDA transfers these data to the ODA for formatting. The ODA provides the formatted data to the CMF for HSDL transmission to the NRS. Within the NRS, VLBI calibration and ancillary data from two DSS are formatted into two-station comparative formats, and provided to the NDS for display in the NOCA. NOCA operators control the VLBI calibration and ancillary data displays through OCI inputs to the NDS.

5. Validation of VLBI Calibration and Ancillary Data. DRS status and configuration data and VLBI calibration and ancillary data are compared in the NRS to standards and limits received from the NDS. Configuration errors and out-of-tolerance data generate alarms, which are passed to the NDS for display in the NOCA. Additionally, the NRS generates alarms for some NRS malfunctions.

6. Display of VLBI Data Accountability. VLBI data accountability and VLBI fringe information are passed to the NRS from the VLBI Processor Subsystem. The NRS formats this data into displays and provides the data to the NDS for display in NOCA. In NOCA, operators control these displays through OCI inputs to the NDS.

7. Provision of Status Data. The NRS collects all alarm messages and passes these to the NOCC Monitor and Control Subsystem (NMC), for inclusion on the Network Performance Record.

III. Functional Requirements of the NOCC Radio Science Subsystem

This section describes the specific functional requirements of the NRS for the time period 1979 to 1983. The fundamental Network Operations requirement levied on the NRS is stated as follows:

> The NRS shall provide the following DSN Radio Science and VLBI System data to the Network Operations Control Center in real-time:
>
> (1) Status
>
> (2) Configuration
>
> (3) Performance
>
> These data shall allow the Network Operations Control Team to:
>
> (1) Evaluate system performance
>
> (2) Effect reasoned real-time control decisions as required.

In addition, the NRS shall possess the general capability to simultaneously process data from either two or three DSS for the following configurations:

(1) Three DSS performing VLBI

(2) Two DSS performing VLBI and one DSS performing Radio Science

(3) Two DSS performing Radio Science

The following subsections A. through H. describe the functional requirements for each of the major functional capabilities provided by the NRS.

A. Functional Requirements for Graphical Display of Radio Metric Data

1. Radio Metric Parameters to be Displayed. The capability shall be provided to display radio metric data parameters at both S- and X-band (where applicable) versus the corresponding data time. The specific radio metric parameters to be displayed are listed in Table 1. The capability shall be provided to display these radio metric parameters with time modulo the full-scale time.

2. Screen Display Capability. Display capability shall be provided as follows:

(1) Full screen

(2) Half-screen vertical (two displays/screen)

(3) Half-screen horizontal (two displays/screen)

(4) Quarter screen (four displays/screen)

The capability shall be provided to display up to 12 parameters simultaneously, consistent with 3-quad-channel capability.

3. Hard-Copy Capability. Hard-copy capability shall be provided in NOCA.

B. Functional Requirements for Graphical Display of Digital SSI Data

The capability shall be provided to replicate, display, update, and distribute the CRT spectrum displays of the SSI.

1. Display Capability

a. Output Points. The number of output points (spectral lines) displayed shall be 400.

b. Amplitude Selection. Full-scale (displayed) amplitude shall be selected from: 18.75 dB; 37.5 dB; 75 dB; 150 dB.

c. Spectral Noise Floor Placement. The spectral noise floor placement within the display shall be selectable.

d. Alphanumeric Notational Information. The following alphanumeric notational information shall be included within each SSI display:

(1) Full-scale amplitude in dB

(2) Full-scale frequency (bandwidth) in kHz

(3) Number of averaged spectra per display

(4) DSS identification (ID)

(5) SSI channel selected

(6) SSI transform size

(7) Maximum power point signal-to-noise ratio (SNR)

(8) Maximum power point frequency

2. Display Update Capability. The capability shall be provided to update the display:

Requirement:	once per 30 seconds
Design goal:	once per 5 seconds

3. Display Distribution and Hard Copy Capability

a. Display Channels. The following SSI display channels shall be provided in NOCC and project Radio Science areas.

(1) Two channels, simultaneous operation

(2) One channel per 64 meter DSS

(3) Project Radio Science area displays shall be slaved to NOCC displays

(4) Channel control capability shall reside with the DSN Operations Chief

b. Hard Copy Capability. SSI display hard-copy capability shall be provided in NOCC.

4. DSS-NOCC Interface. Transmission between the DSS and NOCC shall be via wideband data line. A sample display of SSI digital data to be provided in NOCC is illustrated in Fig. 3.

C. Functional Requirements for Display of DRS Status, Configuration, and Control Data During Radio Science Operations

The NRS shall format and provide to the NDS displays of DRS status, configuration, and control data. A minimum set of these data shall be as follows:

1. DRS Status Display

(1) Program mode (idle, run, playback, etc.)

(2) Hardware status, including:

(a) Frequency Monitor Subassembly (FMS)

(b) Modcomp Computer

(c) Magnetic tape units

(d) Programmed Oscillator Control Assembly (POCA)

(e) Occultation Converter Subassembly

2. DRS Configuration Display

(1) Receiver channel select

(2) A-D conversion mode

(3) Sampling rate (s)

(4) SSI channel select

(5) Receiver filter select

3. DRS Control Data Display

(1) Predicts set ID

(2) FMS frequency

(3) Frequency offsets

(4) Time offsets

(5) Time-tagged Programmed Oscillator (PO) frequency

(6) Time-tagged PO frequency rate

4. DSS-NOCC Interface. Transmission between the DSS and NOCC shall be via HSDL.

D. Functional Requirements for DRS Display During VLBI Operations

The NRS shall format and provide displays to the NDS of DRS status, configuration, and control data during VLBI operations.

1. DRS Status Display

(1) Program mode (idle, run, playback, etc.)

(2) Hardware status, including:

(a) Modcomp Computer

(b) Magnetic tape units

(c) VLBI Converter Subassembly

2. DRS Configuration Display

(1) Receiver channel select

(2) Sampling rate

(3) Receiver filter select

3. DRS Control Data Display

(1) Recorder on-off

(2) Predict set ID

4. DSS-NOCC Interface. Transmission between the DSS and NOCC shall be via HSDL.

E. Functional Requirements for Display of VLBI Calibration and Ancillary Data

1. Display Data Requirements. The NRS shall format and display VLBI calibration and ancillary data from two DSS simultaneously and in two-station comparative format. The displays shall contain the data types listed in Table 2.

2. DSS-NRS Interface. Transmission between the DSS and NRS shall be via HSDL.

F. Functional Requirements for Validation of VLBI Calibration and Ancillary Data

The NRS shall utilize standards and limits to validate VLBI calibration and ancillary data. The data to be validated are:

(1) Configuration

(2) Status

(3) Other calibration and ancillary data (within limits)

The same set of data limits shall be used for validating each parameter at both DSS.

1. VLBI Data Alarms. The NRS shall provide alarms that result from the validation process for the following data types and operating conditions:

(1) All VLBI data types listed in Table 2

(2) DSS not pointing to correct source

(3) DSS not pointing to same source

(4) VLBI predicts missing for DSS

(5) Recorders not running at DSS

G. Functional Requirements for Display of VLBI Data Accountability

The NRS shall format and display VLBI data accountability and VLBI fringe information as provided by the VLBI Processor Subsystem.

H. Functional Requirements for Provision of Status Data to the NOCC Monitor and Control Subsystem (NMC)

The NRS shall provide its operational status to the NMC. In addition, the NRS shall provide all NRS alarm messages to the NMC for inclusion on the Network Performance Record; the expected list of such alarm messages is provided in Table 3.

IV. NRS Planned Implementation Schedule

The implementation schedule for ongoing and future implementation within the NOCC Radio Science Subsystem is as follows:

(1) Open-loop spectra display ..... April 1, 1980

(2) DRS status and configuration displays (Radio Science and VLBI) ..... Nov. 15, 1979

(3) VLBI calibration and ancillary data display ..... Nov. 15, 1979

(4) VLBI calibration and ancillary data validation ..... Nov. 15, 1979

(5) VLBI data accountability display ..... Nov. 15, 1979

References

1. Berman, A. L., "The DSS Radio Science Subsystem – Data Handling of Very Long Baseline Interferometry (VLBI) Data," in *The Deep Space Network Progress Report 42-46*, pp. 115-122, Jet Propulsion Laboratory, Pasadena, California, August 15, 1978.

2. Berman, A. L., "DSN Radio Science System Mark III-78 Real-Time Display Capability," in *The Deep Space Network Progress Report 42-40,* pp. 141-145, Jet Propulsion Laboratory, Pasadena, California, August 15, 1977.

3. Berman, A. L., "Planetary Atmosphere Modeling and Predictions," in *The Deep Space Network Progress Report 42-42,* pp. 125-129, Jet Propulsion Laboratory, Pasadena, California, December 15, 1977.

Table 1. Radio metric parameters to be displayed

Doppler frequency
Doppler pseudo residual
Doppler noise, including:
High rate
Medium-high rate
Medium-low rate
Low rate
High rate (data type 38)
S-X Doppler, discrete
S-X Doppler, summed
Received signal level
Angle 1
Angle 2
DRVID
DRVID noise
Pseudo-DRVID
S-X range
P_r/N_o

Table 2. VLBI calibration and ancillary data types to be displayed

Station configuration data:
Polarization (right or left circularly polarized)
Receiver bandwidth
Station clock offset
Comb generator frequencies and powers
Frequency dwell (time per receiver channel)
Station frequency standard
Antenna Pointing Subsystem (APS) mode
Recorders on-off
Traveling Wave Maser (TWM) used
S- and X-band system noise temperature
Calibration tone data:
Power
Frequency
Phase Stability
Weather parameters:
Temperature
Pressure
Dew point
Water vapor partial pressure
VLBI angle and angle residual data:
Source right ascension, declination
Local hour angle/declination angle and angle rate
Event and alarm conditions
VLBI source identification (ID)
VLBI source strength:
Predicted
Measured
DSS (40-character) alphanumeric information

Table 3. VLBI System alarms to be recorded on the network performance record.

Single Station Alarms:
VLBI data received
Unable to pair DSS
No source tracked
Recording not started
Recording started
Clock offset not entered
Clock offset entered
Comb generator frequency not entered
Comb generator frequency entered
System noise temperature (SNT) not entered
SNT entered
Frequency standard change: XXXX
VLBI status change: XXXX
VLBI mode change: XXXX
APS mode change
Conscan mode change: XXXX
Meteorological Monitor Assembly (MMA) data not received
MMA data received
Status data not received
Status data received
Angle data not received
Angle data received
Comb generator signal missing/
Wrong channel: X
Comb generator signal corrected
Bandwidth SNR not measured
Bandwidth SNR measured
Channel not in use: X
Source change/antenna not moved
Source change: XXXXXXXXX
Session terminated
VLBI Pair Alarms:
Sources do not match
Pointing does not match
Predicts do not match
VLBI pair
End session both DSS
Configuration not identical
Both DSS not recording
Source change one DSS: XX
Source change both DSS

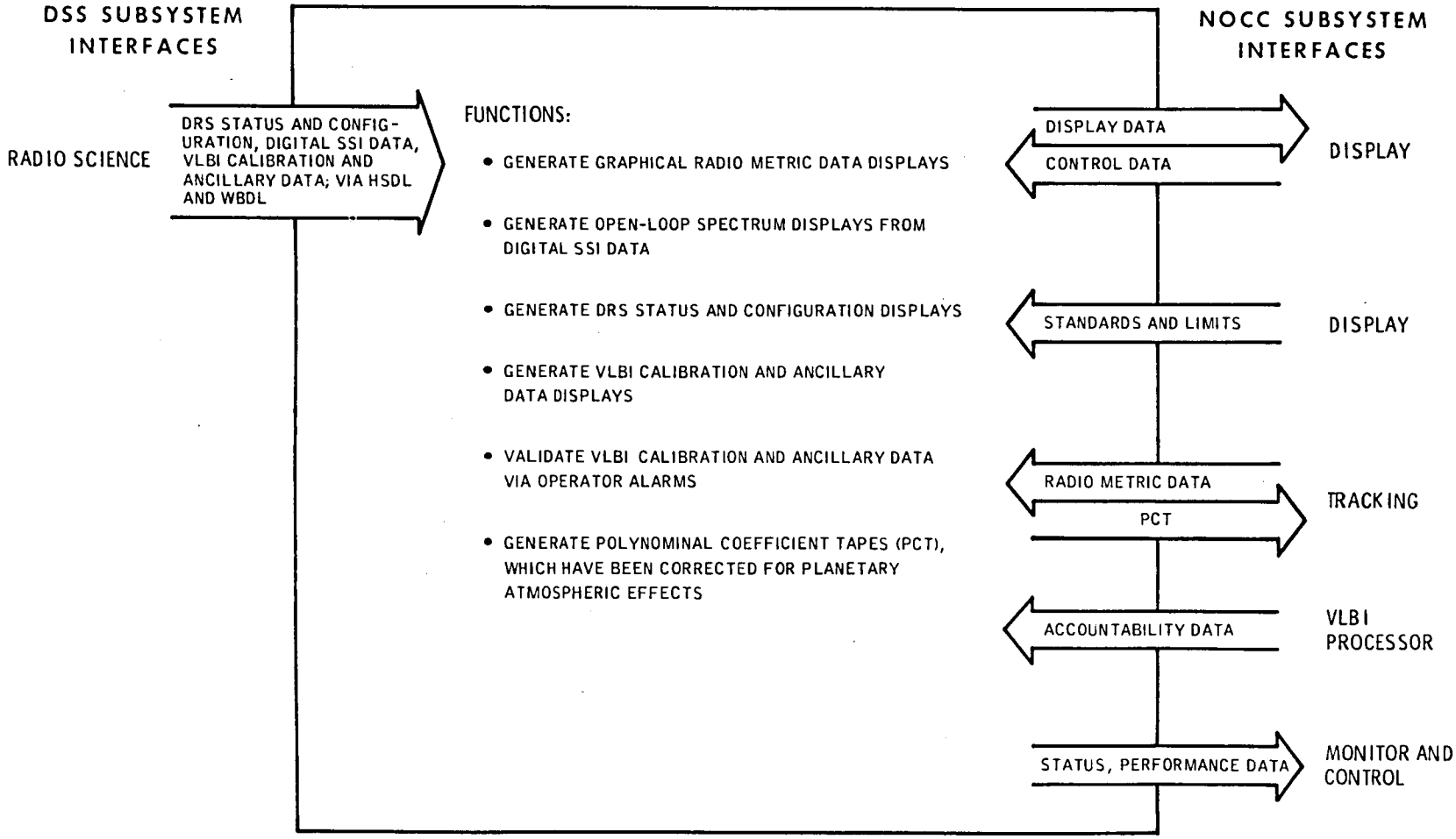


Fig. 1. NOCC Radio Science Subsystem functions and interfaces

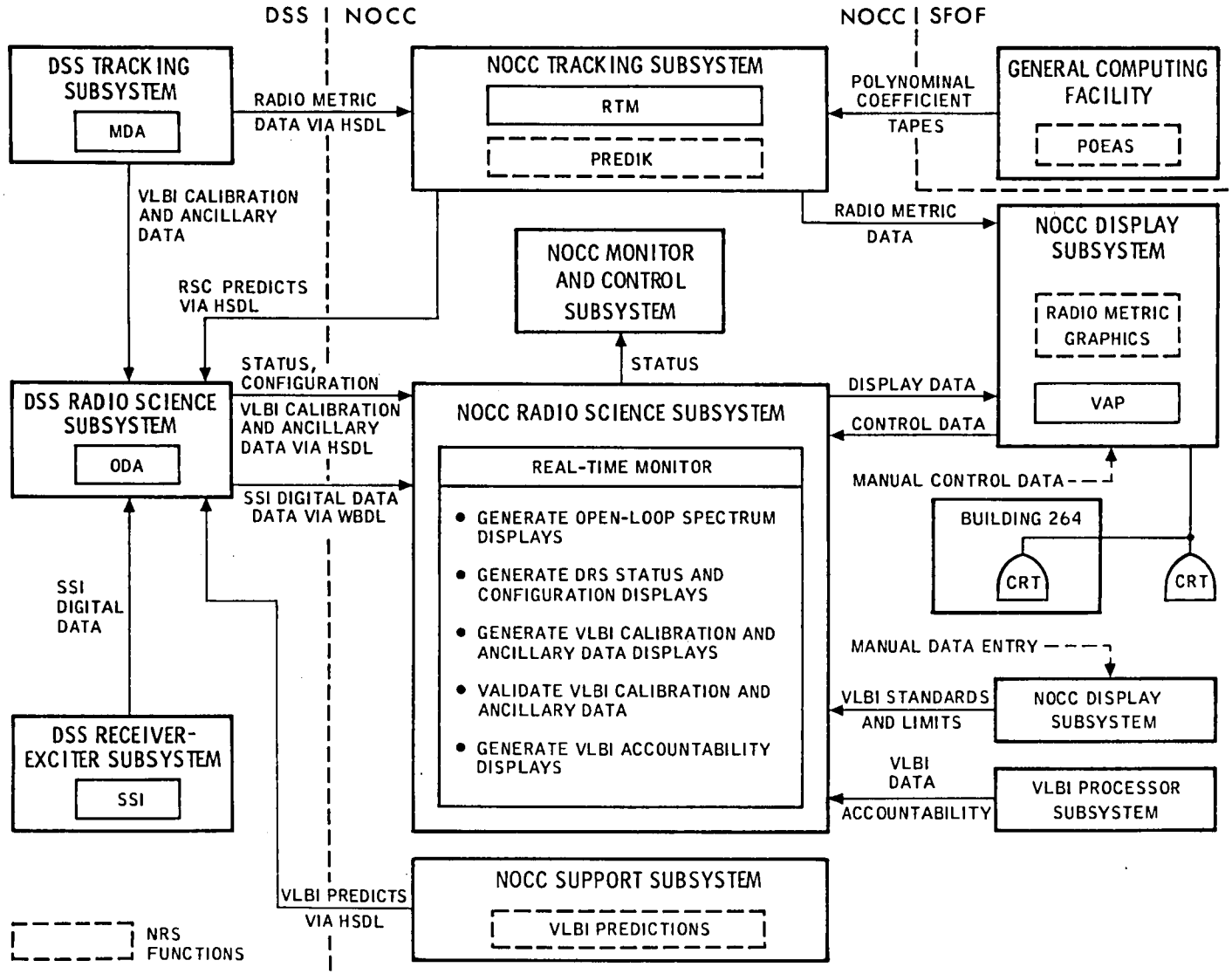
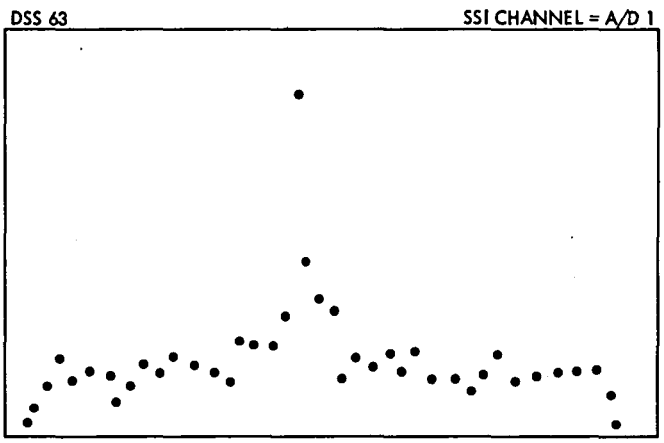


Fig. 2. NOCC Radio Science Subsystem functions and data flow



FULL SCALE AMPLITUDE = 52 dB
FULL SCALE FREQUENCY BANDWIDTH = 10.0 kHz
NUMBER OF AVERAGED SPECTRA = 32
MAXIMUM POWER POINT SNR = 30 dB
MAXIMUM POWER POINT FREQUENCY = 4.55 kHz
TRANSFORM SIZE 4096

Fig. 3. Sample display of SSI digital data

Design of a Quick-Look Decoder for the DSN (7, 1/2) Convolutional Code

C. A. Greenhall and R. L. Miller
Communications Systems Research Section

In a previous article, the authors showed that the DSN (7,1/2) convolutional code could be decoded by a simple "quick-look" method that requires only two shift registers of length 7 for the incoming hard-quantized channel symbols. Such a decoder is being developed for Project Galileo to circumvent an incompatibility between the DSN and the Tracking and Data Relay Satellite System (TDRSS), whose versions of the (7,1/2) code differ in the order of the symbols in each pair. Because signal-to-noise ratios are high during the near-Earth phase of the Galileo mission, quick-look decoding is feasible. The present article derives design parameters and performance figures for the three functions of the decoder: acquisition of node synchronization, generation of the decoded bits, and estimation of channel quality.

I. Introduction

The quick-look decoder discussed here is proposed for a 1200 bit/s telemetry link that will operate while the TDRSS is relaying the signal from the Galileo spacecraft. Figure 1 shows the environment of the proposed decoder. At present, the maximum likelihood (Viterbi) decoder of the TDRSS cannot be used because it expects the symbols in each pair to be reversed relative to the DSN convention. Our proposal is to transmit the hard-quantized symbols (most significant bits) from the 3-bit quantized symbols via the Ground Communications Facility (GCF) to the Mission Control and Computing Center (MCCC) at JPL, where the quick-look decoder would be implemented in software.

We explained the concept of quick-look decoding in Ref. 1. Here, we give a concrete design of such a decoder for the (7,1/2) code. Besides generating the decoded bits, the decoder also acquires node synchronization and estimates channel quality. Working from stochastic models for the channel symbol errors and information bits, we choose the parameters and compute the performance of the algorithms for these services. The parity check bits discussed briefly in Ref. 1 play a crucial role, analogous to the role of the branch metrics in Viterbi decoding. A simple error correction method, not anticipated in Ref. 1, lowers the bit error rate from $7p_e$ to $133p_e^2$, where p_e is the symbol error rate.

The following display recapitulates the encoding and decoding processes given in Ref. 1. All arithmetic is performed modulo 2.

Quantity	Sequence or vector	Formal power series
1. Information bit	b_n	$B(x) = \Sigma b_n x^n$
2. Code vectors	$C_1 = 1011011$	$C_1(x) = 1 + x^2 + x^3 + x^5 + x^6$
	$C_2 = 1111001$	$C_2(x) = 1 + x + x^2 + x^3 + x^6$
3. Channel symbols	s_{1n}, s_{2n}	$S_i(x) = \Sigma s_{in} x^n$ $= C_i(x) B(x), i = 1, 2$
4. Received symbols	s^*_{1n}, s^*_{2n}	$S^*_i(x) = \Sigma s^*_{in} x^n, i = 1, 2$
5. Inversion vectors	$A_1 = 00101$	$A_1(x) = x^2 + x^4$
	$A_2 = 11111$	$A_2(x) = 1 + x + x^2 + x^3 + x^4$
6. Decoded bit	b^*_n	$B^*(x) = \Sigma b^*_n x^n$ $= A_1(x) S^*_1(x) + A_2(x) S^*_2(x)$
7. Parity check bit	p_n	$P(x) = \Sigma p_n x^n$ $= C_2(x) S^*_1(x) + C_1(x) S^*_2(x)$

Reference 1 does not discuss symbol or bit errors. Here, we let

$$E_i(x) = \Sigma e_{in} x^n = S^*_i(x), - S_i(x), (i = 1, 2)$$

$$D(x) = \Sigma d_n x^n = B^*(x) - B(x),$$

be the formal power series for the symbol errors

$$e_{1n} = s^*_{1n} - s_{1n}, \quad e_{2n} = s^*_{2n} - s_{2n},$$

and the decoded bit errors

$$d_n = b^*_n - b_n .$$

The identities

$$C_2(x)C_1(x) + C_1(x)C_2(x) = 0,$$

$$A_1(x)C_1(x) + A_2(x)C_2(x) = 1,$$

give the parity bits and decoded bit errors in terms of the symbol errors:

$$P(x) = C_2(x)E_1(x) + C_1(x)E_2(x), \qquad (1)$$

$$D(x) = A_1(x)\, E_1(x) + A_2(x)\, E_2(x). \qquad (2)$$

These formulas are basic for the analysis of the decoder.

In the DSN version of this code, the first symbols s_{1n} are inverted by the encoder, i.e., s_{1n} (DSN) $= s_{1n} + 1$ (mod 2), to provide enough symbol transitions for the symbol synchronizer. The formula for computing the decoded bits (line 6 of the display) is unchanged because A_1 has an even number (2) of ones. The parity bits obtained from line 7 have to be inverted because C_2 has an odd number (5) of ones. Each parity bit p_n is the sum of 5 of the s_{1k} and 5 of the s_{2k}. If each s_{1k} is inverted then p_n is inverted.

II. Decoding With Correction of Isolated Errors

The basic decoder described in Ref. 1 does no error correction. If symbol errors are rare, however, then most of them can easily be recognized from the parity stream. If $E_1(x) = 1$, $E_2(x) = 0$ (i.e., s^*_{10} is wrong and all other symbols are right), then from Eqs. (1) and (2) we have $P(x) = C_2(x)$, $D(x) = A_1(x)$. The pattern $C_2 = 1111001$ appears in the parity stream p_n, and 00101 in the bit error stream d_n. Thus, if we see the C_2 pattern from time $n - 6$ to time n in the parity stream, then, presuming that $e_{1,n-6} = 1$, we correct the last few decoded bits by adding the vector A_1 to them, modulo 2. Similarly, if we see C_1 in the parity stream we correct the decoded bits by adding A_2. We formalize this as follows:

Correction Algorithm. Keep the last 7 parity bits and decoded information bits in shift registers

$$P = (p_{n-6}, p_{n-5}, \ldots, p_n),$$

$$B^* = (b^*_{n-6}, b^*_{n-5}, \ldots, b^*_n).$$

For each bit time n

If $P = 1011011$ then let $B^* = B^* + 1111100$

If $P = 1111001$ then let $B^* = B^* + 0010100$

Pass b^*_{n-6} to the calling program.

This algorithm corrects all bit errors due to isolated symbol errors; a symbol error at time n is said to be *isolated* if there are no other symbol errors from time $n - 6$ to $n + 6$, inclusive. Most symbol errors are isolated. Bit errors are caused by "bursts" of two or more symbol errors. The derivation below yields the following performance estimates: Assume that the

symbol errors are independent, with probability p_e of occurring. (This is the average number of symbol errors per *symbol*, not per bit.) Then the bit error probability p_d satisfies

$$133p_e^2\,(1 - 36p_e) < p_d < 133p_e^2 \qquad (3)$$

for $p_e \leqslant 10^{-2}$. We shall use the upper bound for computation.

Figure 2 plots the bit error rate (BER), p_d, of the quick-look decoder, and the symbol error rate (SER), p_e, as functions of the bit signal-to-noise ratio E_b/N_o. Formulas used are Eq. (3), together with

$$p_e = Q\left(\sqrt{E_b/N_o}\right),$$

where $Q(x)$ is the probability that a standard Gaussian is greater than x.

According to our present information about the Galileo-TDRSS link, the E_b/N_o, including receiver and demodulation losses, will be 11.3 dB. From Fig. 2 we see that the quick-look decoder margin is 0.5 dB for a BER of 10^{-5}, and 3.3 dB for a BER of 5×10^{-3}.

To derive the estimate (3), we introduce the following assumptions and notation: At time zero, the decoder shift registers are free of symbol errors. The symbol errors e_{1n}, e_{2n}, $n \geqslant 1$, are independent with probability p_e of being 1 and $1 - p_e$ of being 0. The notation e_n designates the pair (e_{1n}, e_{2n}), and $e_n = 0$ means $e_{1n} = 0, e_{2n} = 0$.

Definition. A (symbol error) *burst* is said to occur between bit times m and n $(m \leqslant n)$ if

(1) $e_m \neq 0, e_n \neq 0$;

(2) $e_k = 0, m - 6 \leqslant k \leqslant m - 1, n + 1 \leqslant k \leqslant n + 6$;

(3) $(e_m \ldots, e_n)$ contains no run of 6 consecutive e_k equal to 0.

For $1 \leqslant m \leqslant 6$ a burst can start at m provided all previous bit times are free of symbol errors.

A burst having r symbol errors is called an *r-burst*. A 1-burst is also called an *isolated symbol error.*

Every symbol error in the sequence e_n belongs to some burst. The correction algorithm corrects all bit errors due to isolated symbol errors. We shall use Feller's theory of recurrent events (Ref. 3) to study the bit error rate due to longer bursts.

Definition. The event ξ is said to occur at bit time n, $n \geqslant 7$, if time $n - 6$ is the end of a burst. In other words, $e_{n-6} \neq 0$ and $e_k = 0$ for $n - 5 \leqslant k \leqslant n$.

Each burst is associated with an occurrence of ξ. When ξ occurs, the decoder shift registers are free from symbol errors except in the last position. In the next bit time, e_{n-6} is shifted out of the registers and the situation is exactly like that at time 1. Each time ξ occurs, the process "starts from scratch."

One can prove that ξ is a persistent, aperiodic recurrent event in the sense of Feller. The times between occurrences of ξ are independent random variables with a common distribution. Let M denote the *mean recurrence time*, and let u_n denote the probability that ξ occurs at time n. Then $u_n \to 1/M$ as $n > \infty$. Hence, we can interpret $\lim u_n$ as the *average number of bursts per bit.*

Furthermore, let β denote an arbitrary class of burst types (a burst *type* is the finite sequence of zeros and ones that make up a burst, considered without regard to when the burst occurs), e.g., the 2-bursts, or the bursts of length 10 bit times. Define $\xi(\beta)$ like ξ except that the type of the burst that ends at time $n - 6$ must belong to β. Then $\xi(\beta)$ is also a recurrent event. Let $u_n(\beta)$ be the probability that $\xi(\beta)$ occurs at time n. If ℓ is an upper bound for the lengths of the burst types in β, then $u_n(\beta) = u(\beta)$, a constant, for all $n \geqslant \ell + 12$.

The bit error rate p_d of the decoder can now be written

$$p_d = \sum_{r=2}^{\infty} p_d(r),$$

where $p_d(r)$ is the bit error rate due to r-bursts. (Recall that $p_d(1) = 0$.) As expected, for small p_e the first term $p_d(2)$ dominates the rest. There are 25 individual types of 2-bursts, call them $\beta_1, \ldots, \beta_{25}$ (in no particular order). They are diagrammed below:

1st symbol:	1 01	0000001
2nd symbol:	1 '10 ' ... '	1000000 '
	10	1000000
	01 ' ... '	0000001 '
	11	1000001
	00 ' ... '	0000000 '
	00	0000000
	11 ' ... '	1000001 '

The bit error rate due to 2-bursts is given by

$$p_d(2) = \sum_{k=1}^{25} x_k u(\beta_k),$$

where x_k is the number of bit errors allowed by the decoder when it encounters a burst of type β_k.

Because the 2-burst types are at most 7 bits long, we have $u_n(\beta) = u(\beta)$ for $1 \leqslant k \leqslant 25$ and $n \geqslant 19$. Because each burst is surrounded by 12 good symbols before and after, and contains up to 12 good symbols in its interior, we have

$$u(\beta_k) = (1 - p_s)^{m_k} p_s^2,$$

where $24 \leqslant m_k \leqslant 36$.

Two examples of the computation of x_k are shown below:

Burst type	Parity errors	Bit errors	x_k
1001	1111001	00101	
0000	1111001	00101	
	1110110001	00101101	4
10000	1111001	00101	
00001	1011011	11111	
	1[1111001]011	0010100	
	C_2	001100111	5

In the second example, the pattern C_2 appears "by accident" in the parity stream, starting at bit time 2. At bit time 8 the correction algorithm, behaving as if an isolated first-symbol error had occurred at time 2, adds A_1 to the decoded bit shift register. Notice that this spurious "correction" merely redistributes the bit errors without increasing their number. This is the only type of 2-burst in which C_1 or C_2 appears in the parity stream.

We find that

$$\sum_{k=1}^{25} x_k = 133,$$

and hence

$$133p_e^2 (1 - p_e)^{36} < p_d(2) < 133p_e^2 (1 - p_e)^{24} \qquad (4)$$

We now estimate the contribution of r-bursts, $r > 2$, to the bit error rate. The effect of the correction algorithm will be neglected, for we just saw that it has no effect on the 2-burst performance. An r-burst begins with the symbol pair (1,1), (1,0), or (0,1). Then there are 0 to 5 good bit times, followed by an $(r-2)$-burst (for the (1,1) case) or an $(r-1)$-burst (for the other two cases). Therefore, the number of r-burst types, y_r, satisfies

$$y_r = 6y_{r-1} + 12y_{r-2}, \quad r \geqslant 3$$

$$y_1 = 2,\ y_2 = 25.$$

The solution of this initial value problem is

$$y_r = K_1 z_1^r + K_2 z_2^r,$$

where $z_1 = 3 + \sqrt{21}, z_2 = 3 - \sqrt{21}, K_1 = (273 - 15\sqrt{21})/504$, $K_2 = (273 + 15\sqrt{21})/504$. The rate of occurrence of each type of r-burst is at most $p_e^r (1 - p_e)^{24}$, and each of the r symbol errors gives rise to at most 5 bit errors. Therefore,

$$p_d(r) < y_r \cdot 5r \cdot p_e^r (1 - p_e)^{24}. \qquad (5)$$

Summing (5) over $r \geqslant 3$ and adding the upper bound for $p_d(2)$ from (4), we obtain an expression that remains below $133\,p_e^2$ for $p_e \leqslant 10^{-2}$. The lower bound in (3) comes from the lower bound in (4) for $p_d(2)$.

In view of the crudity of (5), we conjecture that the upper bound in (3) remains valid even when the effect of the correction algorithm on r-bursts ($r > 2$) is included.

III. Node Synchronization

The decoder receives the stream of corrupted symbols

$$s_{11}^*, s_{12}^*, s_{21}^*, s_{22}^*, s_{31}^*, s_{32}^*, \ldots$$

and must decide how to pair them off. We use the method of the *up-down counter*, also used in Viterbi decoders (Ref. 2). In our case, the counter is driven by the parity bits. If a parity bit is zero, the counter is decremented by 1; if a parity bit is 1, the counter is incremented by $k - 1$, where k is a fixed integer $\geqslant 2$. The counter starts at zero and is not allowed to become negative. If it ever reaches a preassigned threshold T then it is reset to zero, the decoder lets one symbol go by, and the decoder informs the calling program that a resync has occurred.

The problem here is to choose k and T. We must choose k so that the average drift rate of the counter is negative for true sync and positive for false sync. If sync is true, then Eq. (1) holds. The counter rarely departs from zero if the SER is low, because parity errors are rare. Eventually, though, it does reach the threshold, causing a *false alarm*, a needless loss of sync. We want the expected time to false alarm, E_{fa}, to be large. To be specific, we shall set T high enough so that E_{fa} is more than 100 times as long as Galileo will be using the TDRSS. To do this, it is necessary to assume a certain cleanliness of the channel. Our *symbol error assumption* is

$$\text{SER} \leqslant 6.13 \times 10^{-3} \tag{6}$$

at which point E_b/N_o = 8 dB, BER = 5×10^{-3} (see Fig. 2).

Let sync be false. Then one expects a relatively high density of parity errors, which force the counter to rise quickly to the threshold. For this to happen, however, it is essential that the information bit stream contain enough information. If, for example, the information bits are all zero, then most of the corrupted symbols and parity bits are zero, and one cannot tell true sync from false. Our *information bit assumptions* are (1) the ratio of ones to total bits is between 5 percent and 95 percent, and (2) the ratio of transitions to total bits is at least 5 percent. Incorporating these assumptions into stochastic models, we can estimate the expected time to resync the decoder, E_{rs}, which we want to be small. The two quantities E_{fa} and E_{rs} specify the performance of the sync algorithm.

The symbol error and bit assumptions define a *design point*, from which the optimization procedure of subsection C below yields the counter parameters

$$k = 8, \quad T = 512.$$

The performance of the algorithm is

$$E_{fa} > 10^9 \text{ bits}, \quad E_{rs} = 460 \text{ bits}.$$

Our design point is negotiable; if the above choice is unsuitable, the authors are willing to recompute k and T.

A. True Sync

Let us assume that node sync is true and that the symbol errors e_{1n}, e_{2n} are independent with probability p_e of being one. Our aim is to get a lower bound on E_{fa}. We begin with a preliminary remark. As we saw in section II, most of the symbol errors are isolated, and each isolated symbol error produces 5 parity errors. Therefore, the number of parity errors n_p in a given stretch of n_b bits is approximately 5 times the number of symbol errors n_e. Moreover, Eq. (1) implies that $n_p \leqslant 5n_e$ provided that the decoder shift registers are initially clean, which, henceforth, we assume.

The up-down counter executes a random walk with a reflecting barrier at −1 and an absorbing barrier at T. Since the steps are not independent it is convenient to bound the motion of this counter, call it Counter 1, by the motion of another (fictitious) counter, Counter 2, that executes a random walk with independent steps. Counter 2 operates each *symbol* time as follows: If the next symbol is correct then the counter is decremented by 1/2; if it is incorrect then the counter is incremented by $5k - (1/2)$. There is a reflecting barrier at − 1/2 and an absorbing barrier at T. Without barriers, in n_b bit times Counter 1 moves up by $(k-1)n_p - (n_b - n_p) = kn_p - n_b$ (probably negative), while Counter 2 moves up by $[5k - (1/2)]\, n_e - (1/2)(2n_b - n_e) = 5kn_e - n_b$. It follows from our earlier remark that Counter 1 advances no more than Counter 2, with or without barriers.

The difference equation method given in Chapter XIV of Ref. 3 can be used to derive bounds on the mean first-passage time of the absorbing barrier T by Counter 2. Using the conventions of Ref. 3, we consider a random walk on the integers with a reflecting barrier at zero and an absorbing barrier at a. It advances by d with probability p and by −1 with probability $q = 1 - p$. The parameters are given by

$$a = 2T + 1, \quad d = 10k - 1, \quad p = p_e. \tag{7}$$

Let D_j be the mean first-passage time for a walk that starts j units above the zero-level. We assume, but do not prove, that D_j is finite. Then, it satisfies the difference equation

$$D_j = pD_{j+d} + qD_{j-1} + 1, \quad 1 \leqslant j \leqslant a - 1, \tag{8}$$

with boundary conditions

$$D_0 = D_1,$$

$$D_j = 0, \quad a \leqslant j \leqslant a + d - 1$$

Our aim is to find bounds for D_1 (see (16) and (17) below). The characteristic equation of Eq. (8) is

$$pz^d + qz^{-1} = 1, \tag{9}$$

which, provided $q - pd \neq 0$, has a simple root at 1 and just one other positive root λ. For any A and B, the sequence

$$E_j = \frac{j}{q - pd} + A + B\lambda^j \tag{10}$$

satisfies

$$E_j = pE_{j+d} + qE_{j-1} + 1 .$$

Solving for the A and B that make

$$E_0 = E_1 , \quad E_a = 0 , \tag{11}$$

we get

$$E_j = \frac{1}{q - pd} \frac{\lambda^a - \lambda^j}{\lambda - 1} + j - a \tag{12}$$

The sequence $-E_j$ is convex. This, together with Eq. (11), implies $E_j < 0$ for $j > a$. Therefore, the difference $\Delta_j = D_j - E_j$ satisfies

$$\Delta_j = p\Delta_{j+d} + q\Delta_{j-1} , \quad 1 \leqslant j \leqslant a - 1 , \tag{13}$$

$$\Delta_0 = \Delta_1 , \tag{14}$$

$$\Delta_j \geqslant 0 , \quad a \leqslant j \leqslant a + d - 1 . \tag{15}$$

We claim that $\Delta_j \geqslant 0$ for $0 \leqslant j \leqslant a + d - 1$. To prove this, let Δ_j assume its minimum value m at $j = r$. Because of Eq. (14) we can assume $r \geqslant 1$. Repeated use of Eq. (13) gives $\Delta_{r+d} = m$, $\Delta_{r+2d} = m$, and finally $\Delta_s = m$ for some s such that $a \leqslant s \leqslant a + d - 1$. This proves $m \geqslant 0$.

Thus we have the result

$$D_j \geqslant E_j , \quad 0 \leqslant j \leqslant a + d - 1 .$$

In particular

$$D_1 = D_0 \geqslant \frac{1}{q - pd} \left(\frac{\lambda^a - 1}{\lambda - 1} - a \right) . \tag{16}$$

A similar argument that replaces (11) by

$$E_0 = E_1 , \quad E_{a+d-1} = 0 ,$$

leads to the upper bound

$$D_1 \leqslant \frac{1}{q - pd} \left(\frac{\lambda^{a+d-1} - 1}{\lambda - 1} - (a + d - 1) \right) . \tag{17}$$

For design, we use the lower bound (16), plus (7) and the bound

$$E_{fa} \geqslant \frac{1}{2} D_1 , \tag{18}$$

which follows from the relationship between Counters 1 and 2.

B. False Sync

If node sync is false then we want the up-down counter to rise quickly to the threshold T. We can neglect symbol errors because they produce more parity errors, which make the counter rise even faster.

The encoder transmits the symbols serially in a stream represented by the formal power series

$$S_1(x^2) + xS_2(x^2) . \tag{19}$$

The decoder sends the coefficients of $1, x^2, x^4, \ldots$ down the first-symbol pipe, and the coefficients of $x, x^3, x^5, \ldots$ down the second-symbol pipe. If sync is false, then the decoder sees the stream whose formal power series is x times (19), or

$$xS_1(x^2) + x^2S_2(x^2) . \tag{20}$$

Therefore, it sends $xS_2(x)$ down the first-symbol pipe, and $S_1(x)$ down the second-symbol pipe. The parity stream comes out as

$$\begin{aligned} P(x) &= C_2(x)xS_2(x) + C_1(x)S_1(x) \\ &= [xC_2(x)^2 + C_1(x)^2]\, B(x) \\ &= K(x)B(x) \end{aligned} \tag{21}$$

where

$$K(x) = xC_2(x^2) + C_1(x^2)$$

$$= 1 + x + x^3 + x^4 + x^5 + x^6 + x^7 + x^{10} + x^{12} + x^{13}.$$

(Note that $f(x)^2 = f(x^2)$ (mod 2) for any polynomial $f(x)$.) Each parity bit is the modulo 2 sum of 10 information bits. The density of parity errors depends on the distribution of zeros and ones in the information bit stream. We shall use two models for the b_n.

Model B. Independent bits. Let the b_n be independent with probability p_b of being one, $q_b = 1 - p_b$ of being zero (p_b is *not* a bit *error* probability). The density of parity errors, μ_p, is the probability that the sum of 10 of the b_n is odd. By considering the expansions of $(q_b + p_b)^{10}$ and $(q_b - p_b)^{10}$ we get

$$\mu_p = \frac{1}{2}[1 - (q_b - p_b)^{10}] =: \mu_{pb} \qquad (22)$$

Model T. Independent transitions. Let $t_n = b_n + b_{n-1}$ (mod 2) be independent, with probability p_t of being one, $q_t = 1 - p_t$ of being zero. (The probability that b_n is one is 1/2.) Let $T(x)$ be the corresponding formal power series. Then

$$T(x) = \sum t_n x^n = (1 + x)B(x).$$

Since $K(1) = 0$ we have $K(x) = (1 + x)L(x)$ and

$$P(x) = L(x)T(x),$$

where

$$L(x) = 1 + x^3 + x^5 + x^7 + x^8 + x^9 + x^{12}$$

Because $L(x)$ has 7 nonzero terms, the density of parity errors is now

$$\frac{1}{2}[1 - (q_t - p_t)^7] =: \mu_{pt} \qquad (23)$$

Let $p_b = p_t \leq 1/2$. Then the bit streams in both models carry the same amount of information, and $\mu_{pt} \leq \mu_{pb}$. Hence, in order to be conservative about the upward drift of the counter, we use model T to compute μ_p. The average drift rate of the counter is $(k - 1)\mu_p - (1 - \mu_p) = k\mu_p - 1$, which must be positive if the counter is to advance rapidly and steadily to the threshold. We shall not be far wrong, then, if we estimate the expected time to resync, E_{rs}, by

$$E_{rs} = \frac{T}{k\mu_p - 1}, \qquad (24)$$

where $\mu_p = \mu_{pt}$.

C. Choice of Sync Parameters

We wish to make a rational choice of the counter parameters k and T. Recall that our chosen design point is defined by

$p_e = 6.13 \times 10^{-3}$ symbol error rate

$p_t = 0.05$ bit transition rate.

The decoder would be used by the Galileo mission for at most one hour, about 4×10^6 bits at 1200 b/s. Let us require that the expected time to false alarm, E_{fa}, be much greater, say 10^9 bits. Keeping p_e and p_t fixed we try a value of k. By trial and error we find the least T such that the lower bound in (18) for E_{fa} exceeds 10^9. Then the expected time to resync, E_{rs}, is obtained from Eq. (24). This gives E_{rs} as a function of k. If k is too large, then the counter under true sync has less negative drift; hence T must be made large in order to make $E_{fa} > 10^9$, and this makes E_{rs} large. If k is too small, then T can be made smaller for a given E_{fa}, but E_{rs} still becomes large because the denominator in Eq. (24) is close to zero. There is a k that minimizes E_{rs}; for our design point we get $k = 8$, $T = 511$, $E_{rs} = 471$. The contribution from false alarms to bit error rate is $E_{rs}/E_{fa} > 4.7 \times 10^{-7}$.

Of course, if the channel is actually better than our design point, then E_{fa} is much greater; for example, if $p_e = 10^{-3}$ then $E_{fa} > 9 \times 10^{22}$. On the other hand, if $p_e = 10^{-2}$ then $1.1 \times 10^5 < E_{fa} < 1.9 \times 10^5$ (from the theory of Counter 2), so that we should expect a false alarm every 2 minutes on the average.

IV. Estimation of Channel Quality (SER)

Assume that node sync is true. As we pointed out earlier, the number of parity errors n_p in n_b bits is about $5n_e$, where n_e is the number of symbol errors. Therefore, the symbol error rate p_e can be estimated from n_p by

$$p_e \approx \frac{n_e}{2n_b} \approx \frac{n_p}{10n_b}. \qquad (25)$$

To use Eq. (25), one could wait until $n_p = 5n$ where n is a fixed integer, say 25. Then the relative one-sigma error in the estimator Eq. (25) is about $1/\sqrt{n}$. If p_e is low, say 10^{-6}, then symbol errors occur only once every 7 minutes; thus, it would take too long to accumulate 25 of them. To accommodate a wide range of SER, we propose the following scheme. Assume that the decoder shares memory with a "main program." The locations n_p, n_b, and p_e are available to both programs. Initially, all three locations are set to zero.

The *decoder* accumulates the parity counts in n_p and bit counts in n_b. When n_p reaches $5n$, the decoder sets $p_e = n_p/(10n_b)$, $n_p = 0$, $n_b = 0$, and begins counting again. We suggest $n = 25$.

The *main program*, whenever it wants an SER estimate, checks p_e. If $p_e \neq 0$ it uses p_e as the estimate and sets $p_e = 0$. If $p_e = 0$ it computes $n_p/(10n_b)$ for itself as the estimate.

In other words, the main program takes the decoder's most recent update, if one has been provided since the main program's last inquiry; otherwise, the main program does the best it can with the most recent bit counts. The decoder updates p_e once every $n/(2p_e)$ bits, on the average.

V. Conclusions

We have derived design parameters and performance of a simple quick-look decoder for the DSN (7, 1/2) convolutional code. The decoder uses parity check bits for correcting isolated errors, finding node sync, and estimating the symbol error rate. An up-down counter, driven by the parity bits, is used to detect false node sync. The performance margin for near-Earth Galileo telemetry is estimated to be 3.3 dB for a bit error rate of 5×10^{-3}.

References

1. Greenhall, C. A., and Miller, R. L., "Quick-Look Decoding Schemes for DSN Convolutional Codes," *DSN Progress Report 42-51*, pp. 162-166, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, 15 June 1979.
2. Gilhousen, K. W., et al., *Coding Systems Study for High Data Rate Telemetry Links*, Linkabit Corporation, San Diego, California, January 1971.
3. Feller, W., *Introduction to Probability Theory and Its Applications*, Volume I, 3rd Edition, New York, 1968.

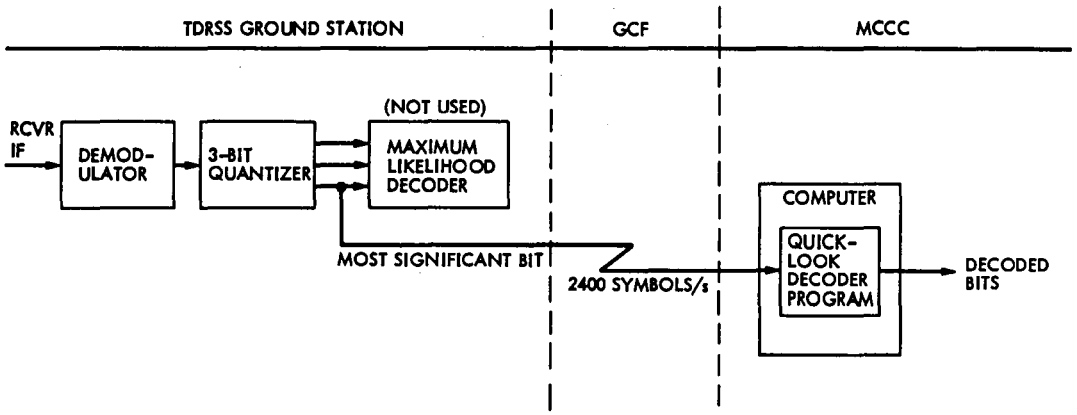


Fig.1. Environment of quick-look decoder

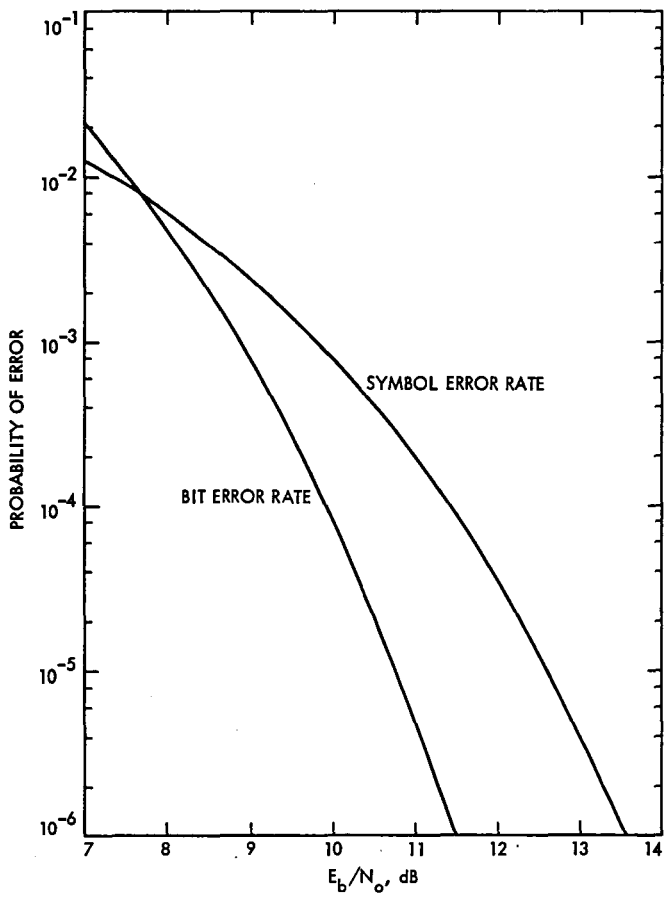


Fig. 2. Performance of quick-look decoder

A Reed-Solomon Decoding Program for Correcting Both Errors and Erasures

R. L. Miller, T. K. Truong, and B. Benjauthrit
Communication Systems Research Section

I. S. Reed
Department of Electrical Engineering
University of Southern California

This article discusses the software implementation of a simplified algorithm for decoding errors and erasures of Reed-Solomon (R-S) code words using the techniques of finite field transforms and continued fractions. In particular, random code words from the (255, 223, 33) R-S code over GF(2^8) are corrupted by random error and erasure patterns, and decoded whenever theoretically possible. A matrix of execution times of this decoder under varying sets of errors and erasure patterns is also included. This matrix demonstrates the relative amounts of time required for decoding different error and erasure patterns, as well as the correctness of the algorithms and the software implementation.

I. Introduction

Recently the authors in Ref. 1 developed a simplified algorithm for correcting erasures and errors of Reed-Solomon (R-S) codes over the finite field $GF(p^m)$, where p is a prime and m is an integer. For a space communication link, it was shown (Ref. 2) that a 16-error-correcting R-S code of 255 8-bit symbols, concatenated with a $k = 7$, rate = 1/2 or 1/3, Viterbi decoded convolutional code, can be used to reduce the signal-to-noise ratio required to meet a specified bit-error rate. Such a concatenated code is being considered for the Galileo project and the NASA End-to-End Data System.

In such a concatenated code, the inner convolutional decoder is sometimes able to find only two or more equally probable symbols. Then the best policy is to declare an erasure of the symbol. If the outer R-S code is able to use the additional information that erasures have occurred, then it is reasonable to assume that the system performance will be enhanced.

In this article, an algorithm based on the ideas in Ref. 1 is used to correct erasures and errors of R-S code words using a finite field transform over $GF(2^8)$. This algorithm is written in FORTRAN V and is implemented on the UNIVAC 1108 computer. A matrix of decoding times for correcting errors and erasures of the code words is given at the end of this paper.

II. The R-S Decoding Algorithm

The algorithm given in (Ref. 1) was used to correct patterns of t errors and s erasures of the words of the (255, 223, 33) R-S code, where $2t + s < 33$ and the symbols belong to the finite field $GF(2^8)$. Define the following five vectors:

$(c_0, c_1, \cdots, c_{254})$ = c, code vector

$(r_0, r_1, \cdots, r_{254})$ = r, received vector

$(\mu_0, \mu_1, \cdots, \mu_{254})$ = μ, erasure vector

$(e_0, e_1, \cdots, e_{254})$ = e, error vector

$(\tilde{\mu}_0, \tilde{\mu}_1, \cdots, \tilde{\mu}_{254})$ = $\tilde{\mu}$, new erasure vector

These vectors are related by $\mathbf{r} = \mathbf{c} + \mu + \mathbf{e}$ and $\tilde{\mu} = \mathbf{e} + \mu$.

Suppose that t errors and s erasures occur in the received vector r of 255 symbols and assume $2t + s < 33$. Then the decoding procedure consists of the following steps:

Step 1: Compute the syndromes S_k $(1 \leqslant k \leqslant 32)$ of the received 255-tuple $(r_0, r_1, \cdots, r_{254})$, i.e.,

$$S_k = \sum_{i=0}^{254} r_i \alpha^{ik} \quad \text{for } k = 1, 2, \cdots, 32 \tag{1}$$

where α is an element of order 255 in $GF(2^8)$. If $S_k = 0$ for $1 \leqslant k \leqslant 32$, then r is a code word and no decoding is necessary. Otherwise,

Step 2: Compute τ_j for $j = 0, 1, 2, \cdots, s$ from the erasure locator polynomial

$$\tau(x) = \prod_{j=1}^{s} (x - Z_j) = \prod_{j=1}^{s} (-1)^j \tau_j x^{s-j} \tag{2}$$

where s is the number of erasures in the received vector, and Z_j $(1 \leqslant j \leqslant s)$ are the known erasure locations. Then compute the Forney syndromes T_i for $1 \leqslant i \leqslant d - 1 - s$ from the equation

$$T_i = \sum_{j=0}^{s} (-1)^j \tau_j S_{i+s-j} \quad \text{for } 1 \leqslant i \leqslant d - 1 - s \tag{3}$$

where τ_j $(1 \leqslant j \leqslant s)$ and S_j $(1 \leqslant j \leqslant 32)$ are known.

Step 3: If $0 \leqslant s < 32$, then use continued fractions to determine the error locator polynomial $\sigma(x)$ from the known T_i's $(1 \leqslant i < 32 - s)$. For the special case $s = 32$, it was shown (Ref. 3) that it is impossible for any decoder to tell whether there are zero or more additional errors. Thus, for $s = 32$, the best policy is not to decode the message at all. If $2t + s \geqslant 33$, it is shown in Ref. 4 that the continued fraction algorithm will not determine the correct error locator polynomial and will (1) either terminate abnormally, (2) terminate normally with a polynomial whose roots do not represent possible error locations, or (3) terminate normally with a polynomial whose roots do represent possible error locations.

In the first two cases, a decoding failure will occur; the decoder will be unable to decode the received word. In the third case, a decoding error will occur; there will be no alarm telling of the inability to decode; the decoder will operate as is $2t + s < 33$. The probability of such miscorrections of errors and erasures of R-S code words is discussed further in Ref. 4.

Step 4: Use a Chien-type search to find the t roots of the error locator polynomial. If t distinct roots cannot be found that represent possible error locations, then declare a decoding failure. Otherwise,

Step 5: Compute the combined erasure and error locator polynomial from the equation

$$\tilde{\tau}(x) = \sigma(x)\tau(x) = \sum_{k=0}^{s+t} (-1)^k \tilde{\tau}_k x^{s+t-k} \tag{4}$$

where $\sigma(x)$ and $\tau(x)$ are now known. Then compute the rest of transform of the erasure and error vector from the equation

$$S_\ell = \sum_{k=1}^{s+t} (-1)^k \tilde{\tau}_k S_{\ell-k} \quad \text{for } \ell > d - 1$$

Step 6: Invert the transform of $\tilde{\mu}$ at the points corresponding to the known error and erasure locations to obtain the amplitudes of $\tilde{\mu}$. That is,

$$\tilde{\mu}(Z_j) = e_i + \mu_i = \sum_{k=0}^{254} S_k Z_j^k \quad \text{for } j = 1, 2, \cdots, s + t \tag{5}$$

where Z_j are the known error and erasure locations of $\tilde{\mu}$. Then subtract from the received word the error and erasure vector to obtain the corrected code word.

III. Program Design and Implementation

The decoding procedure described in the previous section was implemented on the UNIVAC 1108 computer using FORTRAN V. This program was used to correct any combination of t errors and s erasures occurring in the 255-symbol R-S code words, where $2t + s < 33$. The overall basic structure of the program is given in Fig. 1. It is divided into a main program and five major subroutines.

The Main Program: This is the main driver of the rest of the program. It initializes the decoding process and keeps track of the elapsed CPU time.

Input: This subroutine generates a random code vector (polynomial) $R(x)$ for the R-S decoder and then adds errors and erasures $E(x)$ to it.

Step 1: The first 32 syndromes of the received vectors as well as the corrected vectors are calculated in this subroutine. In case the corrected received word is not an R-S code word, the subroutine will output the message, "The corrected received vector is not a codeword." This is helpful in confirming the correctness of the program, as well as indicating that the number of errors and erasures have exceeded the limits allowable by the decoder.

Step 2: This subroutine computes the Forney syndrome vector T (Eq. 3) from the erasure vector Z. The erasure locator polynomial $\tau(x)$ (Eq. 2) is also calculated here.

Step 3: The error locator polynomial $\sigma(x)$ is calculated from the Forney syndrome vector T using the continued fraction algorithm. The product of the error locator polynomial $\sigma(x)$ and the erasure locator polynomial $\tau(x)$ is next computed (Eq. 4). The coefficients of this erasure and error locator polynomial are used to compute the remaining terms $S_{33}, \cdots, S_{255}$. Also, the locations of the errors are now determined. (The erasure locations are known a priori by definition.)

Step 4: This step directly computes the inverse Fourier transform of the vector $(S_1, S_2, \cdots, S_{255})$ to obtain the nonzero error and erasure magnitudes. Finally, the received vector is corrected to provide an estimate of transmitted code word.

IV. Simulation Results

The computation times for decoding numerous code words that were corrupted by errors and erasures are given in Table 1. These results were obtained by computing each entry in Table 1 five times and then averaging. Along any row or column, the computation times tend to increase with the row or column indices until decoding failures occur due to an excess of allowable errors and erasures.

V. Summary

A software Reed-Solomon decoder has been developed for the (255, 223, 33) R-S code whose symbols belong to $GF(2^8)$. The simulation indicates that the decoder correctly reconstructs the transmitted code word from the received word if the code word has been corrupted by t errors and s erasures, where $2t + s < 33$.

Acknowledgment

This work was supported in part by the U.S. Air Force Office of Scientific Research under Grant Number AFOSR 75-2798 and also in part by NASA Contract NAS7-100.

References

1. I. S. Reed and T. K. Truong, "A Simplified Algorithm for Correcting Both Errors and Erasures of R-S Codes," in *The Deep Space Network Progress Report 42-48*, Jet Propulsion Laboratory, Pasadena, California, September 1978.

2. J. Odenwalder, et al., *Hybrid Coding Systems Study Final Report* Linkabit Corp., NASA CR114, 486, September 1972.

3. E. R. Berlekamp and J. L. Ramsey "Readable Erasures Improve the Performance of Reed-Solomon Codes" *IEEE Transactions on Information Theory*, Vol. IT-24, No. 5, Sept. 1978.

4. R. L. Miller, I. S. Reed, and T. K. Truong, "The Probability of Incorrectly Decoding Errors and Erasures and Reed-Solomon Code Words," submitted to *IEEE Trans. on Inform Theory.*

Table 1. Decoder execution times in seconds

Erasures \ Errors	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
0	0.314	0.354	0.397	0.463	0.529	0.483	0.515	0.564	0.608	0.681	0.750	0.760	0.734	0.812	0.777	0.940	0.878	0.592
1	0.342	0.375	0.466	0.472	0.483	0.495	0.549	0.584	0.625	0.661	0.709	0.771	0.746	0.792	0.887	0.867	0.414	
2	0.377	0.427	0.439	0.455	0.485	0.557	0.561	0.616	0.609	0.737	0.813	0.703	0.811	0.899	0.871	0.837	0.577	
3	0.382	0.421	0.455	0.481	0.511	0.571	0.574	0.607	0.744	0.757	0.697	0.779	0.814	0.840	0.836	0.429		
4	0.439	0.456	0.466	0.524	0.594	0.634	0.588	0.629	0.761	0.679	0.710	0.778	0.848	0.943	0.956	0.577		
5	0.437	0.471	0.483	0.582	0.583	0.596	0.609	0.638	0.763	0.759	0.782	0.810	0.868	1.054	0.409			
6	0.439	0.472	0.552	0.572	0.574	0.631	0.696	0.656	0.697	0.724	0.808	0.794	0.930	0.997	0.521			
7	0.474	0.492	0.625	0.612	0.588	0.660	0.782	0.679	0.770	0.817	0.790	0.815	0.943	0.401				
8	0.492	0.610	0.647	0.661	0.640	0.648	0.699	0.692	0.820	0.897	0.784	0.864	0.906	0.521				
9	0.583	0.585	0.665	0.626	0.625	0.664	0.720	0.741	0.781	0.840	0.812	0.918	0.432					
10	0.659	0.607	0.668	0.683	0.730	0.724	0.746	0.767	0.769	0.823	0.936	1.034	0.485					
11	0.561	0.695	0.682	0.661	0.741	0.763	0.733	0.770	0.818	0.811	0.849	0.405						
12	0.606	0.661	0.681	0.756	0.691	0.741	0.753	0.783	0.919	0.830	0.912	0.462						
13	0.656	0.638	0.687	0.737	0.708	0.740	0.800	0.816	0.901	0.858	0.356							
14	0.620	0.760	0.716	0.740	0.735	0.779	0.785	0.831	0.924	0.949	0.449							
15	0.693	0.778	0.719	0.749	0.840	0.863	0.824	0.981	0.883	0.424								
16	0.662	0.735	0.715	0.766	0.758	0.860	0.833	0.907	0.912	0.481								
17	0.802	0.744	0.712	0.800	0.789	0.852	0.841	0.864	0.402									
18	0.858	0.720	0.756	0.825	0.806	0.980	0.859	0.879	0.506									
19	0.852	0.747	0.771	0.791	0.802	0.963	0.897	0.347										
20	0.756	0.740	0.841	0.935	0.831	0.878	0.972	0.428										
21	0.758	0.763	0.954	0.911	0.918	0.899	0.386											
22	0.750	0.805	0.956	0.898	0.875	1.049	0.406											
23	0.769	0.807	0.825	0.858	0.910	0.429												
24	0.811	0.902	0.893	0.892	0.907	0.439												
25	0.829	0.857	0.906	0.945	0.342													
26	0.872	0.920	0.907	0.928	0.372													
27	0.960	0.960	0.935	0.340														
28	0.914	0.908	0.935	0.507														
29	1.028	0.930	0.374															
30	0.985	0.945	0.826															
31	0.960	0.374																

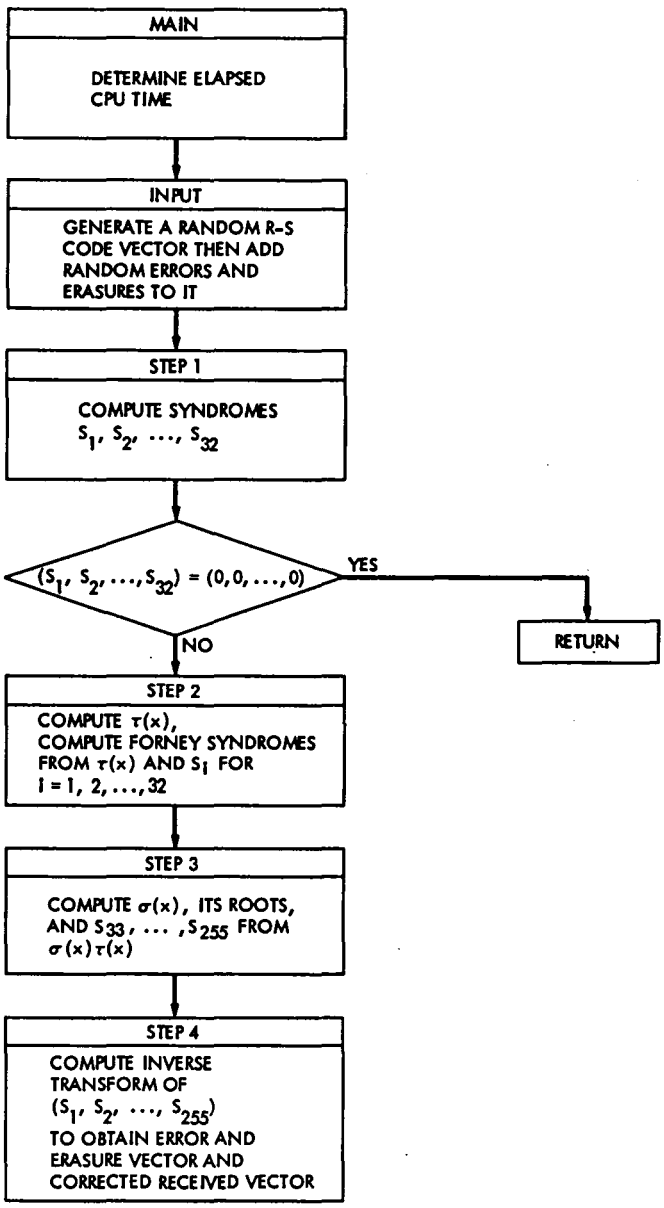


Fig. 1. Basic functional structures of R-S decoding program using transform over GF (2^8) and continued fractions

DSN Tracking System Uplink Frequency Control

G. L. Spradlin
TDA Engineering

The failure of the Voyager 2 spacecraft receiver requires that the DSN maintain the communications link to the spacecraft by constantly tuning the uplink frequency to null the effects of doppler. The following article discusses the planned implementation within the DSN Tracking System to automate this current manual process.

I. Introduction

The failure of the Voyager 2 receiver eliminated the ability of this spacecraft to maintain lock on the Deep Space Station generated uplink frequency, as this frequency is shifted by the time varying effects of doppler. Healthy spacecraft receivers do track the uplink signal, and thus it is standard practice within the DSN to select an uplink frequency that is fixed for a given station view period, and is the predicted spacecraft best-lock frequency (the frequency that results in the quickest acquisition by the unstressed spacecraft receiver), modified by the average doppler effect to be experienced during the station view period.

This standard practice cannot be used for Voyager 2 tracks. Instead, the uplink is sustained by constantly tuning the station-transmitted frequency to compensate for the uplink doppler effect, and in effect provide the spacecraft receiver with a near-constant received frequency (±25 Hz about the best-lock frequency at S-band). Following the receiver failure, procedures and software were hastily developed and put to use to provide the needed communications link with the spacecraft.

These procedures and software were developed to support an emergency spacecraft condition and require continuous support by personnel at JPL and at the supporting station. The numerous points at which human intervention are required and the resulting susceptibility to errors are the major motivating factors behind the implementation of automated uplink frequency control at the station. This implementation is described in the following sections. Figure 1 provides a simplified block diagram of the uplink frequency control implementation.

II. DSS Subsystems

The urgently needed relief from the burden on operations personnel in manually executing the required uplink frequency control sequences dictated that the implementation be accomplished as quickly as possible; and further, that existing capabilities be utilized to the fullest extent possible to avoid the lengthy evolution of wholly new hardware and software. This was the guiding philosophy for the uplink frequency control implementation within the station subsystems, as well as the Network Operations Control Center (NOCC) subsystems at JPL.

A. The Receiver-Exciter Subsystem

A digital frequency controller, the Programmed Oscillator Control Assembly (POCA), was developed several years ago as part of the Block IV receiver-exciter implementation. The high utility of this device has led to its deployment throughout the Network. The POCA is the key component in the existing uplink frequency control capability. However, it also represents the focal point of most errors made in supporting the uplink, as POCA programming is accomplished by manually entering frequencies, times, and tuning rates. This manual activity must be carried on throughout all Voyager 2 tracks as the POCA does not have sufficient storage to accommodate all parameters necessary to sustain uplink frequency control for the duration of a station track. Not infrequently an error is made in the entry of POCA control parameters.

The automated uplink frequency control implementation will also utilize the POCA. However, operator entries will be eliminated with the establishment of an interface between the POCA and the Metric Data Assembly (MDA). Entries to the POCA will be made automatically by this assembly. In addition, this assembly will advance load the POCA storage registers to maximize the length of time that the uplink may be sustained in the event of a failure of either this assembly or its interface to the POCA.

Manual control of the POCA – should it be necessary – will be provided via a central station control terminal, the Data System Terminal (DST). Feedback of the status of the POCA will be provided to the MDA and DST. A report of frequency, frequency rate, and status are included in the radio metric data stream.

A new digital controller is scheduled to make its debut in the Network about August 1980. This new controller includes increased storage capacity, and will become the prime exciter controller in the 64- and 34-meter subnets. The POCA will be retained at these stations as a backup. The 26-meter Network will not be equipped with either POCAs or the new controllers. Use of the new controller will not functionally alter the uplink frequency control implementation. Figure 2 depicts the planned uplink frequency control capability throughout the Network.

B. The DSS Tracking Subsystem

As mentioned earlier, the MDA plays a key role in uplink frequency control implementation by providing the Receiver-Exciter Subsystem with frequency control parameters. These parameters, based upon uplink frequency control predicts, consist of a start tuning frequency, and frequency rate and time pairs that model in a piece-wise linear fashion, the profile of the doppler modified best-lock frequency.

A significant time and effort savings was realized in implementing this capability in the MDA by virtue of a functionally similar capability recently implemented in the Radio Science System. The Radio Science System implementation included development of a new module of the prediction program designed specifically to support frequency control using a POCA, as well as the logic and algorithms necessary to process the predictions and interface with the POCA.

For the uplink frequency control implementation, the prediction processing and interface logic were modified to provide improved reliability by minimizing dependence on the MDA by the Receiver-Exciter Subsystems. Changes were made in the storage of predicts; the MDA will store three sets of control predicts, each set consisting of eight days of predictions. The MDA will also provide feedback to JPL regarding the successful reception of predictions via the High-Speed Data Subsystem. A local station hard copy of the uplink frequency control predictions as received by the MDA will be provided.

The MDA will provide reports via the radio metric data stream of the POCA status and commanded frequency tuning rate any time a change in either status or rate is detected. These parameters are necessary for both monitoring the performance of the station and for the orbit determination process where the frequency rates caused by tuning must be separated from those of the relative earth-spacecraft motion. Also to be provided in the radio metric data stream are the identification of the predicts set in use for frequency control, and any frequency or time biases that may have been centered to offset this predicts set.

C. The Network Operations Control Center (NOCC) Tracking Subsystem

As previously mentioned, a module of the prediction software implemented to support the Radio Science System was designed to fit linear segments to a given ground received frequency profile. It generates values of frequency and time pairs that approximate the given profile within specified limits and assures that this can be accomplished within the precision limits of the POCA. This is exactly the capability needed for the uplink frequency control implementation. Thus the only change required to the prediction software is to create a new predicts type for uplink frequency control and a new station

destination code. Thus, when uplink frequency control predicts are transmitted to the station they will be routed automatically to the MDA. The actual generation of the data to be transmitted will be accomplished without modification to the existing predicts program algorithm. The frequency, frequency profile, curve fitting constraints, and time spans are all operator controlled inputs.

The Tracking Subsystem Real-Time Monitor (RTM) will provide display of predicts reception and acknowledge messages generated by the MDA. The RTM will also provide volatile displays to the Network Operations Control Team (NOCT) and project users of the POCA status, and the commanded uplink frequency tuning rates. These parameters will also be incorporated into the System Performance Record for more detailed nonreal-time analysis.

III. Summary

The Voyager 2 spacecraft receiver failure requires that the Deep Space Stations continuously tune the uplink frequency to compensate for the effects of doppler. This process is currently accomplished through a process that requires extensive human intervention and is error prone. An implementation within the DSN Tracking System will provide automated uplink frequency control by using electrical interfaces to control the transmitted frequency directly from predictions.

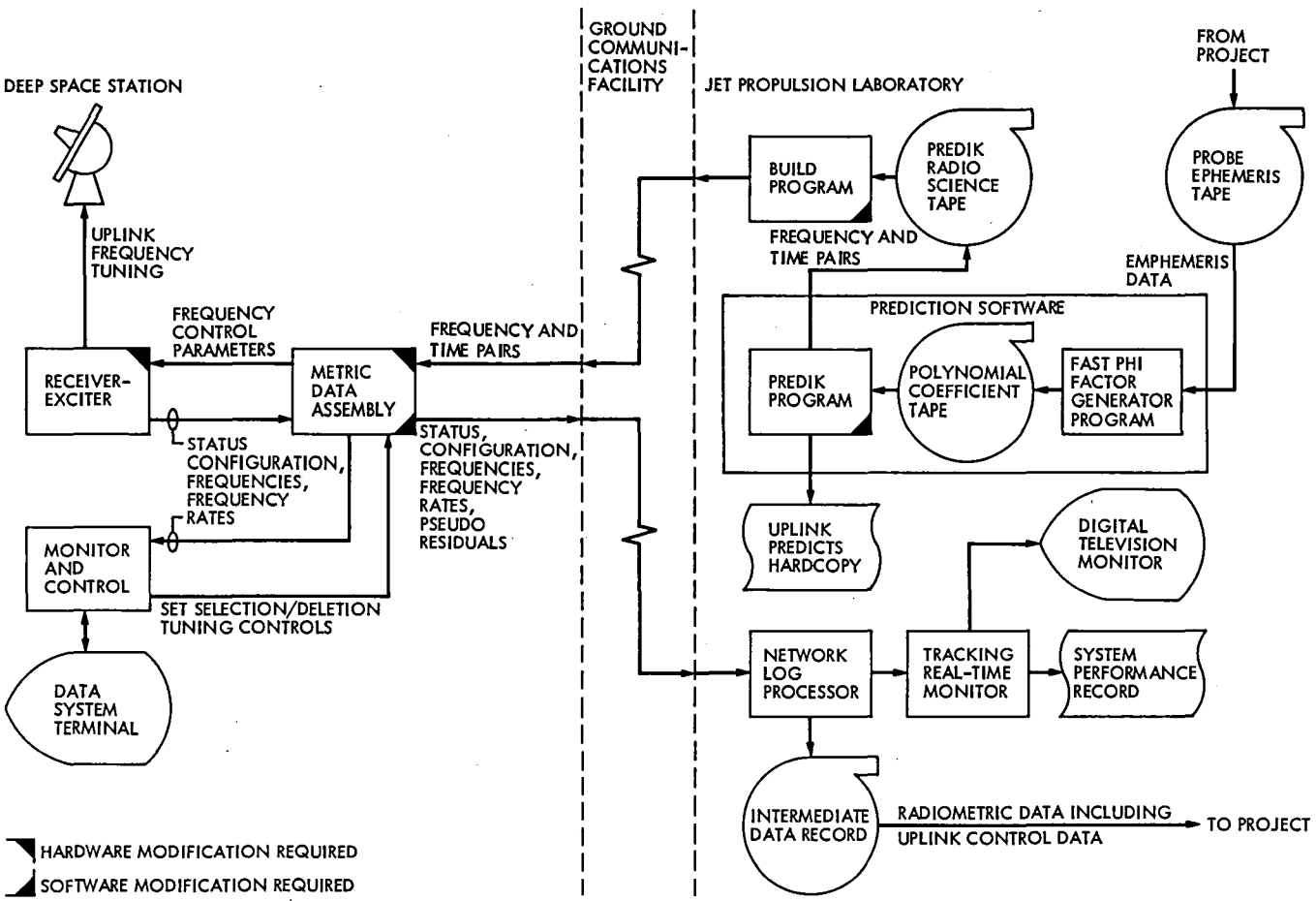


Fig. 1. Uplink frequency control functional block diagram

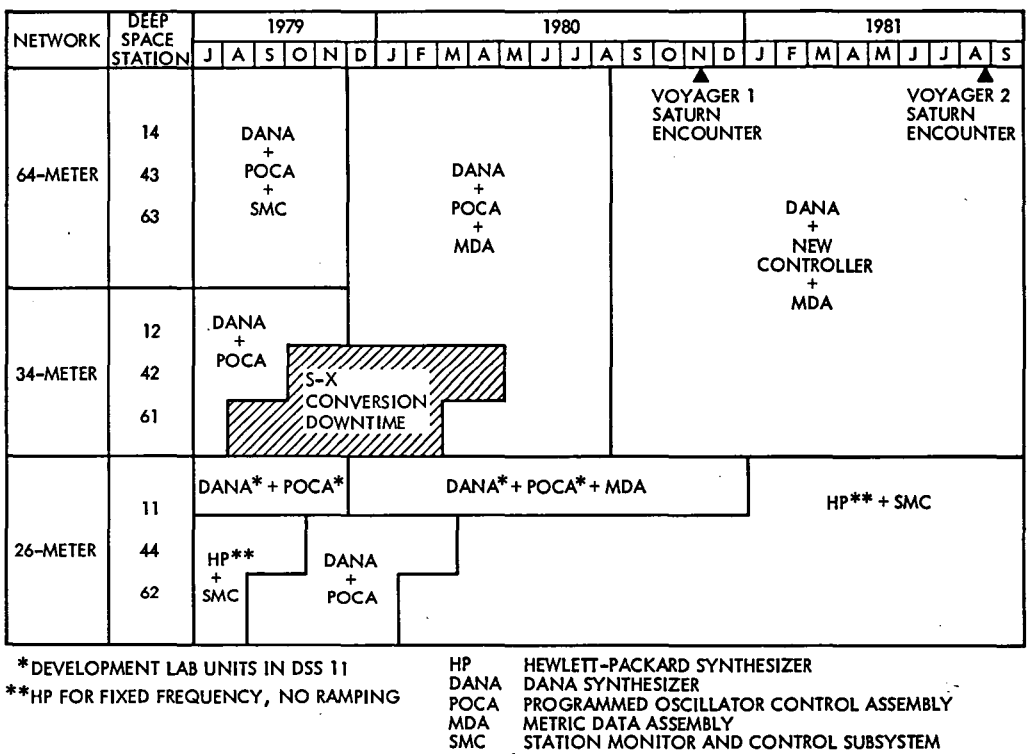


Fig. 2. Uplink control configurations

DSN CONSCAN – A Vector Nomenclature and Method for Determining Parameter Values

T. Taylor
Deep Space Network Support Section

J. Lu Valle
Deep Space Network Operations Section

The DSN CONSCAN algorithm can be conveniently expressed in vector notation. A vector system is developed and used to describe a procedure for determining CONSCAN parameter values.

I. Introduction

The DSN CONSCAN algorithm can be conveniently expressed in vector notation. A vector system is developed in this article and used to describe a procedure for determining CONSCAN parameter values.

CONSCAN requires the specification of four parameter values for proper operation: the CONSCAN radius, the scan period, the loop phase angle, and CONSCAN gain. These values are manually entered into the antenna pointing subsystem (APS) computer prior to CONSCAN tracking.

The values vary from station to station and are different for different tracking configurations at a single station. There is presently no standard procedure for determining these values. The procedure discussed here can fill that gap more efficiently than previous graphical methods.

In the following sections, only geometrical theory is discussed with no consideration of noise or for optimizing closed-loop operation. Those topics are treated in Ref. 1.

II. Theory

CONSCAN causes the antenna pointing vector to describe a small conical scan (hence, the name) around a reference direction. The reference direction is usually moving (e.g., at sideral rate) so that CONSCAN is superimposed on the grosser antenna motion.

During a full revolution about the reference point, CONSCAN correlates the antenna position with the received signal strength of the RF source. This correlation provides an estimate of the source position. In closed loop operation this estimate is used to automatically correct the antenna reference point, or scan center, after the end of a scan. In open loop operation it is used to inform operators of the source position relative to the scan center.

The chief value of CONSCAN is that it maximizes received signal strength to within about 0.1 dB of the peak value of the antenna gain pattern regardless of errors between mechanical and RF boresight. The main weakness is that variations in signal level due to noise or source fluctuations can cause it to track erratically.

In the following paragraphs, a special coordinate system will be developed, and then the CONSCAN algorithm and its parameters will be discussed.

A. Explanation of Coordinates

A DSN antenna is pointed toward an RF source by the time-dependent conputed prediction angles, $\alpha_p(t)$ and $\delta_p(t)$; the hour-angle is driven by $\alpha_p(t)$, and the declination by $\delta_p(t)$. (The angles could be in other spherical systems such as azimuth – elevation or x – y without affecting the validity of the following discussion.)

These predictions are insufficiently accurate to maximize the received signal level, particularly at X-band, which has a very small beamwidth. Therefore the antenna is moved, either manually or by CONSCAN, in an attempt to place the RF boresight nearer the signal source.

For any arbitrary coordinates, α and δ, in the near vicinity of the predicted point, the variables

$$\beta_1 = \alpha - \alpha_p$$

$$\beta_2 = \delta - \delta_p$$

define a nearly rectangular coordinate system if the βs are small enough. Then β_1 and β_2 can be considered the hour-angle (HA) and declination (DEC) components, respectively, of a vector β.

Figure 1 shows the β coordinates as lines of constant β_1 and β_2 superimposed on a small square area of the angular sphere. The predicted point is at the origin.

Figure 1 also illustrates the notation convention to be used, where the points *S*, *C*, and *A* stand for source, scan center, and antenna position respectively. The subscripts, read from right to left, indicate vector direction so that, for example, β_{SC} indicates the vector from the scan center to the source point, and β_{CS} (= $-\beta_{SC}$) is in the opposite direction. Vector addition follows the normal rules so that $\beta_S = \beta_{SC} + \beta_C$.

While the β system is natural for the APS (the numerical values of the β coordinates, if properly scaled, can be used to drive the antenna from the predict point), it is awkward for CONSCAN because of the convergence of hour-angle lines at nonzero declinations. The β_1 coordinates are "squashed," and the computed numerical value of an arbitrary β will not correspond to the great circle angle it spans. This effect can be severe at large declination angles.

A natural coordinate system for CONSCAN is defined by the vector $\mathbf{X} = (X_1, X_2)$ where

$$\left.\begin{aligned} X_1 &= \beta_1 \cos\delta \\ X_2 &= \beta_2 \end{aligned}\right\} \qquad (1)$$

This "unsquashes" the hour-angle component by the factor cos δ where δ is the declination at β ($\delta = \delta_p + \beta_2$).

Under this transformation, the vectors shown in Fig. 1 are not physically changed (that is, they retain the same graphical directions and lengths), but the numerical component values are different so that the computed lengths correspond to the great circle angles required by CONSCAN.

CONSCAN computations can now be made in the **X** system and the results transformed into β when it is time to drive the antenna. For compact notation, Eq. (1) can be written

$$\mathbf{X} = \mathscr{T}\beta; \quad \beta = \mathscr{T}^{-1}\mathbf{X}$$

where

$$\mathscr{T} = \begin{pmatrix} \cos\delta & 0 \\ 0 & 1 \end{pmatrix}; \quad \mathscr{T}^{-1} = \begin{pmatrix} 1/\cos\delta & 0 \\ 0 & 1 \end{pmatrix}$$

The subscript notation will remain the same for both systems.

B. CONSCAN Algorithm and Parameters

CONSCAN drives the antenna clockwise around the scan center, *C*, as shown in Fig. 2. The period of revolution is *P* seconds, where *P* is one of the four CONSCAN parameters. Since there is a 3-second pause at the end of each scan, *P* is usually chosen so that *P* + 3 corresponds to an integer number of minutes (typically one or two minutes).

The parameter *R* is the angular radius of the scan in degrees and is generally chosen so that the antenna power loss, when the source is at a distance *R* from the antenna, is about 0.1 dB.

The vector $\mathbf{X}_{AC}$ represents the antenna position relative to *C*, and has the length *R* in the direction θ as defined in the figure. The source position is represented by *S*, and $\mathbf{X}_{SA}$ is the vector from the antenna to the source. The received power depends upon the angular distance of the source from the antenna; i.e., upon the length of $\mathbf{X}_{SA}$.

CONSCAN computes points (indicated by "i" subscript), by

$$\mathbf{X}_{ACi} = -R\,(\sin\theta_i(t), \cos\theta_i(t))$$

where the angle θ in radians as a function of time is

$$\theta(t) = \frac{2\pi}{P}\,t$$

Points are generated ten times per second; the antenna is driven by the conversion to β coordinates.

$$\beta_{Ai} = \beta_C + \mathscr{T}^{-1}\,\mathbf{X}_{ACi}$$

After a scan is completed, CONSCAN computes an estimate of the vector from the center to the source, $\mathbf{X}_{SC}$(EST), and (in closed-loop operation) the scan center position is updated by

$$\beta_C(\text{NEW}) = \beta_C(\text{OLD}) + \mathscr{T}^{-1}\,\mathbf{X}_{SC}(\text{EST})$$

While the antenna is being driven along its circular path, CONSCAN correlates the position with the received signal strength to find $\mathbf{X}_{SC}$(EST). Except for the effect of signal filtering, to be discussed momentarily, the correlation is the discrete summation:

$$\mathbf{X}_{SC}(\text{EST}) = \frac{2G}{RP}\sum_{i=1}^{N} \mathscr{S}_i\hat{\mathbf{X}}_{ACi}$$

Instead of the summation, we will use the integral

$$\mathbf{X}_{SC}(\text{EST}) = \frac{2G}{RP}\int_0^P \mathscr{S}\hat{\mathbf{X}}_{AC}\,dt \qquad (2)$$

since N, the $10P$ points in a scan, is large.

$\mathscr{S}$ is a voltage (from the receiver AGC circuitry for spacecraft tracking, or from a square law detector for natural celestial sources) that varies with the received signal strength, and $\hat{\mathbf{X}}_{AC}$ (= $-(\sin\theta, \cos\theta)$) is a unit length vector in the direction of $\mathbf{X}_{AC}$. G is a scaling factor.

The integral part of Eq. (2) is simply a vector that points toward the maximum signal, assuming an axially symmetric antenna pattern with no noise; while the factor $2G/RP$ scales the length of this vector. This is illustrated in Fig. 3 where the integral part is represented by **I**, and the value of $\mathscr{S}$ is indicated by the width of the band around the scan center. Then it is apparent that the correlation is very similar to finding the center of mass of a circular ring whose mass (analogous to signal strength) varies along the circumference.

The $(1/RP)$ part of the scaling factor provides normalization, keeping the length of $\mathbf{X}_{SC}$(EST) approximately constant for different radii and periods, while G is essentially a fudge factor that adjusts the length of $\mathbf{X}_{SC}$(EST) to agree with that of $\mathbf{X}_{SC}$(actual). G is called the "CONSCAN gain," and is the third of the four parameters. It is determined empirically, given a particular configuration, and that is the subject of the next section.

Finally, the factor of 2 floating in the scaling factor is rather superfluous given the empirical nature of G, but it is included here for the sake of accuracy. Its origin and purpose are obscure.

Equation (2) would be satisfactory for the CONSCAN correlation except that $\mathscr{S}$ is actually a filtered voltage, so that variations of $\mathscr{S}$ can lag behind the instantaneous signal level changes by an appreciable amount. (The amplitude is also slightly affected, but we ignore that here.) Since the scan is clockwise, this time lag shows up as a clockwise rotation of **I** as shown in Fig. 4 where the entire pattern for $\mathscr{S}$ is shifted. This effect could be compensated by delaying the values of $\mathbf{X}_{AC}$ fed to the integral, but it is easier to just rotate **I** by an angle that will account for signal lag as well as for other system delays.

Then the corrected version of Eq. (2) should be

$$\mathbf{X}_{SC}(\text{EST}) = \frac{2G}{RP}\mathscr{R}_L\int_0^P \mathscr{S}\hat{\mathbf{X}}_{AC}\,dt \qquad (3)$$

where $\mathscr{R}_L$ rotates **I** by the angle L, and is equivalent to the matrix

$$\mathscr{R}_L = \begin{pmatrix} \cos L & \sin L \\ -\sin L & \cos L \end{pmatrix}$$

$\mathscr{R}_L$ is a clockwise rotation for a positive value of L, so that L must always be negative to be commensurate with the DSN CONSCAN implementation. The last of the CONSCAN parameters, L, is also determined empirically.

The actual CONSCAN implementation uses

$$\mathbf{X}_{SC}(\text{EST}) = \frac{2G}{RP}\int_0^P \mathscr{S}\mathscr{R}_L\,\hat{\mathbf{X}}_{AC}\,dt \qquad (4)$$

which is equivalent to Eq. (3) since $\mathscr{R}_L$ is constant with respect to the integration. Equation (4) uses a little more machine time than is necessary, since the rotations are done inside the summation loop rather than outside.

C. Finding G and L

Of the four parameters R, P, L, and G, R and P are usually chosen to fit mission and station requirements, and then G and L are determined empirically to optimize tracking. Typically, the values have been found graphically by plotting CONSCAN data for several different scan centers. This section will develop a method for finding G and L using the vector approach.

Equation (3) is rewritten

$$\mathbf{X}_{SC} = G\mathscr{R}_L\,\mathbf{u} \qquad (5)$$

where

$$\mathbf{u} = \frac{2}{RP}\int_0^P \mathscr{S}\hat{\mathbf{X}}_{AC}\,dt \qquad (5a)$$

Equation (5) is now considered to be exact (not an estimate as before). It is illustrated in Fig. 5 where it is seen that $\mathscr{R}_L$ rotates $\mathbf{u}$, and G scales the length. The same figure shows the relation

$$\mathbf{X}_C = \mathbf{X}_S - \mathbf{X}_{SC}$$

where $\mathbf{X}_S$ and $\mathbf{X}_{SC}$ are a priori unknown.

Substituting from Eq. (5)

$$\mathbf{X}_C = \mathbf{X}_S - G\mathscr{R}_L\mathbf{u} \qquad (6)$$

Now, values for $\mathbf{u}$ can be output from the APS (on the standard print-out) in open-loop CONSCAN either by using values of $G = 1$ and $L = 0$ for the initial setup, or by correcting the printout for any other values. However, Eq. (6) still has four unknowns: the two components of $\mathbf{X}_S$, and the desired values for G and L. It is true that $\mathbf{X}_S$ can be found by doing a manual boresight, but the accuracy is generally poor and the results are subject to error. Instead, we continue to consider $\mathbf{X}_S$ unknown.

By successively tracking at two different scan centers, i and j, two vector equations will be available and sufficient for finding the unknowns.

$$(\mathbf{X}_C)_i = (\mathbf{X}_S)_i - G\mathscr{R}_L(\mathbf{u})_i$$

$$(\mathbf{X}_C)_j = (\mathbf{X}_S)_j - G\mathscr{R}_L(\mathbf{u})_j$$

Differencing these equations and considering that the drift of $\mathbf{X}_S$ is negligible between measurements

$$[(\mathbf{X}_C)_i - (\mathbf{X}_C)_j] = -G\mathscr{R}_L[(\mathbf{u})_i - (\mathbf{u})_j]$$

or

$$\Delta\mathbf{X}_C = -G\mathscr{R}_L\Delta\mathbf{u}$$

Now it is easy to find G and L from

$$G = \frac{|\Delta\mathbf{X}_C|}{|\Delta\mathbf{u}|}$$

$$L = \left[\tan^{-1}\left(\frac{\Delta u_2}{\Delta u_1}\right) - \tan^{-1}\left(\frac{\Delta X_{C2}}{\Delta X_{C1}}\right)\right] - 180\ \text{deg} \qquad (7)$$

where Eq. (7) assumes that the arctangent values are put into the proper quadrants and L is modulo 360 degrees.

If desired, $\mathbf{X}_S$ can also be found by substituting the values for L and G into Eq. (6).

III. Test Results at DSS 12

The procedure was used at DSS 12 on 17 May 1979 using the Voyager 1 X-band signal. Parameters were determined for all three Block III AGC bandwidths, using a radius of 0.008 degree (the beamwidth is approximately 0.1 degree) and a time period of 57 seconds.

A manual boresight was accomplished to locate the approximate position of the source. The APS was then put into open-loop CONSCAN with values of $G = 10$ and $L = -23$ degrees. The value, $G = 10$ was used instead of $G = 1$ in order to provide an extra decimal printout on the APS, which outputs a least digit value of one thousandth of a degree. The value of $L = -23$ was used to approximate previous results for L.

Four points were chosen for the scan center, all about $3R$ from the approximate source position. The value of $3R$ was chosen so that scanning would be on a steeper part of the antenna pattern with consequently less noisy vector outputs than for the flatter portion near the source.

With the manual boresight point established and computationally used as the new origin (or predict point) for convenience, the scan center was moved to the first point, offset from the new origin by $\beta_C = (24, 0)$ thousandths of a degree. Output for several scans was taken, and then the process was repeated for the remaining points at $\beta_C = (-24, 0)$, $(0, 24)$, and $(0, -24)$ successively.

The raw data output from the APS for narrow-band AGC are shown in Table 1, along with the averages and standard deviations. (The first output for each new scan center was eliminated because of errors due to the large antenna movement during the scan.) The averages were then used as the "true" values of $(\beta'_{SC})_i$, the initial estimate of the scan center to the source for scan center "i".

These vectors are now converted to **X** coordinates

$$\mathbf{X}_C = \mathscr{T}\beta_C$$

By comparing Eqs. (3) and (5a),

$$\mathbf{u} = \frac{1}{G'}\mathscr{R}_{-L'}\mathscr{T}\beta'_{SC}$$

where G' and L' are the initial estimates of G and L. The results for the $(\mathbf{X}_C)_i$ and $(\mathbf{u})_i$ for narrow-band AGC are shown in Table 2.

The vectors $\Delta\mathbf{u}$ and $\Delta\mathbf{X}_C$ were then computed between successive points, along with

$$|\Delta\mathbf{u}|,\ |\Delta\mathbf{X}_C|,\ \tan^{-1}\left(\frac{\Delta u_2}{\Delta u_1}\right),\ \text{and}\ \tan^{-1}\left(\frac{\Delta X_{C2}}{\Delta X_{C1}}\right)$$

These values were used to compute G and L for each successive pair of points. The results are in Table 3.

Finally, the values of L and G were averaged. The results for all three AGC bandwidths are given in Table 4. For closed-loop operation, G is usually taken as 0.5 times the open-loop value given so that overshoots due to noise are reduced.

The procedure could have used Δs between all six pairs of points rather than successive pairs. However, drift of the source point can be rapid enough that the difference between, say, points 1 and 4 could lead to erroneous results. In fact, when $\mathbf{X}_S$ was computed using the new parameter values, this drift could be seen and amounted to a total of about 8 thousandths of a degree (0.008 deg) over the one-half hour of data taking for narrow-band AGC. These results are shown in Table 5.

IV. Conclusions

Although development of the vector nomenclature seems tedious, the final result is worthwhile by making the CONSCAN algorithm expressible in a neat, consistent package with the two vector components on an equal footing. Vector relations are more easily visualized than their algebraic equivalents. This can pay dividends for future programming of CONSCAN.

The procedure, as tested, worked adequately and should form the basis for a standard procedure. The data gathering process required about one-half hour for each AGC bandwidth including the manual boresight done between bandwidth changes, so that the total time required (including one-half hour for the initial setup) was about two hours. Most of the data reduction was done using a FORTRAN program written for a Modcomp computer.

Acknowledgment

The authors thank D. Girdner for his help in setting up and running the tests at DSS 12, as well as for his insight into the inner workings of CONSCAN. Thanks are also due J. McClure for his coordination efforts with Voyager and DSS 12.

Reference

1. Ohlson, J. E., and Reid, M. S., *Conical-Scan Tracking with the 64-m-Diameter Antenna at Goldstone*, Technical Report 32-1605, Jet Propulsion Laboratory, Pasadena, California, October 1976.

Table 1. Raw data from APS for four scan centers

i	=	1	2	3	4
$(\beta_C)\,i$	=	(24, 0)	(-24, 0)	(0, 24)	(0, -24)
$(\beta'_{SC})\,i$	=	(-58, 5)	(46, 6)	(-2, -53)	(-10, 79)
	=	(-57, 5)	(53, 0)	(-10, -50)	(-8, 75)
	=	(-53, 6)	(60, 5)	(-7, -48)	(-7, 80)
	=	(-58, 11)	(54, 5)	(-12, -48)	(-11, 69)
	=	(-76, 5)	(55, 0)		
	=	(-61, 12)			
Average	=	(-60.5, 7.3)	(53.6, 3.2)	(-7.8, -49.8)	(-9.0, 75.8)
Std. Dev.	=	(8.0, 3.3)	(5.0, 3.0)	(4.4, 2.4)	(1.8, 5.0)

R = 0.008 degree
P = 57.0 seconds
G' = 10.0
L' = -23.0 degrees

Table 2. Conversion to X coordinates

$$(\mathrm{X}_C)\,i = \mathcal{T}(\beta_C)\,i$$

$$(\mathrm{u})\,i = \frac{1}{G'}\mathcal{R}_{-L'}\mathcal{T}(\beta'_{SC})\,i$$

i	=	1	2	3	4
$(\beta_C)\,i$	=	(24, 0)	(-24, 0)	(0, 24)	(0, -24)
$(\mathrm{X}_C)\,i$	=	(22.67, 0)	(-22.67, 0)	(0, 24)	(0, 24)
$(\beta'_{SC})\,i$	=	(-60.5, 7.3)	(53.6, 3.2)	(-7.8, -49.8)	(-9.0, 75.8)
$(\mathrm{u})\,i$	=	(-4.97, 2.90)	(4.78, -1.68)	(-2.63, -4.30)	(2.18, 7.31)

δ = 19.2 degrees

Table 3. Intermediate computations

Points	Δu	ΔX_C	$\lvert\Delta u\rvert$	$\lvert\Delta X_C\rvert$	$\tan^{-1}\left(\frac{\Delta u_2}{\Delta u_1}\right)$	$\tan^{-1}\left(\frac{\Delta X_{C2}}{\Delta X_{C1}}\right)$	G	L
1, 2	(9.75, –4.58)	(–45.3, 0,)	10.8	45.3	- 25.2	180.0	4.2	-25.2
2, 3	(–7.41, –2.62)	(22.7, 24)	7.9	33.0	-160.5	46.6	4.2	-27.1
3, 4	(4.81, 11.61)	(0, –48)	12.6	48.0	67.5	- 90.0	3.8	-22.5

Table 4. Final values of *G* and *L* by AGC bandwidth

Narrow	Medium	Wide
$G = 4.1 \pm 0.2$	3.7 ± 0.2	3.7 ± 0.0
$L = -24.9 \pm 2.3$	-0.1 ± 4.9	-2.5 ± 2.7

Table 5. Source drift

Point	X_s	Time since boresight
Boresight	(0, 0)	0 min
1	(–0.7, 2.2)	13
2	(–2.1, 2.0)	19
3	(–2.3, 3.6)	24
4	(–4.5, 6.8)	29

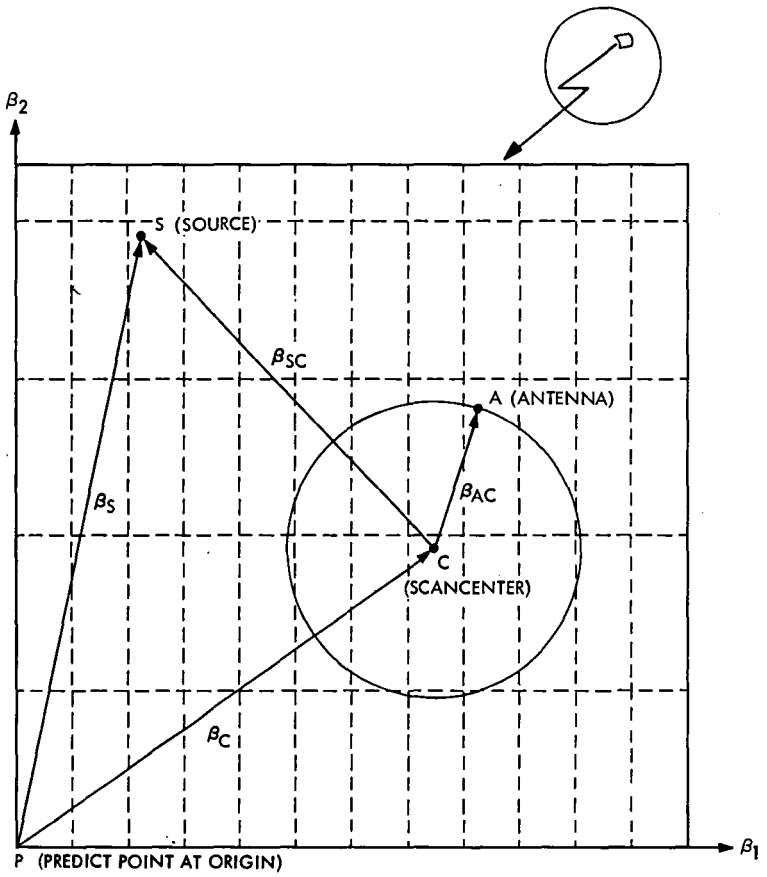


Fig. 1. β coordinate system

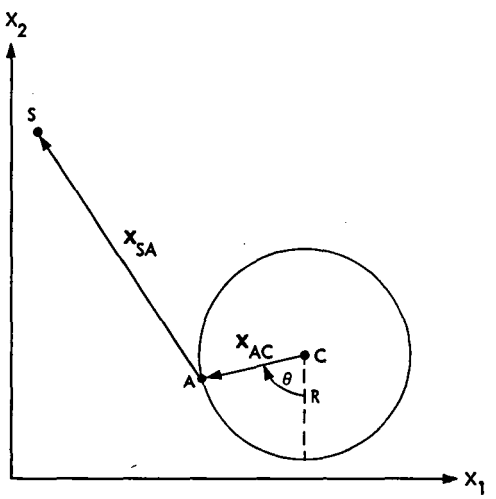


Fig. 2. The scan circle

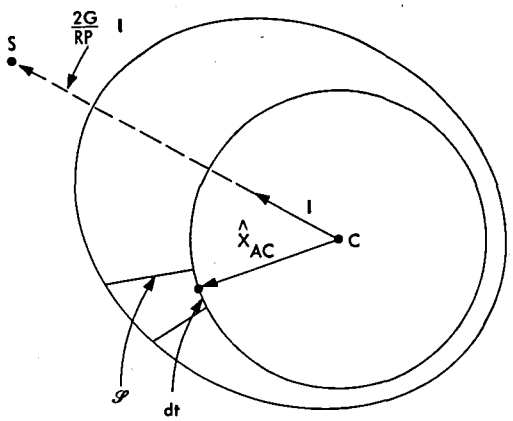


Fig. 3. Signal level variation around the scan circle

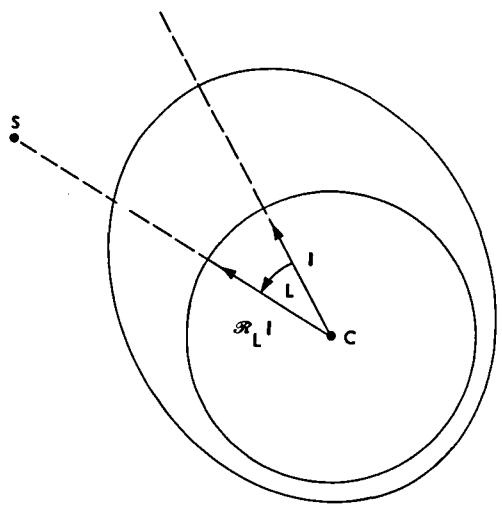


Fig. 4. Effect on filtering on signal level variation

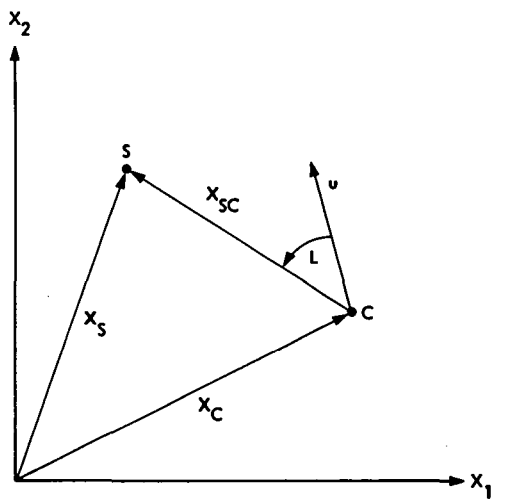


Fig. 5. Vector relationships for determining G and *L*

Antenna Azimuth Bearing Model Experiment

H. McGinness
DSN Engineering Section

A reduced scale model of an antenna wheel and track azimuth bearing indicates that its prototype would have a long life and require little maintenance.

I. Introduction

During the past few years various wheel and track configurations have been studied in order to appraise their suitability for use as azimuth bearings of large antennas. A particular design comprising a single circular track and a novel wheel suspension system appeared to have unique advantages, not only for the antenna application, but for many applications involving a flat roller on a flat track. An analysis of the suspension system is given in Ref. 1.

A reduced scale model of an antenna azimuth bearing employing this concept was constructed. The following describes the model and the test results obtained so far.

II. Purpose of the Model

The model was designed to test several independent things pertaining to an antenna azimuth bearing wherein the supporting wheels are attached to the alidade base and the track is grouted onto a concrete foundation. Among these are the wheel and track resistance to contact surface failure, grout strength, circumferential movement of track on its supporting grout, effect of a large temperature difference between the wheel supporting structure and the track, and the amount of the rolling resistance torque. The model was designed to produce a wheel cyclic loading approximately 3200 times as fast as that of an antenna prototype.

The failure mode of a wheel and track usually is spalling of the contact surface caused by the propagation of fatigue cracks. The maximum shear stress in the neighborhood of the contact surface varies as the square root of the applied load, therefore it is efficient to use high strength materials for both the roller and track. Commercial roller bearings use extremely strong steel of Rockwell C hardness 60 or more. Since the high stress regions exist only near the surface, it would be prudent to design large wheels and tracks to have hard surfaces on relatively soft mild steel in order to minimize material costs. If the track is composed of a continuous ring made of mild steel and covered with replaceable hardened wear strip segments, advantage can be taken of the high material strength in the high stress region. This permits the ring or track runner to be made of much less expensive mild steel, and makes possible a fast replacement of a damaged or badly worn wear strip. For large wear strips it seems best from an economical standpoint to use medium hardness values, say, from 35 to 40 Rockwell C. Although there is a great amount of information on the fatigue failure of both very hard and very soft steels, such as commercial bearing steel and mild carbon steel respectively, there is a paucity of fatigue information on steels in the hardness range of 35 to 40 Rockwell C, when subjected to the cyclical loading experienced by a wheel rolling on a flat track. Ordinary rotating beam cyclical loading tests are deemed inappropriate for the determination of an endurance limit for roller bearing stresses, since in the bearing the peak tensile and compressive stresses are of entirely different absolute values.

The conclusions reached by two independent studies conducted for JPL by bearing engineers, Refs. 2 and 3, were that it is highly desirable that model tests be made at the nominal Hertz stress level of 86184 (10^4) N/m^2 (125,000 psi) using surface hardnesses between 35 and 40 Rockwell C.

The track ring or runner must be made sufficiently flat and level and then grouted to its concrete foundation. The loads imposed on the grout and concrete by a rotating antenna are unusual in that the load at a point is completely on when a wheel is directly above it, and is completely off when the wheel is a short distance away from this point. In most civil engineering structures, a large percentage of the total load is steady. A further complication is that the grout or concrete under the runner will likely experience some tensile stress and this is difficult to evaluate by analysis. The existence of tensile stress may be understood by considering the problem of a small loaded area on the surface of a semi-infinite body, the solution of which gives surface tension just outside the loaded area. The analogous semi-infinite plate problem solution does not give any surface tensile stresses. The actual problem at hand lies somewhere between these two cases. The model tests should establish whether a cement grout is adequate for the proposed loads.

A phenomenon that has been observed with wheel-track systems is that of circumferential displacement of the track with respect to the grout or foundation. This will hereinafter be called track walking. At first glance it might seem that the high friction between the bottom of the track and the grout would be sufficient to resist the torque applied to the runner, and indeed it is. However, if the track is considered as a beam on an elastic foundation and without any shear connection to the grout, other than friction, it will be seen that the beam is displaced upward on both sides of a loaded wheel. If there is no tensile adherence between beam and grout there will be a space between the two. As the wheel advances there can be a rolling or ironing action of the track on the grout, causing the track to advance circumferentially. Whether this action occurs probably depends upon the shear strength between the track and grout away from the immediate vicinity of the wheel. It is important to determine if this action occurs, and if so, to develop adequate auxiliary restraints to prevent it.

Various empirical formulas have been employed to calculate the torque required to rotate a large wheel supported structure. The question arises whether the methods used to derive the torque resistance of a small roller thrust bearing would apply to the wheel and track system. The effect of misaligned wheels on the azimuth torque has been calculated. A model could confirm these theories.

During periods of maintenance or repair, antennas usually are not rotating in azimuth. If during this parked condition a differential temperature occurs between the track and the alidade base structure, which supports the wheels, there likely will be a radial displacement of the wheel with respect to the track. It was not known how much wheel and track damage would result from this scuffing action.

III. Description of the Model

The model being tested is as shown in Figs. 1 and 2. The track runner is a mild steel circular ring of 1.77-m mean radius. Its rectangular cross section is 30.48 by 10.16 mm. Each of eighteen hardened steel wear strips of 2.54-mm thickness and 30.48-mm width is fastened to the ring with 14 screws. The ring is then centered and leveled, and dry-pack cement grout placed in the 9.50-mm space between the ring and concrete pedestal.

The alidade base is simulated by the H-beam frame shown in Fig. 1. Each wheel truck is pivoted to the frame by two ball joints, one of which is adjustable by opposing set screws. This adjustment allows the truck to be rotated about a vertical axis for best wheel-to-track alignment.

The center pivot is a spherical roller bearing. Its inner race is attached to a steel pedestal. Its outer race housing is fastened to a diaphragm that joins a cylindrical shell fastened to the H-beam frame. The flexible diaphragm insures that all the ballast weight loading goes onto the wheels and not onto the center pivot bearing. A sprocket of pitch diameter 0.869 m is attached to the lower edge of the diaphragm-cylinder structure and a roller chain connects it to the drive unit mounted on the floor. Four screw jacks in the corners of the frame are connected by a chain so that one person can operate them in unison, thus allowing convenient wheel alignment and inspection. The flexible diaphragm also allows the frame to be jacked without the necessity of removing the center bearing.

The drive unit output sprocket has a pitch diameter of 0.1317 m and is driven by a 43:1 ratio reducer powered by a 0.75 kW, three-phase electric motor running at 1750 rpm. This gives the H-beam frame an angular speed of approximately 6 rpm. The motor can be driven in either direction.

The wheel loads come from the H-beam frame weight and from the ballast weights stacked onto the frame. Originally it was planned to have a wheel truck in each of the four corners of the H-beam frame. It was thought that the torsional flexibility of the frame would ensure a uniform distribution of loading among the four wheel trucks, provided the flatness tolerance of the track ring was met. Although each cross

section of the track is level when viewed as a vertical radial plane, there are some slopes in circumferential directions. Therefore it was decided to employ three trucks only, making the reactions statically determinate, and adjust the ballast weights so as to give equal truck loadings when rotating at 6 rpm. The effects of centrifugal force on the reactions was considered in adjusting the ballast weights.

When two wheels were used on one truck, the proper angle between the two wheels was established in the metrology laboratory by an optical method. The truck assembly was subsequently aligned to the frame. The outer end of each wheel axle was made square with respect to the wheel axis. A special mirror with parallel surfaces could be spring loaded against the axle end by running a small coil spring through a hole in the axle. One end of the spring was attached to the back side of the mirror and the other end was pinned against the inner end of the axle. An autocollimating jig transit was positioned approximately one meter outside the track and at the elevation of the wheel. The instrument was aligned to a vertical pin extending upward from the center bearing pedestal. Then the frame was rotated until a wheel with mirror attached intercepted the line of sight. The truck was rotated about a vertical axis, by adjusting the opposing set screws on the ball pivot joint, until autocollimation on the wheel mirror was obtained. It is believed that by this method the wheel alignment can be known to within 1 arc minute.

The wheel suspension system is shown in Fig. 2. Each wheel bearing housing frame is connected to the truck frame by a pair of flexure struts at each side of the wheel. The flexure struts extended would intersect at the top of the track surface, thus they act as truss members and are capable of resisting both vertical and horizontal forces applied at the wheel-track contact point. The resistance to a moment, however, is very small. This means that the interface moment between the wheel and track will be very small thereby precluding any high edge loading of the wheels even though the wheels and track may have a relative misalignment. This is an important advantage because in practice it is difficult to prevent a relative angular displacement between the wheel and track. Such displacement, for example, may come from needed manufacturing tolerances, or from a warping of the track foundation. The interface moment magnitude is discussed in Refs. 1 and 9.

There are disadvantages associated with this suspension system. Its intrinsic complexity is tempered somewhat by the fact that it can use smaller wheels because of the nearly uniform loading achieved across the width of the wheel. Perhaps its principal disadvantage compared to conventional designs is its smaller stiffness in the vertical direction. This is due to the small cross-sectional area of the flexure struts. In many instances, however, this may reduce a vibrational frequency by less than 10 percent.

The wheels of the model are tapered so as to be true rolling elements. The mean diameter of the wheel is 50 mm and its width is 12 mm. As originally designed, it was a 1/20 scale model of a 100-m diameter antenna, but can be considered also as a larger scale model for smaller antennas.

IV. Scale Effects

A model geometrically similar to the prototype and properly loaded in the same manner will have the same stresses as the prototype. If the model length is λ times the prototype length, the proper loads to be applied to this model, in order to match the prototype stresses, are λ^2 times the prototype loads. Much significant information can be obtained from a reduced scale structural model. However, it must be realized that the scaling does not apply to the material grain size and probably not to surface finishes. Therefore, the statistical fatigue failure performance of models may differ appreciably from that of their full-scale counterparts. This is known to be true for the spalling failure of commercial rolling bearings. The maximum shear stress, τ_{MAX}, in the neighborhood of the cylindrical roller and track contact is: (from Ref. 4)

$$\tau_{MAX} = 0.179\sqrt{\frac{PE}{DL}} \tag{1}$$

where

P is the normal force between roller and track

E is the common modulus of elasticity of the wheel and track

D is the roller diameter

L is the roller length

From this equation it is clear that the model stress will match the prototype stress if the model is geometrically similar, has the same elastic modulus, and is loaded so that:

$$P_{MODEL} = \lambda^2 P_{PROTOTYPE} \tag{2}$$

Rearranging Eq. (1), we obtain:

$$P = \frac{DL}{E}\frac{\tau_{MAX}^2}{(0.179)^2} \tag{3}$$

Reference 5 presents an empirical formula that establishes the dynamic capacity of commercial roller bearings. The dynamic capacity is defined as that load which at least 90 percent of a large sample of bearings will endure for 1,000,000 revolutions without exhibiting any fatigue-induced spalling of the rollers or races. As applied to a roller thrust bearing, which corresponds to the case at hand, the equation is:

$$P_1 = L^{7/9} D^{29/27} f_c \tag{4}$$

where f_c is a dimensional factor having the dimensions of force/(length)$^{50/27}$. For a geometrically similar model made of the same material as its prototype, the relationship between P_1 model and P prototype, derived from Eq. (4) is:

$$P_{1\ \text{MODEL}} = \lambda^{50/27} P_{\text{PROTOTYPE}} \tag{5}$$

Dividing Eq. (5) by Eq. (2), we obtain:

$$\frac{P_1}{P} = \frac{1}{\lambda^{4/27}} \tag{6}$$

which is the factor by which the reduced scale model load must be increased in order to be a fatigue model of the contact area. Since Eq. (4) pertains to commercial bearing steels that are very hard, and the subject model is made of medium hardness steel, it is not known how accurate the load factor of Eq. (6) would be when applied to the subject model. Equation (4) as applied to commercial roller bearings has been experimentally verified by a large number of tests conducted by various experimenters (Ref. 6). However, there is a theoretical rationale supporting it. This is based upon Weibull's premise that as material volume is increased, the number of potential crack sources is also increased. Lundberg modified this concept to apply to the highly stressed region near the roller-track contact. It would appear that the factor represented by Eq. (6) would not be unconservative when applied to steels of reduced hardness.

V. Model Tests

The tests conducted so far have employed six wheels each loaded to 5560 N (1250 lb) which produced Hertz stresses of 86184 (10^4) N/m^2. The first test ran a total of 40.60 hours, half of which was in the clockwise direction. Since the model turns 375.55 turns per hour, this was a total of 15,247 turns. It is estimated that the DSN antennas make approximately 1000 turns per year, thus the first test corresponds to about 15 years of antenna operation as far as the wheels are concerned. Regarding the antenna track, the test corresponds also to 15 years of operation provided the antenna has six wheels, and to a greater or smaller amount depending upon the number of antenna wheels. Most of the first test was run with no lubricant on the track.

Very small pits soon occurred on the wear strip and wheel surfaces. It was judged that these were not from fatigue but resulted from small particles being broken from the sharp edges of the wear strip mitered ends. The particles were distributed around the track by the wheels. When this action was understood, the wear strip ends were stoned to a small radius and the pitting action seemed to cease.

The wear strips were reground slightly in preparation for the second extended run. Upon removal of the wear strips, it was observed that there was fretting corrosion between the wear strips and their steel support ring. Various lubricants, namely, SAE No. 30 engine oil, bearing grease, chassis grease, thread lubricant, MoS_2, dry graphite, and a commercial product called Cortec were placed on the bottom surfaces of the wear strips. During the second extended run, some of these lubricants were extruded through the wear strip mitered joints and smeared over the upper surface of the wear strips, thus producing an inadvertent lubrication on the wheel track surface. This second test was run for a total of 81.54 hours, approximately half of which was in the clockwise direction. The total time for the two tests was 122.14 hours or 45,869 revolutions. There was no evidence of fatigue failure of wheels or wear strips. Of the several lubricants used between the wear strips and their support ring, only the thread lubricant (Silver Goop) and the Cortec VCJ-309 were completely effective in preventing fretting. Various platings will be tried on future tests.

A. Track Walking

After three hours of counterclockwise rotation of the model, it was observed that there was no circumferential displacement of the track with respect to the runner, i.e., there had been no track walking. After an additional 3.75 hours in the same direction, it was observed that the track had moved 17.8 mm in the counterclockwise direction. After 1.13 hours more of rotation in the same direction, the walking displacement had reached 25.4 mm. In the 4.88 hour period in which the 25.4-mm displacement occurred, the model made 1833 turns, and the average displacement per turn was 0.0139 mm. Initially there was a tensile bonding between the track ring and the cement grout. A short distance from the loaded wheels there was a tensile stress at the bond, per Winkler beam theory. After three hours of cyclical loading the bond failed in tension, thus allowing the track to separate slightly from the

grout. The loaded wheels then produced an ironing or smoothing action on the track, which caused the "wrinkle" to advance in the direction of motion.

Four mild steel tangent links, each having a cross-sectional area of approximately 21 square millimeters, were then installed to connect the track ring to the concrete foundation. Each link was attached to the track ring with one 4.8-mm diameter steel screw in shear. After 1.5 hours of running, one of these screws had yielded and was on the verge of breaking. It is probable that this one screw was resisting the entire walking force since the anchor nuts in the concrete at the other links were loose. Four larger tangent links having a cross sectional area of 91 square millimeters were installed using 6.35-mm diameter screws. These have not failed after many more hours of testing.

B. Grout

After 40.6 hours of operation the track ring was raised from the grout enough to make its surface visible over a distance of approximately 0.30 m. There were no loose particles and no evidence of the track ring having slid over the grout. This verified the concept that the circumferential displacement of the track was of the nature of a rolling action on the grout, rather than sliding.

After an additional 49.13 hours of model rotation, it was observed that the track ring over an angular distance of 80 degrees had been displaced outward, with respect to the grout by as much as 1.4 mm. The outer shoulder of the grout was broken away over this 80 degrees; however, the grout beneath the ring appeared to be sound. The reason for this is not presently understood. Misaligned wheels can exert a large radial force on the track, but it would seem that the frictional resistance between grout and track would be sufficient to resist the induced force from a misaligned wheel. It may be that when circumferential walking is prevented by tangent links, the ironing action of the loaded wheels tends to produce a radial displacement of the track ring. It is important that this effect be understood.

The maximum grout compressive stress, σ_{BR}, has been calculated by the following equation, derived from Refs. 7 and 8:

$$\sigma_{BR} = \frac{0.306P}{(1-\nu^2)^{0.2775} h^{5/6} b^{7/6}} \left(\frac{E_F}{E}\right)^{0.2775} \quad (7)$$

where

P is the wheel vertical load

ν is Poisson's ratio of the foundation

h is the beam (track runner) depth

b is the half width of the beam

E_F is the elastic modulus of the foundation

E is the elastic modulus of the beam

This equation may be used for comparing the calculated compressive stress on the grout with various standards such as building codes, etc. As stated before, the nature of the cyclical loading is uncommon and it is difficult to correlate these calculated compressive stresses to endurance stresses as determined by other methods of cyclical loading.

If there should be a horizontal force applied at the wheel-track contact surface as a result of a misaligned wheel (Ref. 9), the maximum compressive stress on the grout will be more than that given by Eq. (7). This is the primary reason for maintaining good wheel alignment in the plan view.

C. Frictional Torque Resistance

For the case of properly tapered wheels, the frictional torque resistance, T, can be considered as composed of three parts, namely, the rolling resistance of the wheels on the track, the resistance of the wheel bearings, and the sliding resistance caused by misaligned wheels. Each of these components is proportional to the product of total weight, W, and the mean wheel radius, R; therefore, the total frictional torque is:

$$T = WR\left[f_1 + f_2\frac{d}{D} + \mu \sin\theta\right] \quad (8)$$

where

f_1 is the factor for roller thrust bearings

f_2 is the factor for radial ball bearings on the wheel axles

μ is the coefficient of sliding friction

θ is the misalignment angle of the wheels in the plan view

d is the wheel bearing bore diameter

D is the wheel diameter

Using values from Ref. 5 for f_1 and f_2 and setting d/D = 0.30, μ = 0.30, and sin θ = 0.00015, which is the expected value of the wheel misalignment, we obtain from Eq. (8):

$$T = WR\,[0.0011 + 0.0015(0.30) + (0.30)\,0.00015]$$

$$= WR\,(0.0016) \qquad (9)$$

The drive chain was temporarily removed and the model rotated at a uniform angular velocity of approximately 6 rpm by applying a tangential force at a radius of 1.84 m. The force was applied through a force scale and the values varied between 53 and 71 N. Using the average of 62 N, the torque was 114 N · m. The total weight, W, was 33188 N and the mean wheel radius was 1.775 m. From these values, the measured friction coefficient, f_m, is:

$$f_m = T/WR = 114/33188(1.775) = 0.0019 \qquad (10)$$

which is to be compared with the factor 0.0016 of Eq. (9).

Subsequently all the wheels were misaligned by approximately 13.7 arc minutes. This was accomplished by turning each alignment set screw one half turn. The torque required to rotate the model at approximately 6 rpm varied between 158 and 190 N · m. If we use the average of 174 N · m, Eq. (10) yields:

$$f_m = T/WR = 174/33188(1.775) = 0.0030 \qquad (11)$$

Calculating the torque T from Eq. (8) by using the new value of θ = 13.7 arc seconds, we obtain:

$$T = WR\,[0.0011 + 0.0015(0.30) + (0.30)0.00398]$$

$$= WR\,(0.0027) \qquad (12)$$

The value 0.0030 of Eq. (11) is to be compared with the factor 0.0027 of Eq. (12).

Therefore it would appear that the methods of estimating the frictional torque of small bearings can be applied to large bearings and reasonably accurate results obtained.

D. Traction Capacity

The traction capacity was measured by tying the model turntable to the building structure and applying torque to a wheel until it slipped. The ballast weights were adjusted until the wheel reaction was 1214 N. A 6.35-mm diameter bolt was put through the hole in the wheel axle and a nut tightened onto it. The hexagon head of the bolt was then torqued with a torque measuring wrench until slippage occurred. The values recorded were the maximum ones before slippage and correspond to the static capacity. The static coefficient of friction was calculated from the following equation:

$$\mu_s = t/rw \qquad (13)$$

where

t is the measured torque

r is the radius of the wheel

w is the vertical load on the wheel

First the wheel and track were wiped clean with acetone and several torque measurements made. The calculated value of μ_s ranged from 0.351 to 0.395.

Second the wheel was lubricated with SAE No. 30 nondetergent engine oil. The calculated value of μ_s was 0.153.

The wheel and track were wiped with paper, thus removing most of the oil. Then several tests were made. The calculated values of μ_s ranged from 0.242 to 0.286.

It was recommended in Ref. 3 that a small amount of lubrication be provided by means of rubbing a piece of oil impregnated plastic against the track. It is believed that this kind of lubrication would approximate the lubrication obtained when the SAE No. 30 oil was "wiped off".

E. Thermal Expansion

If a heavily loaded nonrotating wheel is forced to slide transversely across the track, some damage will be incurred. For the loads and expected displacements under consideration, it was not known how serious the damage would be. The model turntable was restrained from rotation and the H-beam members of the turntable base were heated by playing torches over their surfaces. A dial indicator was mounted between the track and the end of a wheel axle. At this time all the wheel loads were 5560 N, corresponding to a Hertz stress of 86184 (10^4) N/m^2. The heating continued until the dial indicator read 1.40 mm. During the heating period, the displacement of the dial indicator was not continuous but incremental and accompanied by popping sounds. Two and a half hours after the heating ceased, the indicator read 0.63 mm. Eighteen hours later, when the H-beam members surely were back to their original temperatures, the dial indicator read 0.51 mm. The model was then jacked up exposing the smear marks at all six wheel positions. The width of the marks was approximately equal to the calculated contact area width and the length was approximately equal to the wheel width plus 1.40 mm. Although the marks on both wheels and wear strips were clearly visible, they could not be felt with a finger nail. Examination with a 10X magnifier revealed little damage, which could be polished out easily. Therefore it can be concluded that infrequent transverse sliding of nonrotating wheels across the track will produce only negligible damage.

F. Conclusions

From the tests conducted so far the following conclusions can be made:

(1) The track must be restrained circumferentially if portland-cement-type grout is used. Four tangential links of moderate size are adequate to resist the walking action.

(2) The torque required to overcome the rolling resistance of the wheels can be estimated fairly well by Eq. (8). Since the rolling resistance is usually small in comparison to other resistances, the accuracy of Eq. (8) should be satisfactory.

(3) The traction capacity has been measured and found to be consistent with values obtained from other sources.

(4) The infrequent transverse sliding of a nonrotating wheel across the track produces only negligible damage when the contact stresses and material properties are as herein stipulated.

(5) The wheels, wear strips, and grout have survived 45,869 model turntable revolutions with no sign of fatigue failure. Additional tests must be run at higher wheel loadings in order to compensate for scale effects.

References

1. McGinness, H., *Analysis of a Suspension System for a Wheel Rolling on a Flat Track*, Publication 78-43, Jet Propulsion Laboratory, Pasadena, Calif., Aug. 1978.
2. Rumbarger and Dunfee, *The Evaluation Analysis of Wheel and Track Type Azimuth Bearings for Large Antennas*. Prepared by the Franklin Institute Research Laboratories per JPL Purchase Order No. GK-676593, Dec. 1977.
3. Sibley, L.B., et al., *Azimuth Thrust Bearing Evaluation and Study for 100 m Antenna*. Prepared by SKF Industries Technology Services Division, per JPL Purchase Order No. 954968, Mar. 1978.
4. Timoshenko, S., *Theory of Elasticity*. McGraw-Hill, 1934, page 350.
5. Palmgren, A., *Ball and Roller Bearing Engineering*. SKF Industries, 1959, page 82.
6. Harris, T. A., *Rolling Bearing Analysis*. John Wiley & Sons, 1966, page 340.
7. Biot, M. A., "Bending of an Infinite Beam on an Elastic Foundation." *Journal of Applied Mechanics*, Mar. 1937, page A-6.
8. Hetenyi, M., *Beams on Elastic Foundation*. The University of Michigan Press, 1946, page 207.
9. McGinness, H., "Lateral Forces Induced by a Misaligned Roller," *DSN Progress Report*, March and April 1978, page 253.

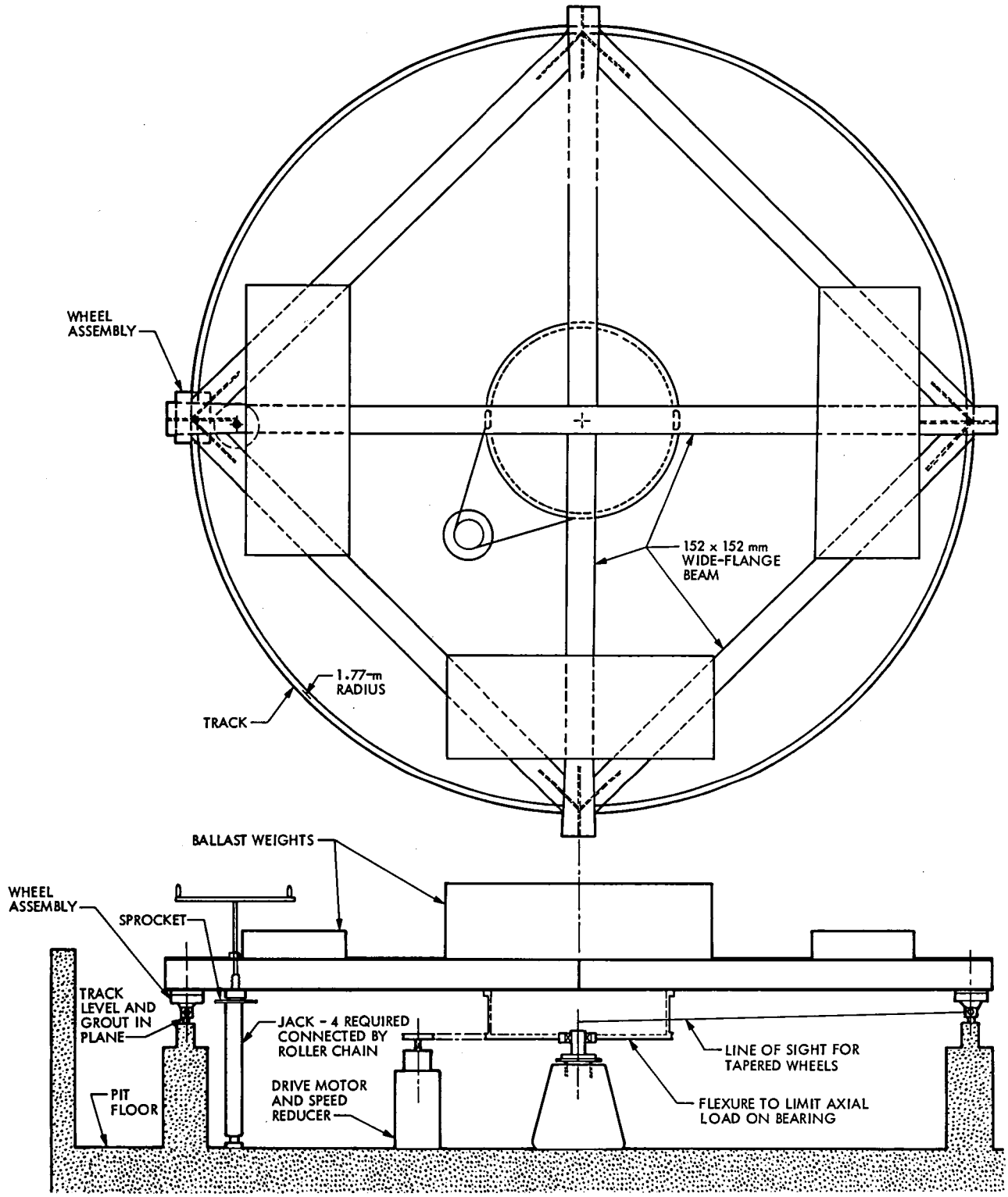


Fig. 1. Antenna azimuth bearing model, general arrangement

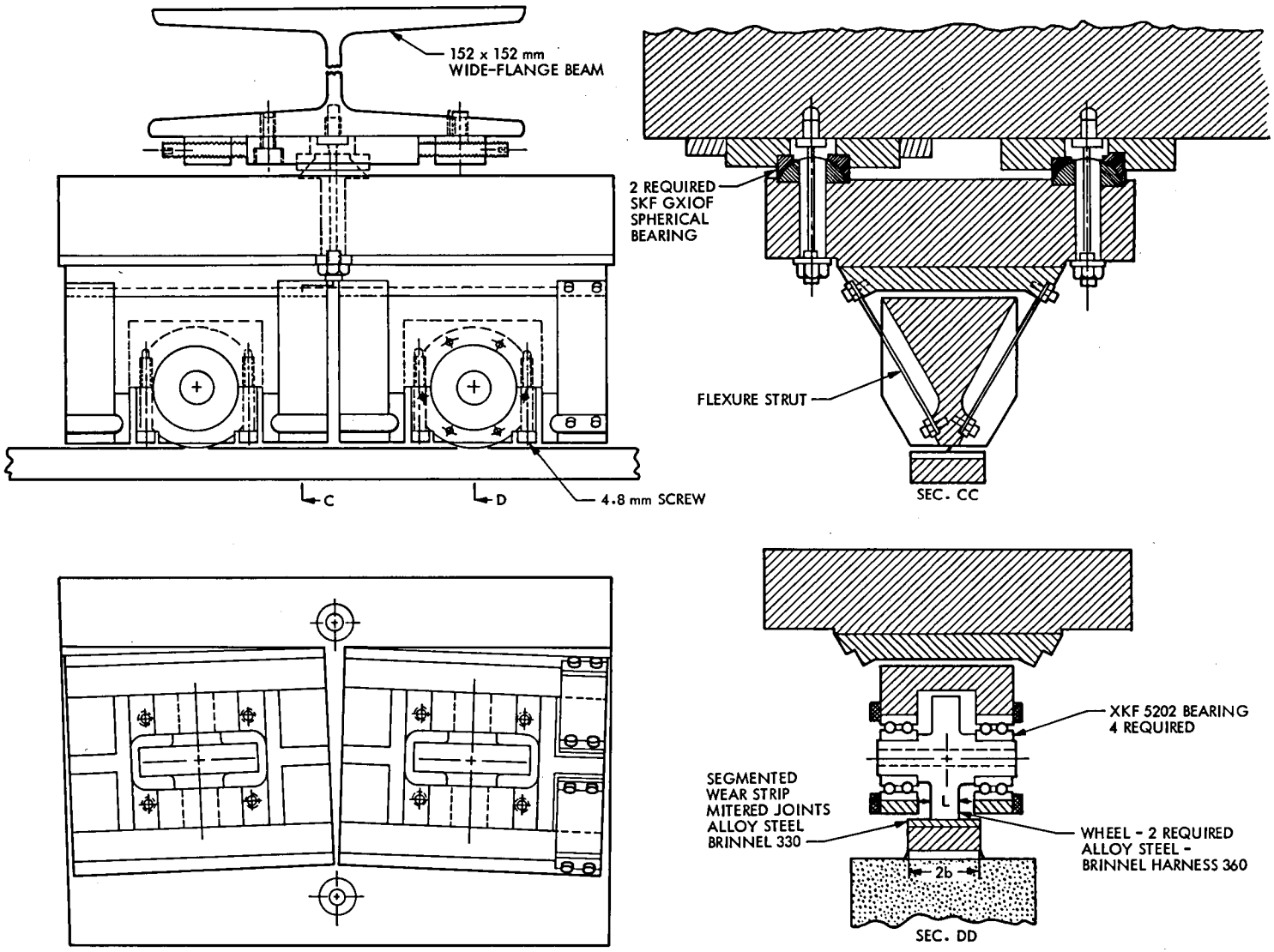


Fig. 2. Wheel suspension system

Radio Frequency Performance of DSS 14 64-m Antenna at X-Band Using a Dual Hybrid Mode Feed

A. J. Freiley
Radio Frequency and Microwave Subsystems Section

The 64-meter antenna X-band system at DSS 14 was evaluated to determine the performance with the new dual hybrid mode feed. The peak system efficiency increased from 42.0 to 45.6 percent resulting in a 0.36 dB increase in antenna gain. The new measured gain is 71.6 dB. Antenna pointing, beamwidth, optimum subreflector focusing, and operating system temperature were unchanged from the previous feed. Some evidence of antenna aging is apparent (from the 1973 measurements reported earlier) both in peak gain and in the pointing angle at which the peak occurs.

I. Introduction

The X-band radio frequency performance of the 64-meter antenna at DSS 14 was measured at 8420 MHz (3.56-cm wavelength) to evaluate the improved antenna gain expected from a dual hybrid mode feed, part of an overall Voyager enhancement program within the DSN. The total Voyager enhancement task is to realize a 2 dB improvement in signal-to-noise ratio (SNR), to improve the imaging science return from the Voyager mission at Saturn and beyond. To achieve this goal the task was separated into three areas: (1) 1.1 dB SNR improvement by real-time arraying with an X-band 34-m collocated aperture; (2) 0.6 dB increase in SNR by reduced noise temperature of the X-band traveling wave maser (TWM); and (3) 0.3 dB increase in 64-m antenna gain through the use of a dual hybrid mode feedhorn to achieve improved aperture illumination.

It is the intent of this article to examine the change in antenna performance due to the installation of the new feed. The remaining aspects of the Voyager enhancement task will be reported elsewhere. Additionally, DSS 14 will have a complete resetting of all main reflector panels as well as a new improved subreflector. The final DSS 14 antenna gain should therefore be 0.5 dB above the already improved levels discussed herein.

To evalute the antenna gain performance, radio metric measurements of selected radio sources were undertaken at X-band prior to, and just after, the installation of the new feed. The radio sources were selected to maintain consistency with previous measurements. While absolute accuracies are always problematical in this work (inaccuracies on an absolute scale are most likely larger than ±0.3 dB), the gain *differences* reported here are considered highly accurate as well as having high resolution (hundredths of a dB).

II. Antenna Modification

The X-band Receive Only (XRO) feedcone has been reconfigured to provide improved performance and capabilities in support of the Voyager mission (see Ref. 1 and 2). The functional block diagrams of the previous XRO system configuration and the new dual hybrid mode system configuration are presented in Fig. 1. Figure 1(a) shows the major functioning

components of the previous XRO system in use prior to October 1978. The dominant features of this configuration were a 22 dB corrugated conical feed horn, selectable polarization and a single low-noise, traveling wave maser (TWM) at 8 K. Figure 1(b) shows the major functional components for the new dual hybrid mode feed. The modification consisted of a new feed assembly with a dual hybrid mode horn, quarter-wave polarizer, orthogonal mode transducer and dual traveling wave masers. The dual hybrid-mode feed horn (Ref. 3) improved the aperture illumination which was expected to increase the antenna system efficiency or gain. The quarter-wave polarizer, orthogonal mode junction, and the dual TWM's provide the capability to route right or left hand circular polarization (RCP or LCP) to either of the TWMs and also provide the simultaneous reception of R and LCP signals to fulfill a radio science polarization requirement. This combination also allows redundancy to receive signals of one polarization if spacecraft system problems arise.[1] The tests reported herein are using the original (8 K) maser. Later installation of the improved (3.5 K expected) machines is yet to be achieved.

III. Technique

The radio metric technique used to evaluate the new system consisted of four parts. The first is the boresighting of the antenna to ensure the RF beam peak of the antenna is coincident with the radio source. The Antenna Pointing System (APS) utilizes the conical scan technique with a total power radiometer to automatically determine the boresight changes and update the pointing commands to maintain optimum pointing (Ref. 4). Conical scan period, radius, gain, and bandwidth are critical parameters that must be set to ensure adequate pointing of the antenna. To minimize tracking errors for radio source observations the scan radius should be near the half power points of the main beam for maximum sensitivity to pointing errors. Limitations on the scan radius and scan period are imposed by encoder errors and RF/IF system gain fluctuations which effect tracking accuracy. The conical scan gain constant varies inversely with the strength of the radio source observed, so caution should be employed when observing more than one source in rapid succession. Selecting the optimum set of conical scan parameters for each source is an iterative process within certain limits until adequate pointing is achieved.

[1]The Voyager spacecraft is unique in that, for efficiency reasons, redundant final amplifiers at X-band are "hardwired" to RCP and LCP flight orthomode ports. Thus, an amplifier failure on the spacecraft results in a switchover which includes a polarization change as well as the amplifier change *per se*.

At X-band the focal length change of the 64-m antenna due to large-scale structural deviations induced by gravity are a critical parameter to be monitored during precision gain measurements. Gain errors of 0.5 to 1 dB are likely if attention is not given to focusing. Therefore, while tracking a radio source, the optimum focus setting was determined by stepping the subreflector through 1 to 2 inches of travel near the optimum position. The focus is then updated to this new position prior to taking noise temperature measurements.

The third part of the radio metric technique is the actual antenna gain or efficiency measurement. The radio metric technique consists of the on-off source operating system temperature measurements (T_{op}) using the noise-adding radiometer (NAR) (Ref. 5). From two off-source and one on-source measurement, the increase in system temperature due to the radio source, ΔT_a (antenna temperature) can be determined for a particular elevation angle. The antenna efficiency at that elevation angle is the ratio of antenna temperature, ΔT_a, to the source temperature, T_s (the temperature observed by a 100 percent efficient antenna). By observing sources at various elevation angles the antenna efficiency with elevation can be characterized. A more detailed description of the on-off measurements and the calculation to obtain antenna efficiency is given in Ref. 6.

The calibration of the NAR is an important part of the measurement scheme. The ambient load is the well-calibrated physical temperature standard used to calibrate the NAR noise diode. NAR calibration is achieved when the NAR measurement of the system temperature on the ambient load equals the physical temperature of the load plus the maser and receiver follow-on temperature.

IV. Radio Source Calibrators

The standard radio source calibrators – 3C48, 3C123, 3C274 (Virgo A), 3C295, 3C380 and DR21 – were selected to maintain consistent calibrations both before and after the feed replacement. The assumed flux density and other related source parameters are given in Table 1. The assumed flux densities, S, are the result of a series of radio source ratio measurements, reported in Ref. 7. The source temperature, T_S, is the standard value on which the system efficiency is based and is determined from the flux density at the frequency of interest for a 100 percent efficient antenna by the equation

$$T_S = \frac{S A_p}{2k}$$

where A_p is the physical area of the antenna and k is Boltzmann's constant (1.380622 X 10^{-23} W/Hz K). The typical corrected peak antenna temperatures are also listed for reference in any future work. The source resolution connection, C_r, was applied to the measured antenna temperature to correct for the systematic error resulting from partial resolution of the radio source. The relatively strong sources, 3C274 and DR21, were the standard sources predominantly used to determine the optimum focus settings for this series of antenna measurements. All of the listed sources are considered to exhibit stable flux levels with time and were used to determine the gain and efficiency of the antenna system.

V. Radio Metric Data

The radio metric measurements at DSS 14 were taken to determine the increase in antenna gain and other parameters due to the installation of the modified XRO feed system. Between July and October 1978, baseline performance of the antenna system was defined (with difficulty) and during November and December 1978, the improved performance was determined.

The antenna pointing system generally performed well during the observations. A comparison of pointing offsets prior to and following the feed modification indicated no notable change as determined by the conical scan technique. The beamwidth of the antenna at X-band (8420 MHz) was 0.038 degrees (137 arc sec) and remained constant before and after the feed replacement. The scan radius used to boresight the antenna ranged from 0.015 to 0.020 degrees and the scan period was about 60 seconds. At one point during the baseline performance observations, the master equatorials I and II were not in operation, causing unacceptable pointing variations. The observations during this session exhibited large fluctuations resulting in the cancellation of the remaining portion of that observing session. Data gathered at this time was not included in the baseline performance. Without the master equatorial, the antenna pointing system is *not capable* of maintaining pointing accuracies necessary for high quality X-band antenna use.

The subreflector axial focus of the antenna as a function of elevation angle was measured using the NAR. While tracking on a radio source the subreflector was positioned in steps of 0.1 inches through a 1- or 2-inch range about the optimum position. At each step the NAR would sample the operating system temperature and by inspection the operator could determine the optimum focus setting at that elevation. Each focus measurement characterized the gain loss due to defocusing of the antenna, the peak of which is the point of optimum focus. As part of the normal data reduction, the system temperature data was curve fitted and normalized to the peak allowing interpolation between the steps to determine the exact setting; corrections for the systematic errors were also applied to the measurements of system efficiency. Typical corrections for improper focus ranged about 1 to 2 percent. Figure 2 shows the typical normalized gain loss data. The defocusing gain loss is −3 dB when the subreflector offset from the optimum is about one wavelength (1.2 to 1.4 inches). This elaborate measurement was one of a few which characterized the defocusing gain loss. Normally 5 to 7 steps are sufficient to determine the optimum axial focus setting. The defocus gain loss measurements were conducted at various elevations and were found to have no significant elevation-dependent variations, nor was there any change or variation due to the installation of the new feed. No noticeable change in the shape of the defocus gain loss curve was expected and none was observed. The optimum axial focus position for DSS 14 as a function of elevation angle is given in Fig. 3. No bias change in the optimum focus was detected due to the feed change. Winds beyond about 15 miles per hour (24 kilometers per hour) have been noted as affecting the axial focus at X-band. Winds blowing into the aperture (low-elevation angle case) cause the optimum focus to move in the outward direction, which is consistent with a more shallow reflector. Automatic axial focusing of the 64-m antenna subreflector (as a function of elevation angle) is being implemented as part of the substantial overall upgrade of the 64-m network in readiness for the Voyager Saturn encounter. These measurements should be used as design data describing the optimum focusing function of the DSS 14 subreflector. Comparable measurements from the other 64-m antennas are needed since the optimum axial focus function will be unique to each antenna.

At the core of the measurement of antenna parameters is the determination of the operating system temperature. During the course of these tests the operating sytem temperature was measured using the normal DSN Y-factor instrumentation and the NAR. Initially the measurements of operating system temperature varied from session to session and agreement between the Y-factor and NAR measurements was poor during a given session. When the system is operating normally and correctly the results of each T_{op} measurement system (Y-factor and NAR) should agree to within the accuracies of the measurements. The Y-factor instrumentation T_{op} measurement varied from 28 to 34 K while the NAR T_{op} measurements ranged from 24 to 28 K. Maser tuning and receiver linearity were found to be the two major causes of the disagreement. Once the problems were understood the T_{op} measurements were stabilized by setting maser tuning parameters (gain and bandwidth) more precisely and reducing the maser gain to improve receiver linearity over the operating range (from ambient load $T_{op} \approx 300$ K to operating system $T_{op} \approx 25$ K). Maser gain was nominally maintained at 42 dB. The operating system temperature at 8420 MHz prior to the

feed modification was about 29 K at zenith; following the implementation of the new feed T_{op} at the same frequency also was about 29 K (Y factor measurements). An early XRO feed (1973) with the dichroic plate system operating with a 23 K operating system temperature (Ref. 8). No assured explanation is available for the higher system temperature indicated during the past year at DSS 14; however the Y factor noise instrumentation is suspect and may require work.

The intent of the feedhorn changes was to improve the antenna gain or efficiency of the 64-m antenna. Figure 4 shows the measured increase in overall antenna system efficiency. The system efficiencies in this figure are as would be observed in spacecraft and radio science missions; the measurements have not been corrected for atmospheric loss effects nor the waveguide loss effects, but have been corrected for effects of source size (C_r in Table 1). Therefore, Fig. 4 includes the normal atmospheric extinction. The original feed performance peaked at 52 degrees elevation with a measured system efficiency equaling 42.0 percent (July 1978). The measurement scatter of this data set was attributed to the T_{op} measurement difficulties previously mentioned. The curve fit is biased to best describe the base performance in spite of the measurement difficulties encountered during that period. The dual hybrid feed performance also peaked at about 52 degrees elevation with a system efficiency of 45.6 percent (November, 1978). The lower measurement scatter is due to the well-controlled linearity and maser tuning and is more typical of this class of measurements. With the installation of the dual hybrid-mode feed, the antenna system efficiency increased by 3.6 percent, which will increase peak antenna gain by 0.36 dB. The antenna gain of DSS 14 at 8420 MHz has increased from 71.27 dB (42.0 percent efficient) to 71.63 dB (45.6 percent efficient) as determined by this series of measurements.

When comparing the July 1978 performance (original feed) with that of 1973 (Ref. 8), the system efficiency with elevation has pronounced differences in shape. The 1973 measurements show a peak near 40 degrees elevation with a system efficiency of 42.6 percent. The July 1978 measurements show a very slight loss in peak efficiency (42.0 percent) but a more important shift in the elevation angle at which the gain peak occurs, as well as reduced gain at low-elevation angles. This is a serious effect (about 0.5 dB gain loss at 20 degrees elevation) compared with the 1973 performance. Presumably, the scheduled resetting of the main reflector panels at DSS 14 will correct the low-elevation angle problem which might be interpreted qualitatively and loosely as "aging." Nevertheless, the dual hybrid mode feed performance definition is totally based on 1978 measurements, and the new feed yields the anticipated improvement (Ref. 3).

VI. Conclusions

The performance of the new dual hybrid mode feed installed on the 64-m antenna at DSS 14 has been evaluated at X-band (8420 MHz) using radio metric techniques. The peak system efficiency at 52 degrees elevation increased, as expected, from 42.0 to 45.6 percent resulting in a 0.36 percent increase in antenna gain over that measured with the original feed. The shape of the efficiency curve with elevation angle is noticeably different from 1973 measurements. The antenna pointing was examined and no significant shifts were detected. The operating system temperature was not affected by the feed change remaining at about 29 k (Y-factor measurements) at zenith; it remains higher than expected but relatively stable during later phases of the measurements reported here. Additional work on the noise instrumentation is required; the Y-factor instrumentation is prone to system linearity difficulties.

References

1. Hartop, R. "New X-Band Microwave Equipment at the DSN 64-meter Stations," in *The Deep Space Network Progress Report 42-48*, pp 126-128, Jet Propulsion Laboratory, Pasadena, California.

2. Hartop, R. W., "New X-Band Antenna Feed for DSN 64-meter stations," in *The Deep Space Network Progress Report 42-52*, pp. 71-74 Jet Propulsion Laboratory, Pasadena, California.

3. Thomas, R. F., Bathker, D. A., "A Dual Hybrid Mode Feedhorn for DSN Antenna Performance Enhancement," in *The Deep Space Network Progress Report 42-22*, pp. 101-108, Jet Propulsion Laboratory, Pasadena, California.

4. Ohlsen, J. E., Reid, M. S., "Conical-Scan Tracking with the 64-m-Diameter Antenna at Goldstone," JPL Technical Report, pp. 32-1605, Pasadena, California, October 1, 1976.

5. Batelaan, P. D., Goldstein, R. M., Stelzried, C. J., "A Noise-Adding Radiometer for Use in the DSN," *JPL Space Program Summary*, Vol. II, pp. 37-65, Jet Propulsion Laboratory, Pasadena, California.

6. Freiley, A. J., Batelaan, P. D., Bathker, D. A., "Absolute Flux Density Calibrations of Radio Sources: 2.3 GHz," JPL Technical Memorandum 33-806, Jet Propulsion Laboratory, Pasadena, California, December 1, 1977.

7. Turegano, J. A., Klein, M. J. "Calibration Radio Sources for Radio Astronomy: Precision Flux Density Measurements at 8420 MHz," (to be published).

8. Freiley, A. J., "Radio Frequency Performance of DSS 14 64-m-Antenna at 3.56 and 1.96 cm Wavelengths," JPL Technical Report 32-1526 Vol. XIX pp. 110-115, Jet Propulsion Laboratory, Pasadena, California.

Table 1. Radio source calibrations

Source	Flux density 8420 MHz $S^{(a)}$ (Jansky)	Source temperature (100% efficient antenna) T_s (Kelvin)	Typical antenna temperature T_a (Kelvin)	Source resolution correction C_r	Source position (1950.0)	
					Right ascension (hr-min-sec)	Declination (deg-min-sec)
3C48	3.33	3.88	1.77	1.000	01 34 49.8	32 54 20
3C123	10.05	11.7	5.34	1.007	04 33 55.2	29 34 14
3C274	46.0	53.6	24.5	1.085	12 28 17.3	12 39 48
3C295	3.69	4.30	1.96	1.000	14 09 33.5	52 26 13
3C380	4.87	5.68	2.59	1.000	18 28 13.5	48 42 41
DR21	21.4	25.0	11.4	1.030	20 37 14.2	42 09 07

[a]Flux density values from Ref. 7.

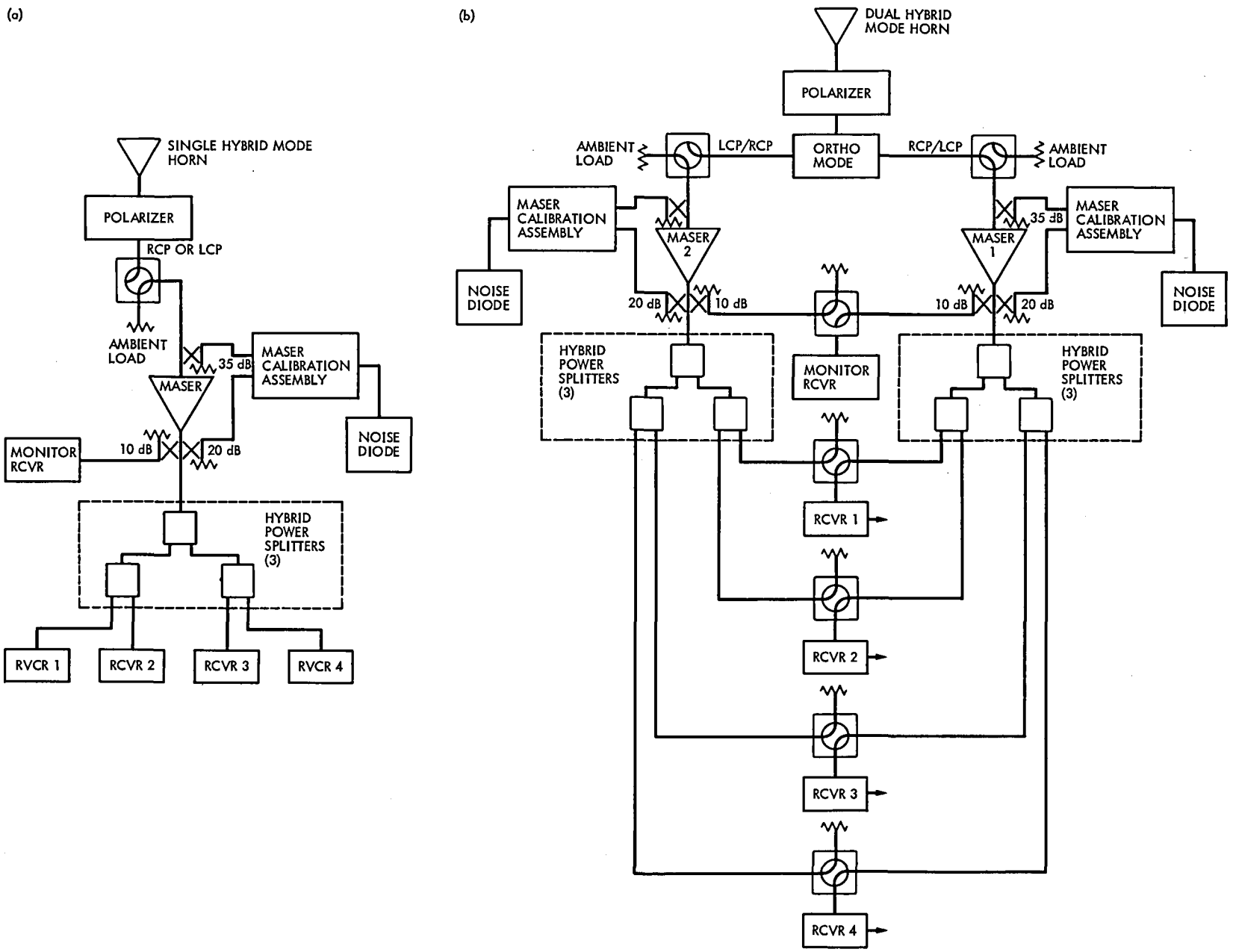


Fig. 1. Functional block diagram of (a) original XRO feed system (prior to October 1978), and (b) modified XRO feed system (after October 1978)

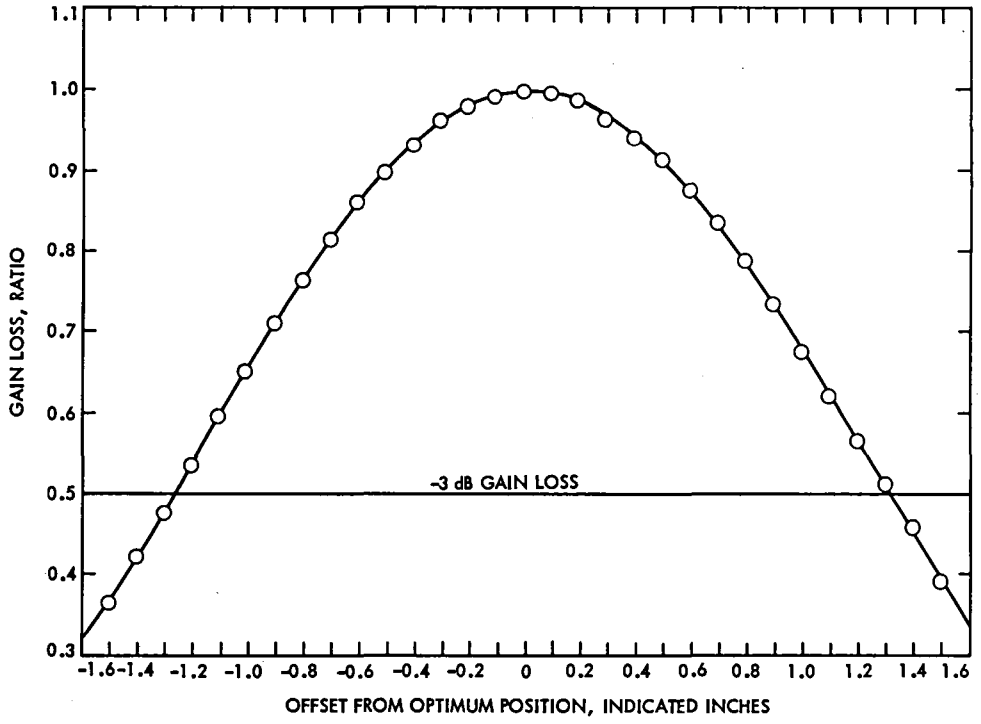


Fig. 2. X-band normalized defocused gain loss as a function of offset from optimum position

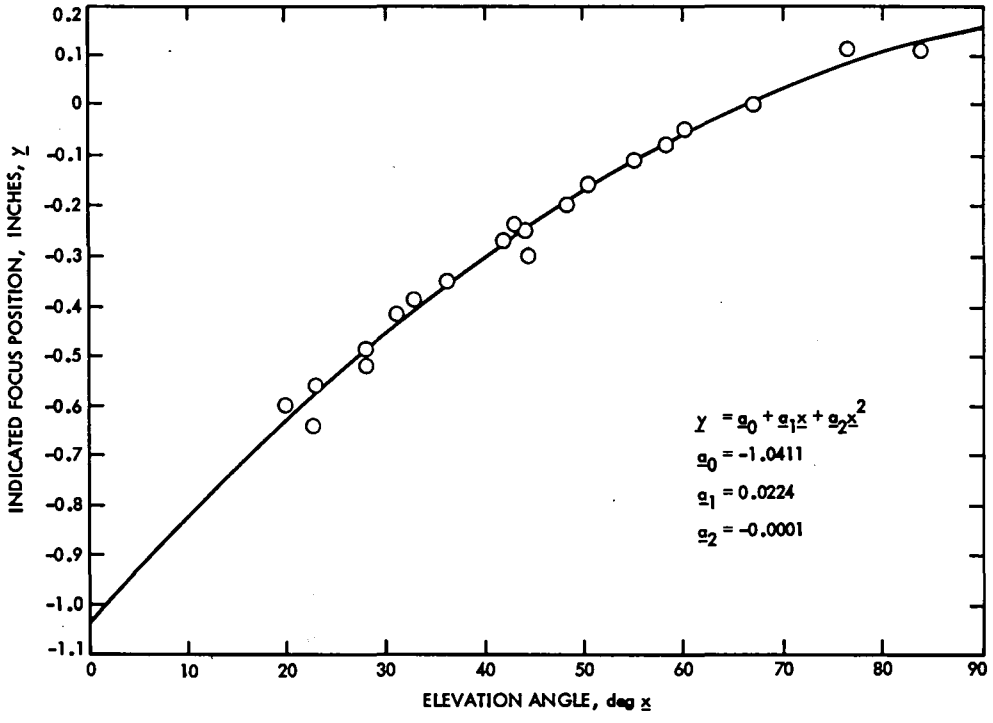


Fig. 3. X-band optimum axial focus position as a function of elevation angle

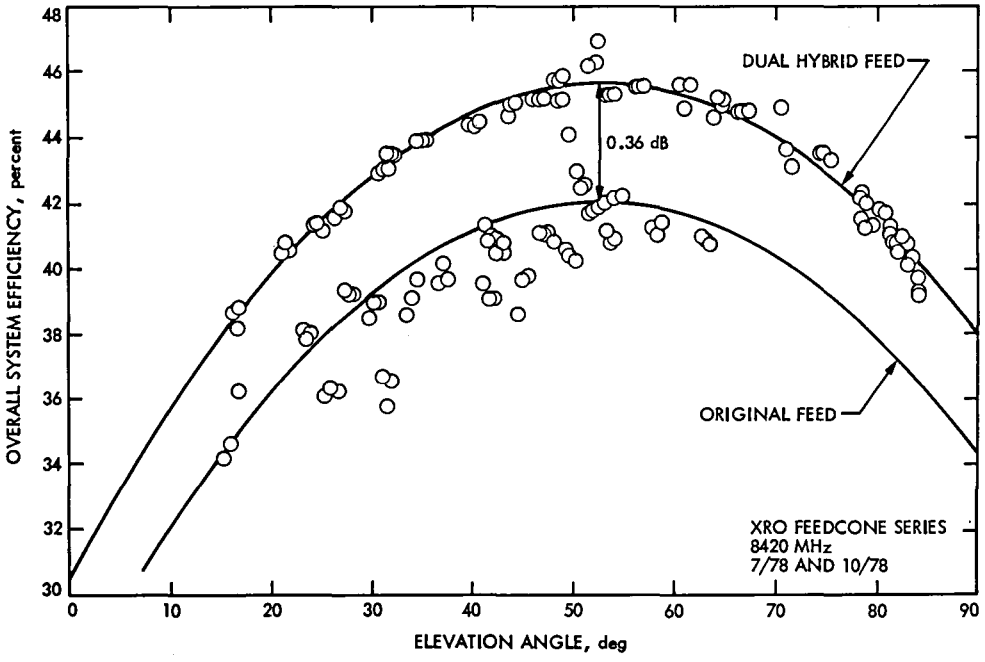


Fig. 4. DSS 14 overall antenna system efficiency at X-band

An Experimental Investigation of the Effects of Antenna Pointing Errors on Range Delays

T. Y. Otoshi
Radio Frequency and Microwave Subsystems Section

T. H. Taylor
Deep Space Network Support Section

The effect of antenna pointing errors has been a possible source of error on range and VLBI time delay measurements that has not been previously investigated. This article presents the results of some S-band tests performed specifically to study this effect. The test procedure involved the intentional mispointing of a 26-meter antenna while ranging to a zero delay device (ZDD) connected to a collimation tower antenna. The observed range changes were less than 0.25 meters when the collimation tower target was within the 3 dB points of the main beam of the 26-meter antenna.

I. Introduction

A possible source of error on ranging and VLBI time delays is the unknown change in range (time) delays occurring when the antenna pointing angle changes from boresight (on target) to an angle offset from boresight. In this article, the boresight angle is defined as the pointing angle at which the target is on the peak of the main beam. An offset from boresight can occur if the antenna pointing system is inaccurate, or when an error has been made in locating the target boresight position, or when using conical scanning.

To study the effects of mispointing the antenna, far-field ranging tests were attempted with the DSS 14 antenna and the Viking Orbiter Spacecraft during April 1977. The results of these tests were unsatisfactory because of a large differenced range minus integrated doppler (DRVID) noise level of about 7 ns (corresponding to 1 meter of range delay). For the test data to be meaningful, it was necessary that range changes of 0.1 meter could be observed. Integrating data over longer periods to minimize noise did not succeed because the data processing at that time did not adequately remove the charged particle delay variation due to the earth's ionosphere and solar plasma.

Recently, some success has been obtained by DSS 42 personnel with collimation tower measurements for ranging. The collimation tower at DSS 42 is located approximately 3587 m (2.23 miles) from the 26-m antenna, or $0.7\,D^2/\lambda$ away at 2295 MHz where D is the 26-m dish diameter and λ is the wavelength. Although this collimation tower is not in the far-field of the 26-m antenna, some of the earlier data at DSS 42 showed that good agreement had been obtained between the optical (survey) delay and measured delays to the collimation tower. It became increasingly apparent that if the collimation tower method could be proven to have minimal effect from ground multipath and near-field errors, then an antenna

pointing test could be done with the DSS 42 collimation tower similar to that attempted previously with the Viking Orbiter Spacecraft. Such a test was subsequently designed and requested of DSS 42. The purpose of this article is to document the test procedure and test results of the pointing error test that was successfully performed by DSS 42 personnel with the Mu-II ranging system and the collimation tower and the 26-m antenna at DSS 42.

II. Test Requirements and Procedure

The purpose of the test was primarily to determine the effects of antenna pointing errors on range delays and secondarily to determine if a collimation tower could be used for this and other types of measurements. The test setup consisted primarily of (1) a 2.44-m- (8-ft-) diameter parabolic antenna with a zero delay device installed on top of the DSS 42 collimation tower and (2) a 20-kW transmitter, a 26-m antenna, a ranging machine, and all of the tracking hardware and software normally used for station delay calibrations and spacecraft tracking.

The test plan was as follows:

(1) Configure the 26-m antenna for RCP polarization and the ranging modulation for 3-dB carrier suppression.

(2) Point the 26-m antenna to the collimation tower and perform range measurements at the following elevation angle offsets (in degrees) from boresight:

0.00, 0.24, 0.22, 0.20, 0.16, 0.12, 0.08, 0.04, 0.00, -0.04, -0.08, -0.12, -0.16, -0.20, -0.22, -0.24

At each offset angle, record the AGC signal levels in dBm and range values corrected by a suitable number of DRVIDs.

(3) Restore the antenna to the collimation tower boresight and repeat the measurements for the same offset values in azimuth.

(4) With the antenna pointed at boresight, use a step attenuator in the S-band line to the ZDD on the collimation tower and step the attenuator so that the signal drops 2, 4, 6, 8, 10, and 12 dB from that obtained at boresight. Record AGC and range values corrected by suitable number of DRVIDs. This step is required for correcting range versus signal level changes for the tests performed above.

If a step attenuator at the collimation tower is not practical, then use the dish-mounted ZDD to obtain a received signal level versus range calibration curve at approximately the same signal levels as the collimation tower measurements.

III. Test Results

The test procedure described in the previous section was capably carried out by DSS 42 personnel and the results are shown in Figs. 1 and 2. The test results shown in the upper half of the figures indicate that when the target was within the 3-dB points of the 26-m antenna, the range changes were less than 0.25 m peak-to-peak for either elevation or azimuth angle offsets. At the larger angle offsets, most of the larger range changes could be explained by the fact that the signal level becomes weak and data becomes increasingly corrupted by noise. For example, ranging power-to-noise ratio was about 15 dB at -160 dBm AGC signal level as compared to 35 dB at -140 dBm.

Specific test parameters of importance to note for the plots of Figs. 1 and 2 were RCP polarization, 3-dB carrier suppression, and test frequencies of Channel 11 (2112.4 MHz uplink and 2294.0 MHz downlink). The test was done at these frequencies because the collimation tower ZDD local oscillator had already been set up for Channel 11.

The boresight coordinates of the DSS 42 collimation tower antenna for the test were assumed to be AZ = 312.324 deg and EL = 3.857 deg, which corresponds to HA = 59.028 deg and DEC = 30.566 deg. However, as may be seen in the AGC boresight curves in the figures, these coordinates were in error by about 0.032 degree in elevation and 0.050 degree in azimuth.

Further discussion should be given at this time concerning the AGC boresight curves. The AGC curves show the signal level change to be twice as large as the antenna pattern gain change as a function of offset angle. This result can be explained by the fact that the test was done for a round-trip measurement. In this particular test setup, there is a signal loss on the uplink frequency signal when mispointing the 26-m antenna at the collimation tower at a look angle off the peak of the main beam, and if the downlink frequency is close to the uplink frequency, then the same loss occurs when the translated signal returns from the collimation tower to the 26-m antenna at the same offset angle. Since the uplink frequency (2112.4 MHz) and downlink frequency (2294.0 MHz) were sufficiently close to each other, the 6-dB points on the AGC received signal level versus offset curves were assumed to be the approximate 3-dB points of the 26-m antenna pattern.

IV. Discussion of Data Reduction and Error Limits

Some of the specific details of test and data reduction will now be discussed. The Mu-II Ranging System (Ref. 1) was fortuitously available at DSS 42 at the time of this test and, therefore, was used in place of the Planetary Ranging Assembly. In the Mu-II System, the doppler can easily be disabled with a switch and, therefore, the DRVID data becomes range (group delay) change only. The Mu-II ranging machine was programmed to provide a DRVID integration time of 16 seconds. Typically, 9 to 10 DRVID values were then obtained and averaged at each offset angle. The average DRVID was then added to the acquisition range (obtained only one time at the start of test) to arrive at a delay value at each offset angle for the range measurements to the collimation tower. Then corrections were made for range changes due to signal level changes in the receiver. This was accomplished with the calibration data obtained from Step 4 of the test procedure and shown in Fig. 3.[1] A weighted least square method was used to compute the best fit curve as well as to determine calibration errors. The corrected range values were then converted from ns to meters and a constant was subtracted from the result to provide the final range change values shown on the plots in Figs. 1 and 2. Within the limitations imposed by noise, these plots represent best achievable accuracies on the measurement of range changes as a function of antenna pointing angle when using a ranging system. The error bars are plus and minus one standard error (Refs. 2 and 3) bars due to (1) measurement scatter, and (2) errors on the calibration corrections for signal level changes. The error limits do not include estimate of other possible sources of errors such as ground multipath or near-field effects.

[1]The range correction for a signal level change from −140 dBm to −160 dBm turned out to be only about 0.3 m for the particular receiver used for the test (see Fig. 3).

Possible causes of time delay changes occurring when the antenna pointing goes from boresight (on target) to off boresight are (1) geometric time delay changes with respect to the antenna center of rotation, (2) differences in antenna phase versus frequency responses at different look angles (Cha, Ref. 4) and, (3) near-field effects. The magnitude of contributions from these error sources are presently not known, but will be studied in the near future.

The described collimation tower tests were necessarily performed at this time because the 26-m antenna at DSS 42 will be updated to the 34-m-diameter size at the end of September, 1979. After the update, the DSS 42 antenna will be more in the near-field of the collimation tower at S-band. It is likely that future time delay experiments with a collimation tower technique can be done easily only at DSS 13.

V. Concluding Remarks

The results of this article indicate that if the target is within the 3-dB points of the antenna beam, the range change will be less than 0.25 meters. A portion of this change is attributed to noise and possible measurement errors, but there still appears to be some small residual effects. Although these residuals are negligibly small for ranging, they could be significant for VLBI.

The test results presented in this article should be analyzed further and compared to the results of a theoretical antenna study. The theoretical study should consist of (1) finding out if group delay changes actually occur through different parts of the antenna pattern, and (2) finding out what errors occur due to the fact that the collimation tower is not in the far-field of the 26-m antenna. To investigate the magnitude of errors of antenna mispointing on VLBI time delays, antenna pointing error tests should be performed with VLBI radio sources and use of VLBI measurement techniques.

Acknowledgments

The test results described in this article were obtained by Alan Robinson and Alan Bailey at the DSN station in Tidbinbilla, Australia. They and other personnel at the DSS 42 station are to be commended for their cooperation and for the technical skills and initiative they have displayed in obtaining data for this special ranging test.

Ted Cullen, of the Radio Frequency and Microwave Subsystems Section, derived the equation and wrote the weighted least squares computer program for the range-versus-signal-level calibration curve and associated error calculations.

References

1. Martin, W. L., and Zygielbaum, A., *Mu-II Ranging*, Technical Memorandum 33-768, Jet Propulsion Laboratory, Pasadena, Calif., May 15, 1977.

2. Eisenhart, C., "Expression of the Uncertainties of Final Results", *Science*, Vol. 160, June 14, 1968, pp. 1201-1204.

3. Ku, H. H., "Notes on the Use of Propagation of Error Formulas", *Journal of Research of the National Bureau of Standards – C. Engineering and Instrumentation*, Vol. 70C, No. 4, Oct.–Dec. 1966, pp. 263-273.

4. Cha, A. G., private communication, January, 1979.

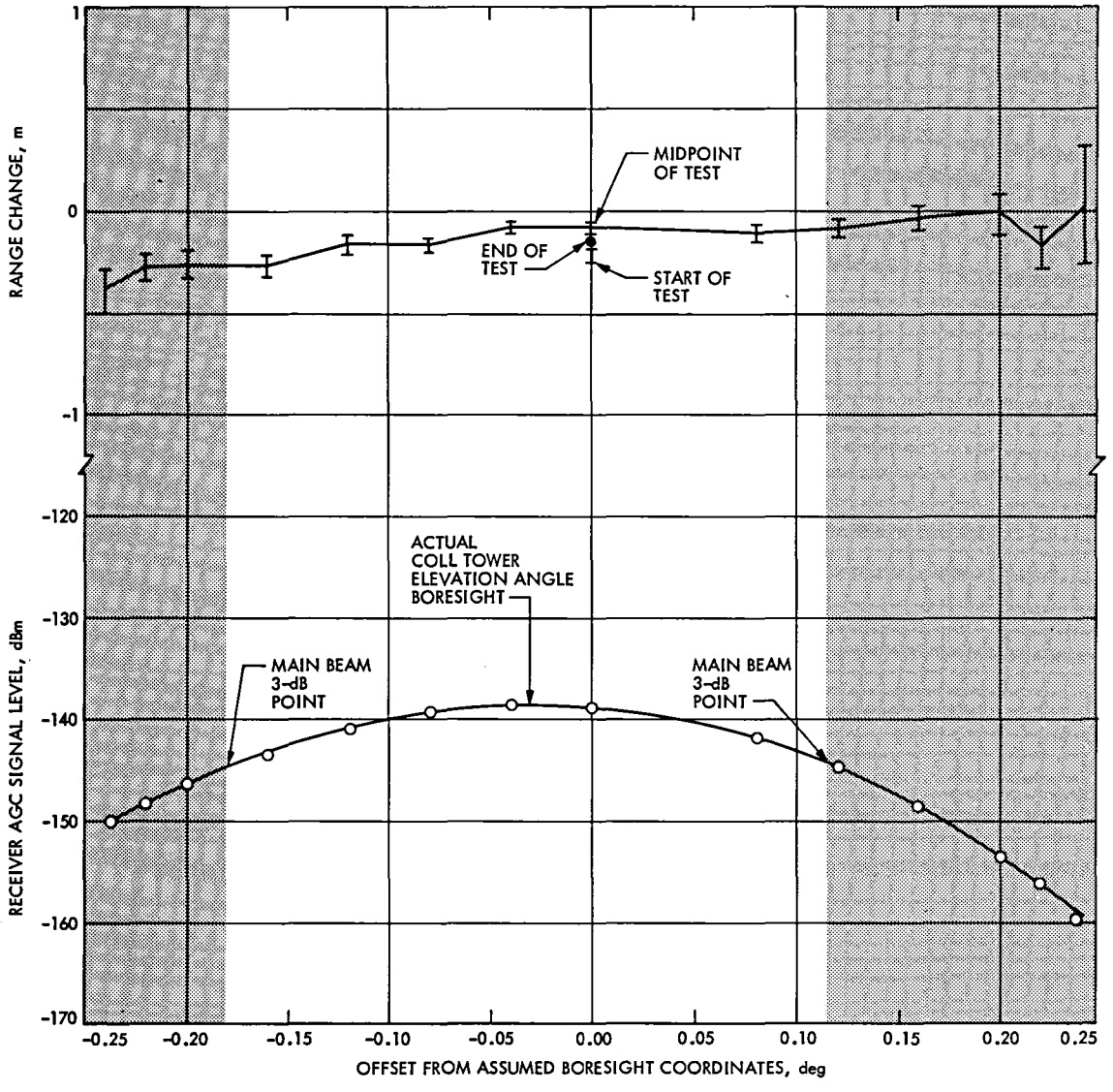


Fig. 1. Range delay change versus elevation angle offsets from boresight. Tests were performed with 26-m antenna and collimation tower at DSS 42 on 1979 Day of Year 50

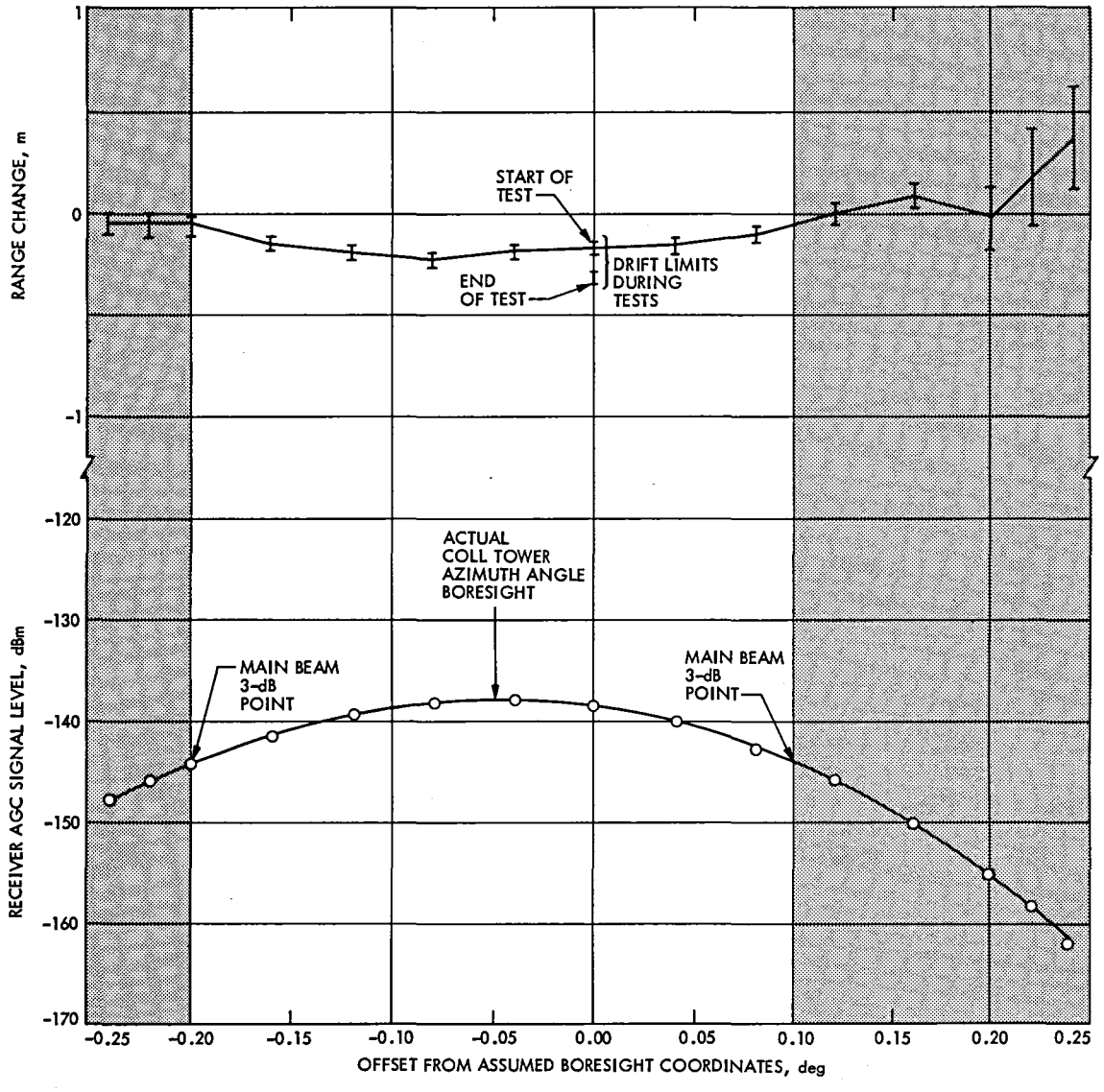


Fig. 2. Range delay change versus azimuth angle offsets from boresight. Tests were performed with 26-m antenna and collimation tower at DSS 42 on 1979 Day of Year 50

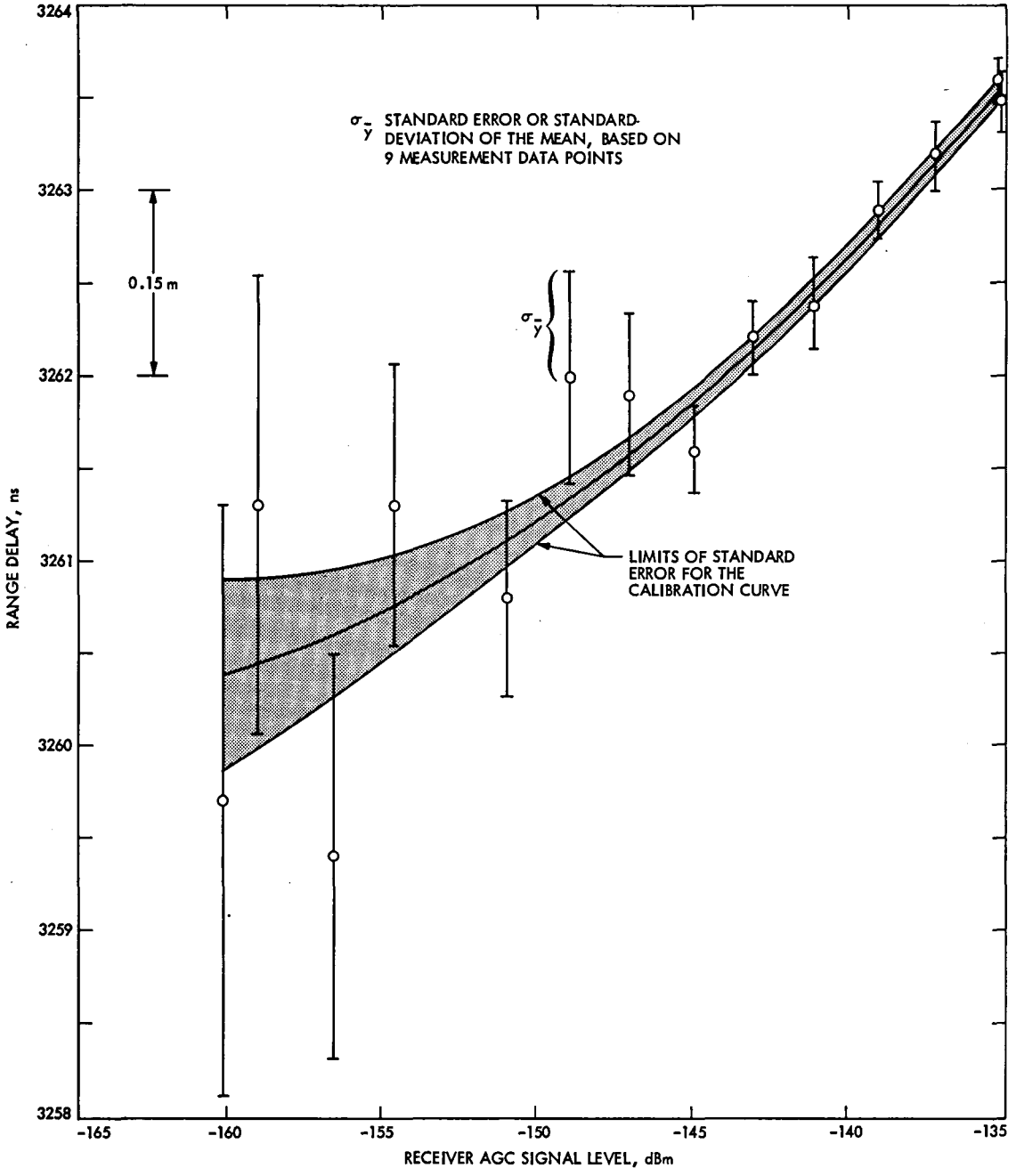


Fig. 3. Weighted least squares calibration curve of station range delay versus signal level characteristics for ZDD mounted on 26-m antenna at DSS 42

S-Band Ultralow-Noise Traveling-Wave Maser

D. Trowbridge and J. Loreman
Radio Frequency and Microwave Subsystems Section

Two S-band traveling-wave maser (TWM) systems with effective input noise temperatures of 2.0 K at 2295 MHz have been supplied to the Deep Space Network. These TWMs are used on the 64-m antennas at Deep Space Stations 43 and 63 to meet the requirements of the Voyager and Pioneer projects. The TWMs use shortened and cooled signal input transmission lines to reduce noise and are equipped with superconducting magnets and solid-state pump sources to provide improved stability performance.

I. Introduction

The Voyager and Pioneer projects have increased requirements for traveling-wave maser (TWM) sensitivity and stability at S-band. Two TWM systems have been built in response to these requirements and installed at Deep Space Stations (DSSs) 43 and 63. Interchangeability with existing Block III TWMs was required. The TWMs have shortened and cooled input transmission lines that reduce the noise temperature as compared to previous operational systems (Ref. 1). Superconducting magnets and solid-state pumps are used to improve TWM gain, phase, and group delay stability. Previously reported research and development (R&D) TWMs (Ref. 2) used the shortened input transmission lines; they were described in detail and will not be covered here. The TWMs provide 45 dB net gain with a 3-dB bandwidth from 2265 to 2305 MHz. Optimum noise performance is achieved at 2295 MHz. An effective input noise temperature of 2.0 K and an overall system temperature of 8.0 K were measured during the evaluation of these systems using a low-noise horn looking at cold sky. The new S-band TWMs are identified Block IV TWMs.

II. Maser Description

The Block IV TWM assembly shown in Fig. 1 is supported by a new frame that is similar to X-band TWMs implemented into the DSN this year. The Block IV TWM package design is smaller in volume and weight, approximately 80 kg, and enables the maser to be shipped at lower cost. Previous TWMs (Ref. 1) utilized a much larger frame and included support equipment on the same frame. The Block IV TWM package enables the refrigerator-maser assembly and pump assembly to be separated from the support equipment, thereby simplifying removal and maintenance procedures. The Block IV TWM assembly is similar in function to a previous R&D model TWM (Ref. 2). The Block IV TWM uses a similar shortened and cooled signal input transmission line that reduces its noise contribution to only 0.4 K; previous Block III TWMs (Ref. 1)

used input signal transmission lines that contributed 2.1 K. The Block IV TWMs are equipped with superconducting magnets and solid-state pump sources to provide improved stability performance. The use of solid-state pump sources will also increase reliability over previous klystron tube pump sources used on Block III type TWMs.

III. Structure Detail

The Block IV TWMs use modified Block III TWM comb structures returned from the DSN for conversion. The converted comb structures are similar to Block III units as reported in Ref. 2 with the exception of modifications described here. A large slowing factor was used on the original Block III TWM comb structures and, because of high resulting gain, magnetic field distortion was necessary to reduce gain to the specified value. During this conversion, the electrical length was reduced from 40 meters to around 35 meters. The reduced electrical length was achieved by grinding a bevel on the ruby top edge on the surface contacting the comb. The structure lid-to-comb-finger top clearance was decreased to further reduce the slowing and shift the maser bandpass lower in frequency to compensate for the upward shift caused by ruby shaping. Reduced slowing, used mainly to obtain the proper gain without magnetic field distortion, results in reduced structure loss and reduces the TWM noise temperature. The desired bandwidth is obtained by use of two field spreading coils attached to the TWM structure. The coils create two magnetic field strength levels within the maser structure and result in a stagger-tuned bandpass. The coils are movable along the length of the maser comb structure and are used to equalize the gain of the lower and higher frequency sections of the maser. Separation of the coils reduces the overall net gain and increases the maser bandwidth by producing a third intermediate, magnetic field region. A soft iron shim is used under one of the coils to reduce the amount of field spreading current required.

IV. Solid State Pump

The solid-state pump assembly uses a Gunn-effect oscillator (for the pump frequency source), a modulator, and protective circuitry similar to a previous solid-state pump source (Ref. 3). A commercially available overvoltage protect circuit and variable attenuator are used with the Block IV pump source assembly shown in Fig. 2. The overvoltage protector prevents damage to the Gunn-effect oscillator caused by excessive power supply output voltage due to improper adjustment, improper connection, or failure of the power supply. The variable attenuator is used to reduce the pump power to a level that saturates the TWM without subjecting the cryogenic system to excessive heat loads. The Gunn-effect oscillator provides up to 200 mw of power electronically tunable from 12.6 to 12.9 GHz. The modulator can be adjusted to sweep the pump frequency across any portion of this range at a 100-KHz rate.

V. Superconducting Magnet

The superconducting magnet (SCM) provides a uniform field near 0.25 tesla as required for S-band maser operation. The SCM is similar to a previously reported SCM (Ref. 4). The magnet operates in a persistent mode in the vacuum environment of the closed cycle refrigerator (CCR) at the 4.5 K temperature. The circuit consists of a pair of superconducting magnet coils, a superconducting switch (S-C switch) and a resistive shunt circuit. The SCM coils and the S-C switch are constructed of superconducting wire of copper-clad niobium-titanium (NbTi), coated with Formvar insulation. The SCM coils and S-C switch are constructed from one continuous piece of superconducting wire with a zero-resistance superconducting welded junction at a heat sink attached to the 4.5 K heat station of the CCR. The S-C switch is a portion of the superconducting wire (with the Formvar and copper removed) inside a small radiation shield, where it is exposed to radiation from a light bulb to heat the wire. The resistive shunt circuit is constructed of brass rod and a specific length of a stainless steel screw. The shunt circuit is connected at one end to the superconducting wire welded junction and the other end (stainless steel screw) to a ground point on the SCM base flange. The shunt circuit is in parallel with the SCM coils and S-C switch circuit. The entire circuit operates at a temperature of 4.5 K, and is mounted on the CCR at the same heat station as the maser structure.

VI. Support Equipment and Instrumentation

The following new assemblies were replaced or incorporated in the Block IV maser instrumentation: (1) replacement of the calibration assembly, (2) new maser monitor assembly, (3) new superconducting magnet control assembly, (4) replaced klystron power supply with the pump control assembly, (5) new magnet charging supply, and (6) new junction boxes for maser and monitor equipment.

The TWM assembly is mounted on a common base plate with the support equipment rack and is shown in Figs. 3 and 4. The support equipment rack provides calibration signal injection points for gain, noise temperature, and signal path length calibration of the TWM. A switch is provided to connect a calibration signal to a pre- or post-TWM injection point for gain calibration of the TWM. A post-TWM directional coupler is provided for output signal connection to the monitor receiver. The monitor receiver remains unchanged and

provides the same visual bandpass indication as used with previous Block III masers.

New operating controls and indicators for the Block IV TWM are provided by the maser monitor assembly, which is located on the support rack assembly (Fig. 4); and the superconducting magnet control assembly (Fig. 5) and the magnet charging supply, which are located in the control room cabinets. These assemblies are used in conjunction with the monitor receiver instrumentation to charge the superconducting magnet to the correct field strength for maser operation. The maser monitor assembly contains a field strength meter and a bridge amplifier that amplifies the current unbalance that exists in a magnetic field sensitive resistance bridge located on the maser amplifier body within the superconducting magnet. The bridge amplifier output is adjusted to provide zero deflection on the superconducting magnet control field strength meter with no magnetic field. It is adjusted for one-half scale deflection at the operating magnetic field strength of 0.25 tesla. The field strength meter provides the operator an indication of magnet charging rate. The magnet unlock control applies power to the S-C switch heater (light bulb) in the SCM assembly. This control is interlocked with the magnet charging power supply to prevent accidental unlocking of the SCM and its discharge to zero field.

The maser pump control, shown in Fig. 6, provides the tune, modulate, and Gunn bias voltages for the solid-state pump source used with the Block IV TWM. The pump control is all solid-state and is expected to greatly reduce power supply failures associated with the previously used klystron power supplies.

VII. Performance

The following performance requirements were established for the Block IV TWM:

(1) Gain and bandwith: 45 ± 1 dB at maximum between 2270 and 2300 MHz with -1 dB bandwidth of 30 MHz minimum and -3 dB bandwidth of 40 MHz minimum.

(2) Gain stability: stationary short term ±0.03 dB for 10 seconds any position, and long term ±0.5 dB for 12 hours any position.

(3) Effective noise temperature: 2.5 K ± 1.0 K, 2285 to 2300 MHz and 2.5 K + 2.0, -1.0 K, 2270 to 2285 MHz.

The gain and bandwidth were measured prior to shipment and the above specifications were met. The noise temperature was measured using a horn and ambient temperature load (microwave absorber). The TWMs effective temperatures were 2.0 K at 2300 MHz, 2.5 K at 2285 MHz, and 4.0 K at 2270 MHz. The gain stability and system noise temperature (includes feed horn, maser, and receiver noise temperature contributions) as reported by DSS 43 and 63 are listed in Table 1.

The Block IV TWMs have met or exceeded all the performance goals established. A third unit is to be installed at DSS 14 in the near future and one additional unit at each 64-m station will bring the total number of Block IV TWMs in the DSN to 6.

References

1. Trowbridge, D., "Block III Implementation Program" in *The Deep Space Network Progress Report*. Technical Report 32-1526, Vol. XVIII, pp. 130-135, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1973.

2. Clauss, R., and E. Wiebe, "Low Noise Receivers: Microwave Maser Development" in *The Deep Space Network Progress Report*. Technical Report 32-1526, Vol. XIX, pp. 93-96, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1974.

3. Quinn, R., "Low Noise Receivers: Microwave Development" in *The Deep Space Network Progress Report*. Technical Report 32-1526, Vol. XIV, pp. 46-49, Jet Propulsion Laboratory, Pasadena, Calif., April 15, 1973.

4. Berwin, R., E. Wiebe, and P. Dachel, "Superconducting Magnet for a Ku-Band Maser" in *The Deep Space Network Progress Report*. Technical Report 32-1526, Vol. V, pp. 109-114, Jet Propulsion Laboratory, Pasadena, Calif., Oct. 15, 1971.

Table 1. TWM performance

Parameter	DSS 43		DSS 63	
Gain stability				
short term	0.04 dB p-p		<0.05 dB p-p	
long term	0.33 dB p-p		0.05 dB p-p	
System noise temperature at:	RCVR 1	RCVR 2	Monitor RCVR	RCVR 3
2265 MHz	19.8	19.8	–	–
2270 MHz	–		19.91	–
2275 MHz	17.3	17.4	–	18.49
2285 MHz	16.3	16.3	18.16	16.78
2295 MHz	15.8	15.8	–	16.67
2300 MHz	–		18.71	–
2305 MHz	16.3	16.3	–	–

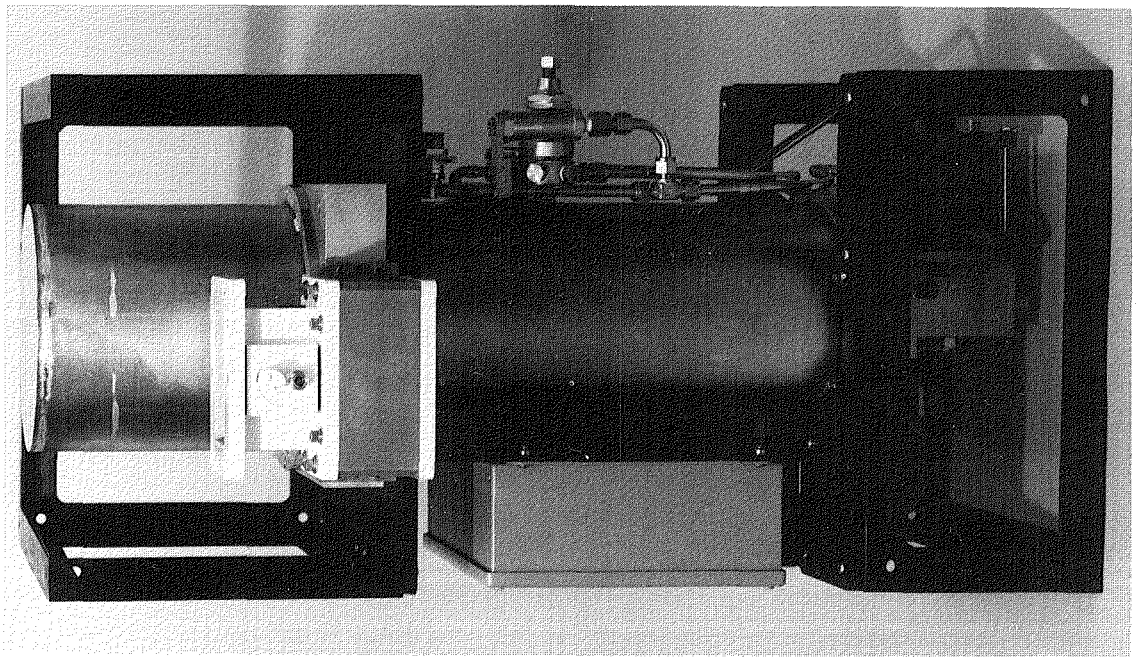


Fig. 1. Block IV traveling-wave maser (TWM) assembly

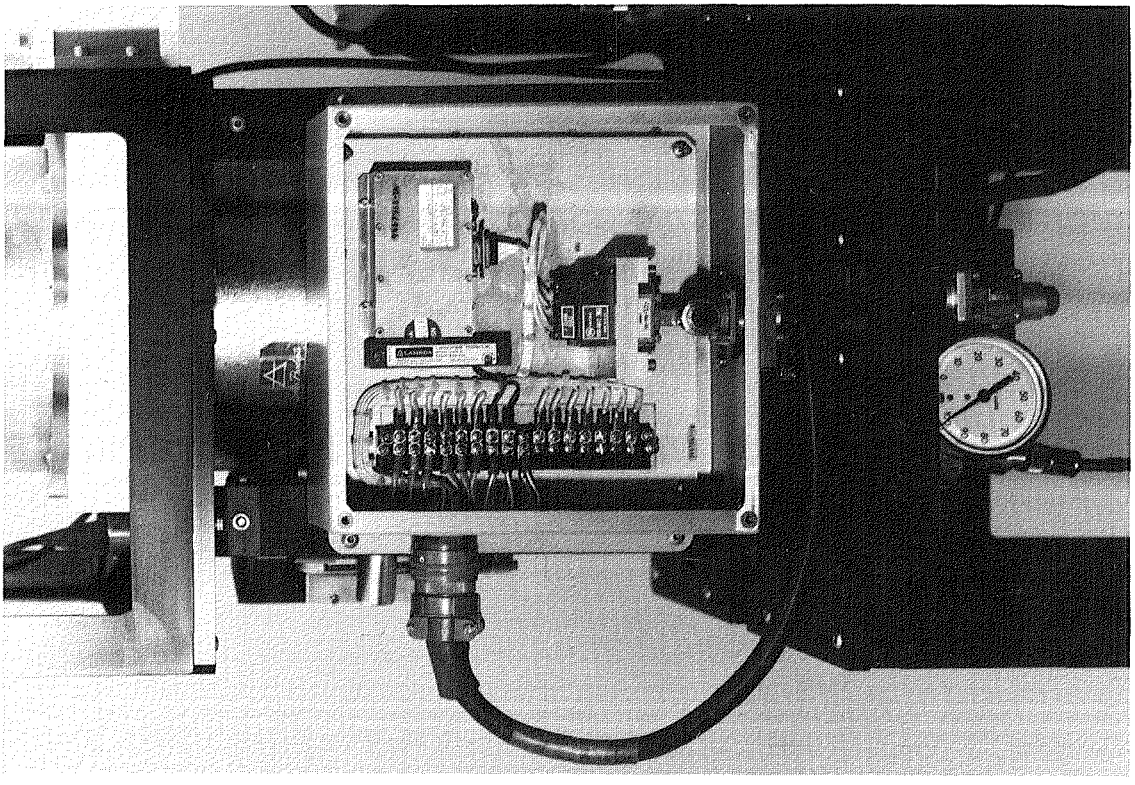


Fig. 2. Pump source assembly mounted on Block IV TWM

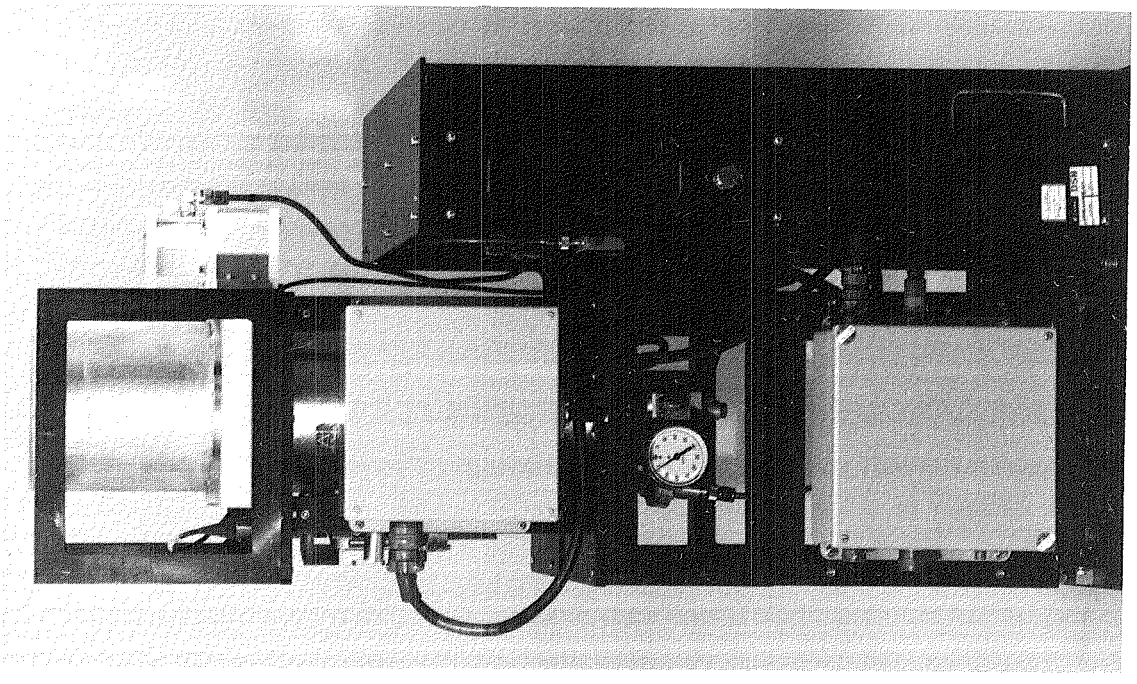


Fig. 3. TWM/CCR and support rack assembly, front view

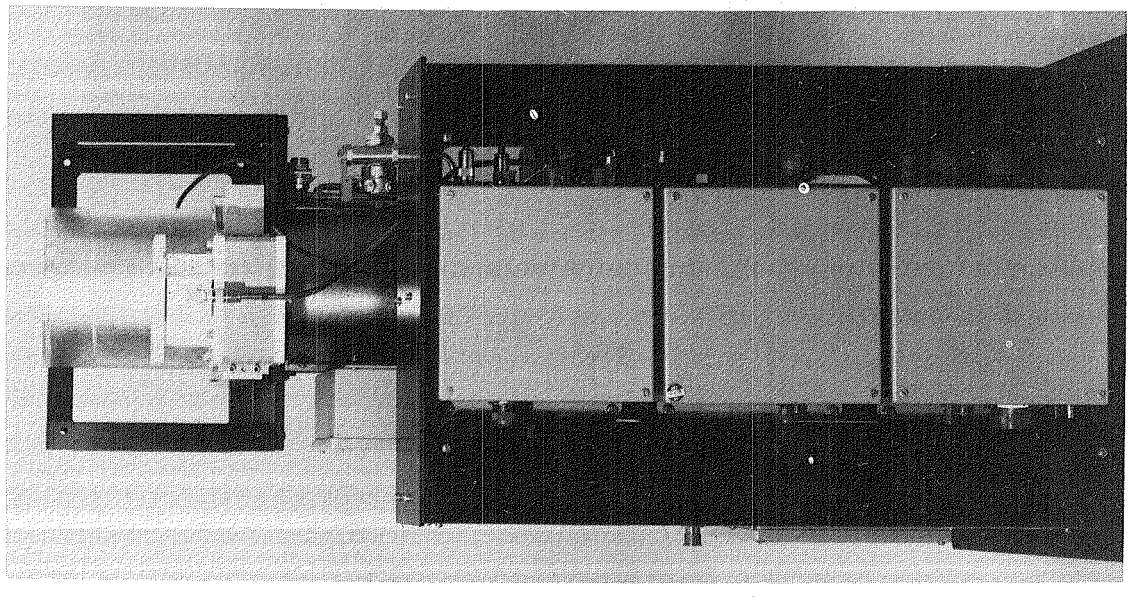


Fig. 4. TWM/CCR and support rack assembly, side view

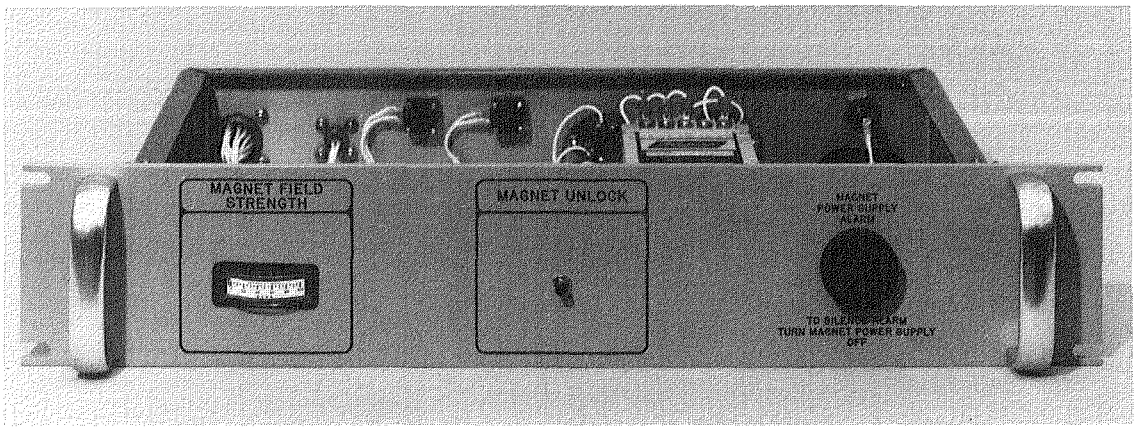


Fig. 5. Superconducting magnet control

Fig. 6. Maser pump control

An Improvement in Fourier Transform Accuracy

R. C. Bunce
Deep Space Network Support Section

Most Fourier transform algorithms, when seeking a single frequency at the sample-time midpoint, simply choose the largest component of the set. The error is at least ±0.5 component, and often larger. This article describes an addition to such algorithms that reduces the variance error by three-to-one (nominal). The addition achieves this by averaging all *components within a predetermined "window," selected as a function of frequency rate, and it is quite easy to mechanize within the existing algorithm.*

I. Introduction

The Fourier transform, particularly the fast algorithm or "FFT", is used in numerous data-reduction applications in the Deep Space Network (DSN). In many of these applications, the transform acts as a kind of "digital filter," yielding a single largest component representing an input sinusoid. The usual input signal is initially contaminated by additive noise, which the transform "spreads out" into a low-order spectrum, attenuated "around" the representative component. The spectrum in use is the "discrete" form of the "finite" Fourier transform, finite implying a bounded time sample, and "discrete," a finite number of sample points.

When the sinusoid is not precisely synchronized with the sample-point period (usual case), or when the frequency is time-variant, or both (normal tracking conditions), the signal transform itself is a spectrum. Under normal conditions, this spectrum contains from four to ten or more significant components. If accuracy is important, the frequency chosen must be associated with a given "time-tag" within the time sample.

Normally, the first-order frequency excursion, or "ramp" assumption is acceptable for DSN applications using sample periods of (up to) a few seconds.

The present process is to simply select the "largest" component as the estimate, associating this with the sample-time center.

This article describes a supplementary averaging procedure that gives promise of a significant improvement in accuracy over the "largest component" method. If implemented, it would require an addition to the existing algorithms, but the addition would be neither particularly lengthly nor difficult to mechanize. It would require one additional a priori input: an estimate of the ramp-rate, or "$\dot{F}$", as present on the sample. This could be obtained from RF doppler, as modified by any additional corrective predicts.

The "$\dot{F}$" values encountered in tracking normally vary from about 1.0 Hz/sec (or less) up to 20 Hz/sec (or more). A nominal value of 5 Hz/sec was chosen for the model to follow.

II. The Finite Fourier Transform of a "Frequency Ramp"

The continuous form of the finite Fourier transform, or power spectral density, is (if noiseless):

$$P(f) = \left| \int_0^{\tau} e^{-j2\pi ft} F(t)\, dt \right|^2 \qquad (1)$$

where, for the case in question:

$$F(t) = \sin (2\pi f_0 t + \pi \dot{f} t^2 + \phi_0)$$

This is the envelope of the components obtained by the discrete expression. With the discrete algorithms, components are spaced "$1/\tau$" apart, across a range "$N/2\tau$"; N the discrete sample point-count. "f_0" is the frequency at the start of the sample (sample length "τ"), and $\dot{f}$ is the "ramp-rate", Hz/sec, during τ. $\dot{f}$ is considered constant, thus implying neglect of higher-order terms, if any. A model program to simulate Eq. (1) in discrete form was coded for this study. At first glance, Eq. (1) appears reasonably straightforward; it would seem logical to simply evaluate the envelope (given $\dot{f}$), then "slide it across" sample data to obtain an estimate of f_0 at $t = 0$, or some similar "best fit" result. This was, in fact, the initial approach.

However, the envelope of Eq. (1) proved very cumbersome to calculate, and this approach was discarded in favor of the simpler routine to follow.

III. Frequency-Time Estimate by Component Averaging

Refer to Fig. 1; this is a typical discrete transform (Fig. 1) that simulates Eq. (1) by the model program when:

$$\tau = 1 \text{ sec}$$

$$f_0 = 22 \text{ Hz}$$

$$\dot{f} = 5 \text{ Hz/sec}$$

The transform base is 50 Hz long, with components 1.0 Hz apart. $N = 100$ points. These are the "standard conditions" of this study. The "bulk" of the study is an investigation of the transform of Fig. 1 under various signal-to-noise ratios (SNR). The purpose was to evaluate the accuracy with which the center frequency of the spectrum could be related to the center of the time sample, $\tau/2$, or 0.5 sec.The true frequency at 0.5 sec. was 24.5 Hz in all cases.

The "inherent" error in the above case, strong signal, by the "largest component" method, is 0.5 Hz; the true center frequency is midway between two components, one of which is chosen. In general, this strong signal error will have a standard deviation of 0.35 Hz, the well-known "$1/\sqrt{12}$" constant.

However, as the noise increases (artificially "injected" in the model program), the "largest component" became more and more likely to appear at locations other than "adjacent to" the center. Refer to Figs. 2, 3 and 4 for typical results. In one run, at -5 dB SNR, this component actually appeared at *2.0 Hz*, a value that would have led to data discard.

The random-displacement naturally led to increased "sigma", or standard deviation. Model results for this, based on 12 samples at each of three SNR, are shown on Fig. 5 (outer set), with estimated extrapolations. These describe the error, using present techniques, for the stated (Fig. 1) conditions, plus-noise, as obtained by the model program.

Even though the envelope of Eq. (1) is unreasonably complex to express in its total form, it *is* easy to show that it is not only symmetric above the center $f(\tau/2)$, but that it attenuates ("dwindles") rapidly outside of the waveform frequency excursion due to $\dot{f}$. The details are omitted.

These characteristics of the envelope led directly to the concept of *component averaging*, which relies on the signal power of *all* signal components present, rather than upon a "single representative." The process is two-fold:

(1) Using an estimate of the envelope width, or "spread" obtained from an a priori measure of $\dot{f}$, a sequential set of component power summations is taken, within a "window" of this width, across the entire spectrum. The region of *greatest total power* is assumed to contain the center frequency.

(2) The second step involves averaging the power-frequency components in the "window." Even though envelope symmetry does not imply component symmetry, when several components are in the "window," the heights (power) of the corresponding "pairs" (closest frequency "fit" on either side) is very nearly a linear function of their absolute offset from center; they "average out."

Said in another way, the envelope derivative on opposite sides of center is approximately of the same slope and linear over short corresponding absolute intervals. Such a derivative is simply the argument of Eq. (1). Investigation showed little discernible error over intervals of "$1/\tau$" in n.

This and several "intuitive" observations from the model all suggested that a very good estimate of the desired center frequency could be obtained by simply averaging the (normalized) power-frequency products of the components within the "maximum power window", the second step of the procedure as stated.

These steps can be described notationally for inclusion in existing algorithms. Let:

N = total number of components in discrete spectrum (about 100 to 1000)

τ = sample length, sec

$\dot{f}$ = ramp-rate (a priori), Hz/sec

Then:

$$SR = \text{search range}$$

$$= INT\,[2\dot{f}\tau]$$

The digit "2" is not critical. It is chosen here from machine observations of the extent of significant "side lobes" in the transform assisted with various "$\dot{f}$".

FW = center frequency "window" chosen for averaging, low-side index

$$= K \text{ of } MAX \text{ of } \left\{ \sum_{m=K}^{K+SR} C_m \right\} \tag{2}$$

where

$$K = 1, 2, \cdots N - SR$$

C_m = (power) size of m^{th} Fourier component

(Step 1)

And, finally:

$$\hat{F}\left[\frac{\tau}{2}\right] = \frac{\sum_{n=FW}^{FW+SR} F_n C_n}{\sum_{M=FW}^{FW+SR} C_n} \tag{3}$$

where

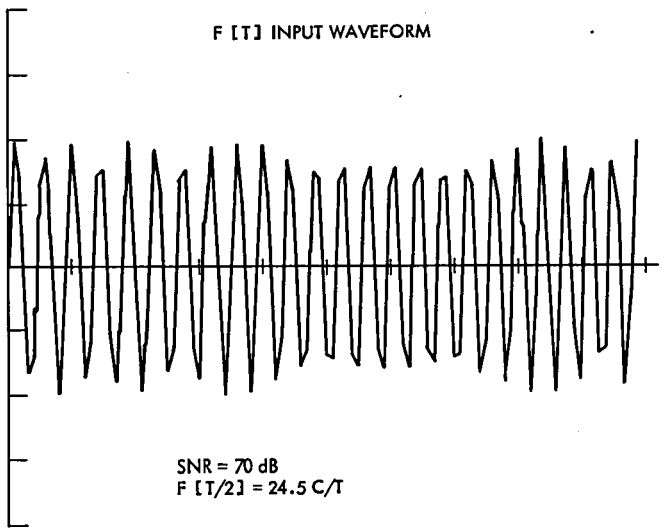
F_n = frequency of component C_n

(Step 2)

The model statistical results, using the above, with $SR = 8$, are shown in collected form on Fig. 5. The accuracy advantage under stated conditions is evident. Improvement up to nearly 3 to 1 was accomplished in the mid-SNR region, and the strong-signal residual was barely evident. Signal-to-noise is that within a 100-Hz bandwidth.

IV. Conclusions

The component summation addition to existing Fourier transform algorithms in DSN applications appears to have considerable potential as a minor tool for accuracy improvement. The data on Fig. 5 demonstrates its behavior with respect to the present process, using the model and levels described herein.



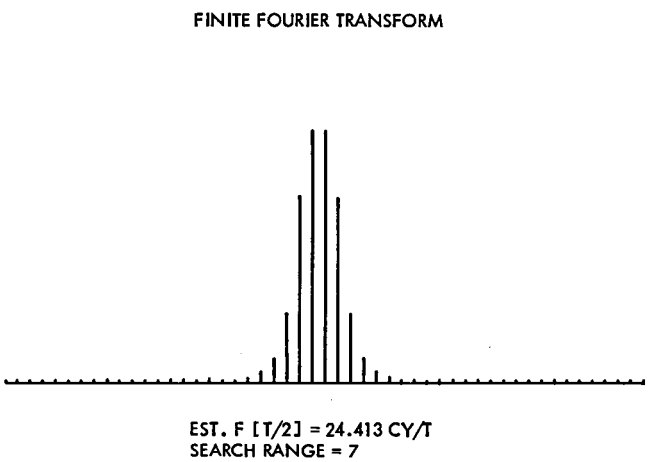
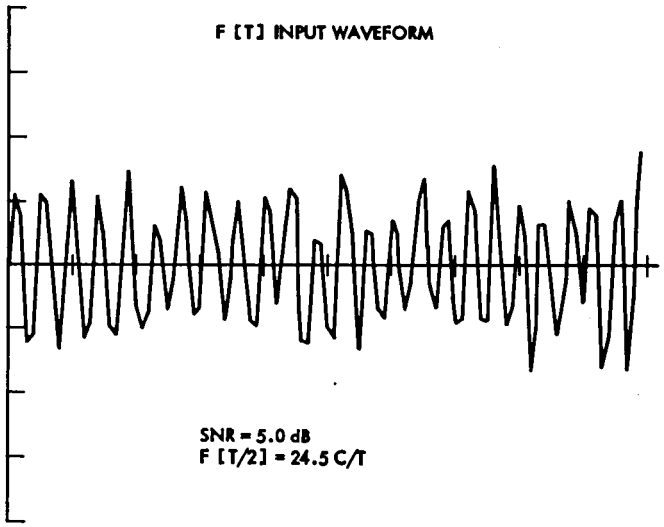


Fig. 1. Transform of ramped sinusoid, noise-free



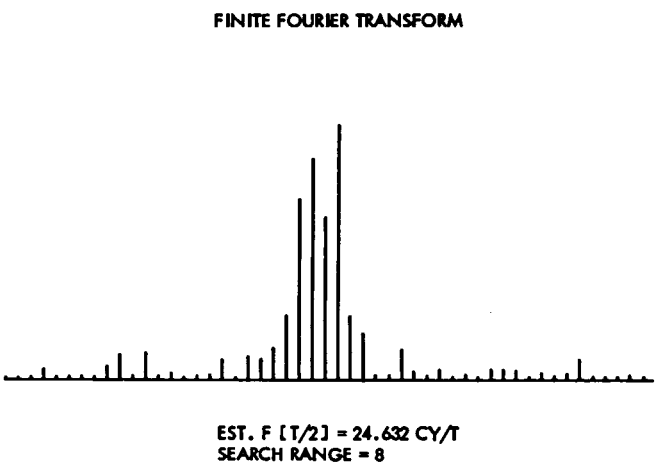


Fig. 2. Transform of ramped sinusoid, SNR = +5 dB (typical) in 100 Hz

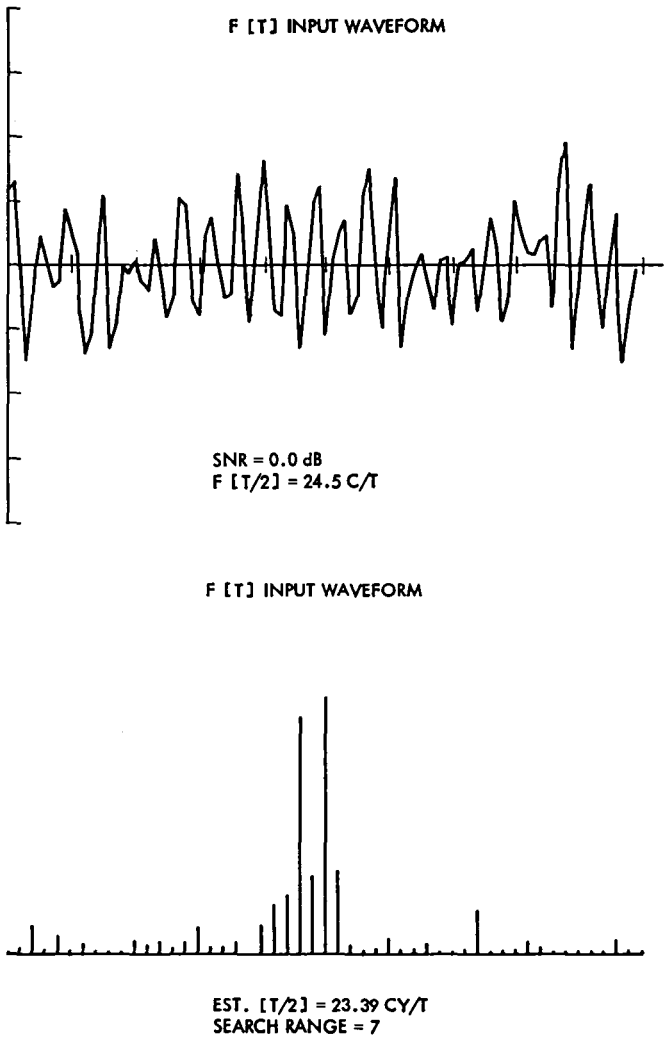


Fig. 3. Transform of ramped sinusoid, SNR = 0 dB (typical) in 100 Hz

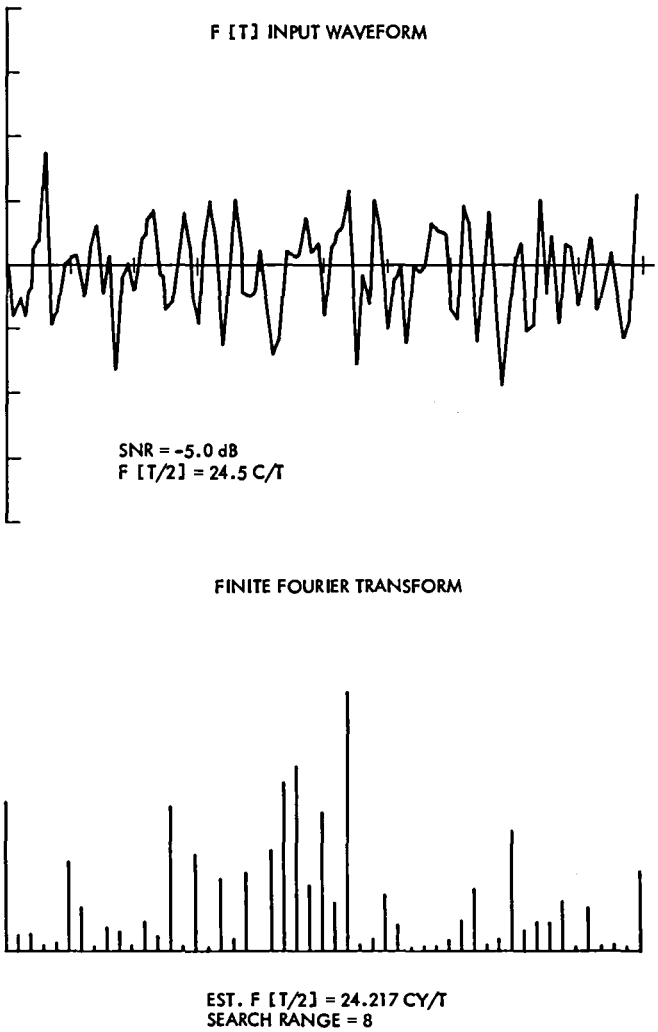


Fig. 4. Transform of ramped sinusoid, SNR = −5 dB (typical) in 100 Hz

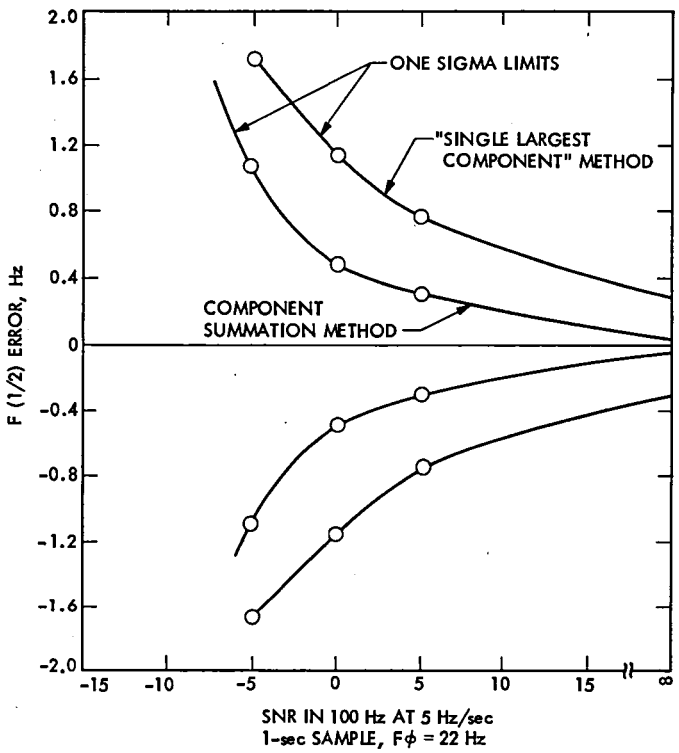


Fig. 5. Statistical comparison, Fourier transform accuracy, two methods

DSS Receiver Thermal Noise Model Comparisons

R. C. Bunce
Deep Space Network Support Section

DSN receiver thermal noise is a component of the Doppler jitter noise measured during DSN system testing. This thermal noise has been modelled theoretically through an evolution of approximations over a period of 20 years. The model is required to predict the expected test result. This article compares the two latest and most accurate models in order to determine if changes in the test algorithm are required for desired accuracy. Also introduced is a new and highly simplified model that exhibits differentials of the same order as the more complicated algorithms. The study concluded that the three models are indistinguishable within the nominal operating range of the receivers.

I. Introduction

DSN receiver thermal (phase) noise is one contribution to overall system noise, as measured during DSN system testing. The test is mechanized to extract the noise as "Doppler jitter," sampled to include the majority of the frequency spread, and reduced finally by the Allen variance technique to reject long-term drift effects.

The receiver thermal contribution is predominant at lower signal levels, but tends to be "swamped out" at strong signal levels by VCO noise, quantization noise, and similar sources.

However, for "pass-fail" criteria, a reasonable theoretical model of expected thermal noise vs. signal level is required across the entire range. Variations of this model have developed historically over more than 20 years. In order to consider possible present "update" of the existing test model, this study was conducted. It compares the two most "recent" models, and introduces a simplified exponential model with comparable accuracy. These thermal models are, of course, variance-added to models of other contributions to obtain final estimates for total system Doppler jitter evaluation.

II. Receiver Thermal Noise Major Parameters

The prime parameter in DSN receiver design is a quantity labelled $2BL_o$. This number, displayed on equipment push-buttons for selection, represents the square-and-double-sided phase noise bandwidth if the loop operated entirely linearly, without detector signal-to-noise ratio (SNR) gain, and if the output SNR (design point) variance were unity. $2BL_o$ is, therefore, entirely hypothetical, for these conditions are inexact. However, loop filter time constants and gain calculations are determined for specified values of $2BL_o$, and all equipment is structured by this design parameter, with conventional linear optimization.

Now, given $2BL_o$ and a system operating temperature, an exact (minimum) reference *signal level* can be expressed. It is

the level when signal and noise power are equal and if the noise bandwidth *is* exactly $2BL_o$, as referred to the receiver input. Said another way, it is the signal that would exhibit an (undetected) phase variance of unity if passed through a filter (double-sided) with a width of $2BL_o$ Hz (open loop). This is the reference level used in DSN testing. It has the value:

$$W_o = KT(2BL_o)$$

$$W_o \text{ (in dBm)} = 10\log(T) + 10\log(2BL_o) - 198.6$$

W_o = reference signal level, W or dbm

T = system temperature, K

K = Blotzman's constant, W/K/Hz

$2BL_o$ = reference (loop) noise bandwidth

In some literature, W_o is 3 dB less, $W_o/2$, at a point where the hypothetical "linear-loop" variance is unity. This discussion avoids that definition for conformance with the DSN testing approach, and as a common base in model comparison.

Using the above, W_o in this discussion takes the role of "minimum signal level" to be investigated. This is reinforced by the fact that *actual* receiver thermal noise, all factors included, results in a variance very close to unity at W_o. This is, therefore, in the vicinity of the original definition of threshold, where the receiver is no longer reliably phase-stable; a true minimum.

Results of this study are presented as jitter (standard deviation, deg, the square-root of variance) vs. margin, dB for various models and $2BL_o$. Given W_o, margin (M) is easily defined by the ratio of actual input signal level to W_o:

$$M = \frac{W}{W_o}$$

$$M \text{ (dB)} = 10\log\frac{W}{W_o}$$

$o \leqslant M \text{ (dB)} \leqslant 50$ (model spread)

W = actual signal level, W or dBm

Also:

$$M \text{ (dB)} = W \text{(dBm)} - W_o \text{ (dBm)} \text{ (for test settings).}$$

III. Predetection Signal-to-Noise Ratio

The phase-lock-loop predetection limiter effects are functions of the predetection signal-to-noise ratio, or PH in most literature. These effects are normally isolated as two-fold: the signal suppression factor, α (signal gain), and the limiter performance factor, Γ (SNR transfer). Both effects are strictly functions of PH; the Γ factor (EXACT) model is quite complicated. Model expressions for these effects are given in the Appendix. However, these expressions are invalid without a sound definition of PH itself.

Since PH occurs at a predetection point, and can be considered as due entirely to input levels, it is easily defined from the predetection filter noise bandwidth, and previous notations:

$$PH = \frac{W}{KT(BC)} = \frac{W_o}{KT(BC)} \times \frac{W}{W_o}$$

$$= \frac{KT(2BL_o)}{KT(BC)} \times M = M \times \frac{2BL_o}{BC}$$

BC = Predetection (two-sided) noise bandwidth, Hz.

IV. The System Signal-to-Noise Ratio Conversion

This concept, introduced by Yuen, describes a very clear-cut model approach. It states that expected output thermal variance (as on Doppler measures), can be modelled by two discrete and simple steps, independent of the quantity of nonlinear (sinusoidal) detectors in the signal path:

(1) Determine the total linear theory variance to output; just consider all detectors linear. Double SNR (divide variance by two) to account for detection enhancement. The figure obtained is:

$$\sigma_\varrho^2 = \sigma_\varrho^2(M) = 1/p$$

p = system SNR ratio

σ_ϱ^2 = system phase variance (linear computation)

(2) Convert the system linear variance from the Gaussian to nonlinear modulo 2-Π (sinusoidal detector) value in a sinusoidal conversion spectral density expression. Yuen proved this valid.

These steps are used for the organization of both the EXACT and approximate models.

V. Expected Doppler Receiver Thermal Noise, $E(\sigma)$

The general algorithm for thermal noise, using conventional loop design by Jaffee/Rechtin can be organized as follows:

Inputs:

$2BL_o$ = Design point loop noise bandwidth, Hz

BC = Predetection noise bandwidth, Hz

T = System temperature, K

W(dBm) = Input signal level, dBm.

Prime quantities (all models):

$$W_o(\text{dBm}) = 10\log(T) + 10\log(2BL_o) - 198.6$$

$$M(\text{dB}) = W(\text{dBm}) - W_o(\text{dBm})$$

$$M = 10^{\frac{M(\text{dB})}{10}} \qquad 1 < M < 10^5$$

$$PH = \frac{BC}{2BL_o} \times M.$$

The algorithm proper follows two major steps:

STEP NO. 1:

$$\sigma_\varrho^2 = \frac{1}{p} = \frac{1}{2M} \times \left[\frac{1}{3} + \frac{2}{3}\frac{\alpha}{\alpha_o}\right] \times \Gamma = \text{system variance (linear computation)}$$

$\alpha = F(PH)\,|_M$

$\alpha_o = F(PH\,@\,M = 1)$

$\Gamma = G(PH)\,|_M$

} Expressions vary among models. See the Appendix.

STEP NO. 2:

$$E(\sigma) = H_1(\sigma_\varrho^2) \text{ (Simon/Lindsey/Yuen) (EXACT)}$$

$$H_2(\sigma_\varrho^2, \alpha/\alpha_o)\,|_M \text{ (Tausworthe) (approximate)}$$

The algorithm, thus, involves the calculation of W_o, M, PH, α, Γ, σ_ϱ^2, and $E(\sigma)$. The first three are elementary, requiring compatible interpretation only. The last three are model-peculiar and somewhat complicated. The factor Γ is not presently available expressed in exact form.

Three models of the above are given in detail in the Appendix:

(1) The original Jaffee-Rechtin linear model

(2) The Tausworthe linear spectral analysis model (in use for DSN testing)

(3) The Simon/Lindsey/Yuen EXACT Model

The first model, linear, is plotted throughout the literature and this plot will not be repeated. The second model, on a straight plot, was indistinguishable from the third, or EXACT model. This latter model is plotted, in the range of interest, in Fig. 1.

Using this EXACT model as a standard, the differential between it and the second, or Tausworthe (presently in use within the testing algorithm), is shown in Fig 2, as a function of margin. Except for a 2 degrees difference at low margin (M = 1), the two models agree within 0.5 deg, with a small bias. The RMS of the differential is 0.4 deg, two bandwidth ratios, inclusive. If the 2 deg differential near M = 1 is disregarded, the overall RMS differential drops to 0.18 deg.

There appears to be little point to changing the DSN algorithm to the EXACT model, for differentials on the order of those above are essentially "masked" by other noise contributions in the system.

However, there is an elementary approximation to the EXACT model, developed by the author; the description follows. It might be worthwhile to use this in testing for algorithm simplicity.

VI. The Exponential Approximation Model

A quick glance at Fig. 1 suggests that the full model, with decibel abscissa and standard deviation ordinate, appears to approach the form of a simple exponential decay.

When this form was tried, with coefficients to fit the end points (margin 0 dB and 50 dB), the resulting exponential fit was within 3 deg throughout. The error plot had the form of

the Gaussian derivative. This led to the following two-term exponential model:

$$E(\sigma) = 55.12\, e^{-a_0 \times M} + M \times a_1 e^{-a_2 \left(\frac{M}{10}\right)^2}$$

Main Approximation | Trim Term

M = Margin in dB

A variation-of-parameters closure program was applied to (1), vs. the EXACT model, and minimum RMS differential for the receiver bandwidth ratios was obtained (see Table 1). The deviations of these approximations from the EXACT model is shown on Fig. 3. The overall RMS of the deviations is 0.22 deg. This level, as with the Tausworthe model, would normally be masked by other contributions.

V. Conclusions

Three nonlinear models for receiver thermal noise were investigated. Within the range of interest (0 dB to 50 dB margin), differences between the models were indistinguishable.

One of the models, the Tausworthe or linear spectral density approximation, is presently in use for DSN system testing. There appears to be little practical advantage to altering the algorithm at this time.

Another model, the Simon/Lindsey, or EXACT algorithm, is complicated to program, but generally considered the most accurate noise prediction available. It might be reasonable to incorporate this model at some future date during general software update.

Finally, the third model, or Exponential Approximation, by the author, is by far the simplest expression for thermal noise that is presently on hand. It is seen as useful in abbreviated programming and, possibly, as a future general-purpose design and test-predictive expression, replacing even the EXACT model when simplicity is desired.

Table 1. a_0, a_1, a_2, coefficients

Block	$BC/2\,BL_o$	(Typ)	a_o	a_1	a_2
III	166.7	2000/12	0.11288	0.58027	0.21658
IV	66.67	200/3	0.11198	0.55579	0.26151
IV	200	2000/10	0.11364	0.59611	0.21125

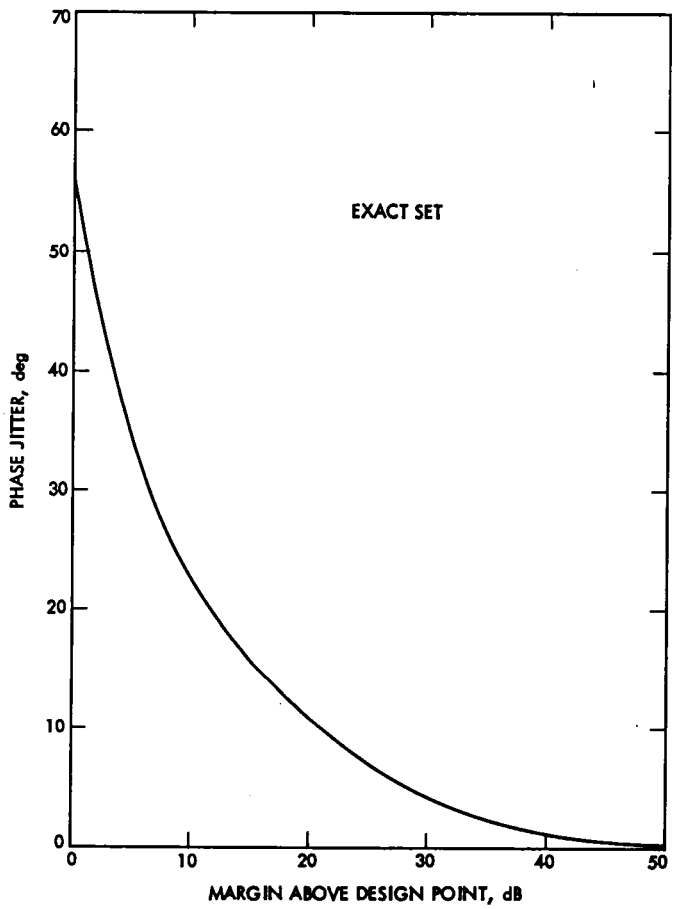


Fig. 1. Receiver thermal noise exact model

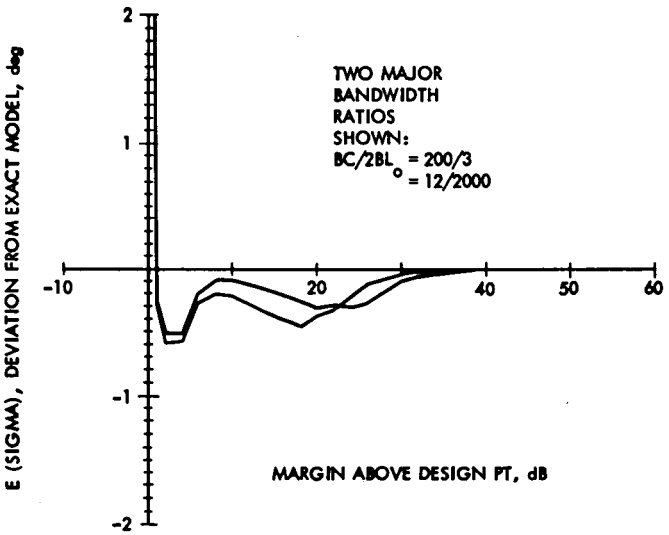


Fig. 2. Deviation between exact and linear-spectral-density approximation models

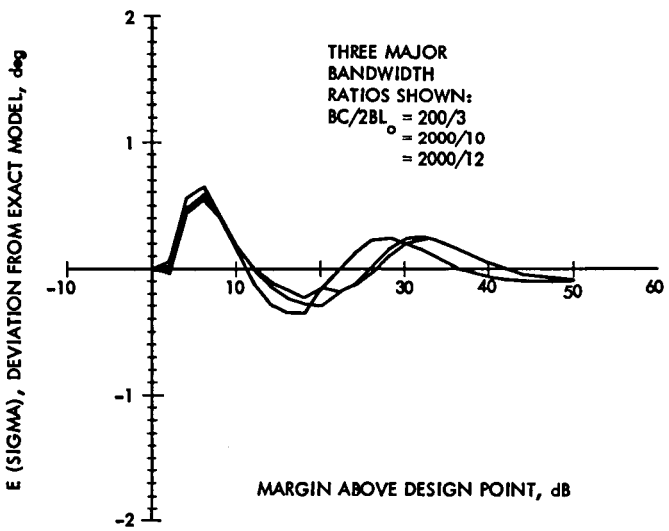


Fig. 3. Deviation between exact and exponential approximation models

Appendix

Receiver Thermal Noise Models

(A) Original Linear Model

Jaffee/Rechtin Linear Loop

$$(1)\quad \sigma_\varphi^2 = \frac{1}{2M}\left[\frac{1}{3} + \frac{2}{3}\frac{\alpha}{\alpha_o}\right] \times \Gamma$$

$$(2)\quad \frac{\alpha}{\alpha_o} = \left[\frac{1 + \frac{4}{\Pi} \times \frac{1}{PH}}{1 + \frac{4}{\Pi} \times \frac{1}{PH} \times \frac{1}{M}}\right]^{1/2}$$

$$PH = PH\,(1)$$

$$(3)\quad \Gamma = 1$$

$$(4)\quad E(\sigma) = \sigma_\varphi \qquad \text{(no conversion)}$$

$$\sigma^o = \frac{180}{\pi}\sigma_\varphi$$

This model applies in the strong-signal ($M > 20$ dB) range only. It considers the detector completely linear, and the limiter as a device that simply holds its power output constant. The model is not considered in the comparison. $E(\sigma)$ is in error by nearly 50 percent at unity margin. It does, however, provide criteria for DSN receiver design, which normally operates near strong-signal levels.

(B) Study Model No. 1

Tausworthe Nonlinear Approximations (used for DSN testing)

$$(1)\quad \sigma_\varphi^2 = \frac{1}{2M} \times \left[\frac{1}{3} + \frac{2}{3}\frac{\alpha}{\alpha_o}\right] \times \Gamma$$

$$(2)\quad \alpha = \left[\frac{0.7854\,PH + 0.4768\,PH^2}{1.0 + 1.024\,PH + 0.4768\,PH^2}\right]^{1/2}$$

$$\alpha_o = \alpha\,[PH(M = 1)]$$

$$(3)\quad \Gamma = \frac{1 + PH}{0.862 + PH}$$

$$(4)\quad \sigma_\varphi^2 = \frac{\left(1 - e^{-a^2}\right)^2}{a^2}$$

$$(5)\quad (1+r)\frac{\left[1 + \frac{2a^2\left(1 - e^{-a^2} - a^2 e^{-a^2/2}\right)}{r\left(1 - e^{-a^2}\right)^2}\right]^{1/2}}{\frac{1 + r\left(1 - e^{-a^2}\right)}{a^2}}$$

$$r = 2\frac{\alpha}{\alpha_o} = r(M)\,\sigma_\varphi^2 = \sigma_\varphi^2\,(M)$$

$$\sigma^2 = \frac{\pi^2}{3}\left[1 - e^{-3a^2/\pi^2\,(1\,+\,0.13a^2)}\right]$$

$$\sigma^o\,(M) = \frac{180}{\pi}\,\sigma^2\,(M)$$

NOTE: Parametric equation in α^2 until agreement to 10^{-4} in σ^2. Place σ_φ^2 in σ^2 for modulo 2π.

a^2 is the variance of a phase process and is iterated before being converted to modulo 2π. The algorithm is called linear spectral analysis approximation, and also uses algebraic approximations to EXACT expressions for α and Γ.

(C) Study Model No. 2

Simon/Lindsey and Yuen EXACT Conversion (near approximation)

$$(1)\quad \sigma_\varphi^2 = \frac{1}{2M} \times \left[\frac{1}{3} + \frac{2}{3}\frac{\alpha}{\alpha_o}\right] \times \Gamma$$

$$(2)\quad \alpha = \frac{\pi}{4}\,PH\,e^{-PH/2}\left[I_o\left(\frac{PH}{2}\right) + I_1\left(\frac{PH}{2}\right)\right]$$

$$\alpha_o = \alpha \;@\; M = 1$$

$$(3)\ \Gamma = \frac{1 - e^{-PH}}{\frac{\pi}{4} PH \times e^{-PH} \times \left[I_o \left(\frac{PH}{2} \right) + I_1 \left(\frac{PH}{2} \right) \right] \left[1 + \left(\frac{3.448}{\pi} - 1 \right) e^{-PH(1 - \pi/4)} \right]}$$

(INEXACT)

$$(4) \qquad \rho = \frac{1}{\sigma_\varrho^{\,2}}$$

$$E(\sigma^2) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\phi^2 \, e^{\rho \cos(\phi)} \, d\phi}{I_o(\rho)}$$

$$= \frac{\pi^2}{3} + 4 \sum_{N=1}^{\infty} \left[(-1)^N \frac{I_n(\rho)}{I_o(\rho)} \cdot \frac{1}{N^2} \right]$$

$$\sigma^o = \frac{180}{\pi} \cdot \sigma$$

$$I_n(x) = \left(\frac{X}{2} \right)^2 \sum_{M=0}^{\infty} \frac{\left(\frac{X}{2} \right)^{2M}}{M!\ (M+N)!}$$

This is the most accurate model presently available. It was used as the standard for comparison.

Deep Space Network Feasibility Study of Terminating Southern California Edison Electrical Service to Goldstone

J. L. Koh
Deep Space Network Support Section

A preliminary study on cost comparison of purchased power versus generated power for the Goldstone Deep Space Communications Complex (GDSCC) shows that there are economic incentives to generate the A-C power requirements for the complex. The justification can only be sustained if the waste heat from the cooling water and/or exhaust systems is recovered for reuse.

I. Introduction

The purpose of this article is to explore the feasibility of eventually terminating Southern California Edison's electrical service to the Goldstone Deep Space Communications Complex (GDSCC) by totally supplying electrical power to the complex from existing GDSCC power generation systems. The advantages of this method of operation include:

(1) Utilization of the diesel generator waste heat to reduce the overall complex energy consumption by 25 to 50 percent;

(2) More consistent power supply source, hence better system reliability;

(3) Optimization of equipment and manpower usage.

It is understood that present GDSCC power generation systems cannot generate power as economically as Southern California Edison when comparing kW for kW; however, the fact that there is a need to maintain a standby diesel power generation system to support Edison's commercial power has, in itself, increased the kWh cost on a purchase basis, as compared to the cost of running on 100 percent GDSCC-generated power alone.

II. Cost Comparison

Attached is an overview analysis of kW-for-kW cost comparison, between commercially purchased power and DSN-generated power (for GDSCC alternating current (A-C) power requirements), based on March 1979 operating costs and 1978 consumption. Table 1 summarizes the details of cost estimates (see Appendix A for calculations), which indicate that the difference is about $0.102/kWh in favor of Edison's purchased power. Offsetting the higher-generated cost/kWh is the waste-energy available in the diesel engine cooling water and exhaust systems which, if recovered for reuse, would substantially reduce the generated power cost. Table 2 shows that the waste-energy has an equivalent of about 2,766,900 liters of diesel fuel per year, worth about $321,624 at $0.11624/liter. Reuse of this waste energy could also substantially reduce the current liquified petroleum gas (LPG) consumption. In 1978, about 374,300 liters of LPG were consumed at GDSCC. At $0.108 per liter, this is equivalent to an expenditure of approximately $40,000 per year. Data sources used in the cost estimates and calculations are listed in Table 3.

III. Energy Conservation and Recovery Systems

Waste energy constitutes about 67 percent of diesel fuel consumption in power generation. Recovering 25 percent of this waste energy for reuse could reduce the total power consumption by about 25 to 50 percent, depending on reuse efficiencies. This reduction should contribute significantly to JPL's goals in energy conservation programs. The waste energy recovered could be used for:

(1) Comfort heating

(2) Absorption chillers

(3) Water distillation

(4) Steam generation, and

(5) Preheating of equipment, etc.

Various successful systems for recovering waste energy from diesel generator power plants have been in existence for several years. In the DSN, two diesel generator waste heat utilization systems are in operation. At DSS 61/63 heat from the power plant engine cooling water system provides comfort heat for the complex. The DSS 62 waste heat recovery system provides both comfort heat and operation room rack cooling. The system most suitable for GDSCC will depend upon the type and cost of the equipment needed to recover waste energy and how this energy is utilized. In-depth studies will be required prior to initiation of a design effort.

IV. Action Plan

Present analysis indicates that there are potential economic incentives to further pursue the concept of supplying (all) power for GDSCC from existing power generation systems. However, before the concept can be implemented the following actions need to be carried out:

(1) A thorough analysis to accurately define the generated kWh cost;

(2) Determine total fossil fuel utilized for generated power vs. purchased power;

(3) Establish diesel fuel availability;

(4) Establish equipment viability in using the potential waste heat energy;

(5) Estimate a cost factor for the change over; and

(6) Draw up a schedule for implementation.

Table 1. Cost comparison between purchased power and generated power

Power consumed (kWh cost)	1978 Actual purchased power mode		Generated power (100% time on generators)
	91.7% time on commercial	8.3% time on generators	
1978 Power consumption	13,236,555 kWh/yr	1,194,560 kWh/yr	14,431,115 kWh/yr
Diesel fuel cost, kWh	$0.0370	–	–
Diesel fuel cost, kWh	–	$0.0332	$0.0332
Preventive maintenance cost/kWhr	–	$0.0100	$0.0067
Overhaul cost/kWh	–	$0.0097	$0.0032
Lube cost/kWh	–	$0.0043	$0.0009
Operator delta cost/kWhr	–	$0.0135	$0.0060
kWh cost	$0.0370	$0.0707	$0.0500
Average cost/kWh	$0.0398[a]		$0.0500

[a]Computations: ($0.0370 X 0.917) + ($0.0707 X 0.083) = $0.0339 + $0.0059
= $0.0398

Table 2. GDSCC—Power generation potential energy savings[a]

Fuel Expenditure:

30% Diesel fuel converted to electricity

30% Diesel fuel loss due to engine cooling water system

37% Diesel fuel loss due to engine exhaust

3% Diesel fuel loss due to radiation and other contributing factors

Energy available for reuse	Computations
Estimated liters of diesel/kWh generated or total liters of diesel/year	= 0.28617 l/kWh = 0.28617 l/kWh X 14,431,115 kWh/yr = 4,129,730 liters of diesel/yr
Equivalent fuel loss due to cooling water system	= 4,129,730 X 0.30 liters of diesel/yr = 1,238,900 liters of diesel/yr
Equivalent fuel loss due to exhaust system	= 4,129,730 X 0.37 liters of diesel/yr = 1,528,000 liters of diesel/yr
Total energy available for reuse @ 100% efficiency	= 1,238,900 + 1,528,000 liters of diesel/yr = 2,766,900 liters of diesel/yr
This energy is worth	= 2,766,900 X $0.11624/yr = $321,624/yr

[a]Based on 1 kWh generated requiring 0.28617 liters of diesel fuel.

Table 3. Data sources

Caterpillar Model G389 and G399 specifications

GDSCC power consumption ECM[a]/400

GDSCC diesel fuel consumption ECM/200

GDSCC LPG consumption ECM/100

GDSCC utilities summary ECM/600

GDSCC meter reference file ECM/610

GDSCC monthly utility usage reports

GDSCC experience on power generation

1978 total power = 14,431,115 kWh

1978 generated power = 1,194,560 kWh

LPG cost/liter = $0.108 (Mar. 1979 price) (1 gal = $0.41)

Diesel cost/liter = $0.1162 (Mar. 1979 price) (1 gal = $0.44)

Lube cost/liter = $0.449 (Mar. 1979 price) (1 gal = $1.70)

1 gallon = 3.7853 liters

1 kilogram diesel = 17,100 kcal equivalent heat value (1 pound = 19,500 Btu)

Edison power cost/kWh = $0.037 (Feb. 1979 price)

[a]ECM = Energy Consumption Management Files.

Appendix A

Cost Comparison of Commercially Purchased Power vs. Generated Power for GDSCC—Cost Estimates and Calculations

1978 Power and Diesel Consumption

	DSS 10	DSS 11	DSS 12	DSS 14	Total[a]
Purchased power kWh (metered at station)	6,823,434	–	–	6,413,121	13,236,555
Generated power kWh	–	213,120	171,040	810,400	1,194,560
Total power consumption kWh	–	–	–	–	14,431,115
Purchased diesel fuel	–	–	–	–	341,847 liters (90,300 gal)

Fuel Cost ($/kWh)

Basis:

Percent power generated = $\frac{1{,}194{,}560 \text{ kWh}}{14{,}431{,}115 \text{ kWh}} \times 100 = 8.3\%$

Liters of diesel fuel/kWh = $\frac{341{,}847 \text{ l/yr}}{1{,}194{,}560 \text{ kWh/yr}} = 0.28617 \text{ l/kWh}$

Diesel fuel cost = $0,11624/liter

Diesel fuel cost $/kWh

Diesel fuel cost/kWh = $0.11624/l × 0.28617 l/kWh
= $0.0332/kWh

[a]For entire GDSCC.

Preventative maintenance costs – ($/kWh)

Basis:		Computations:
(1)	100% generation (i.e., 14,431,115 kWh/yr)	Model G398 costs $1.20/engine-hour-run Model G399 costs $2.00/engine-hour run Smaller model cost $1.00/engine-hour-run
(2)	8.3% generation (i.e., 1,194,560 kWh/yr)	Above cost X (factor of 1.5)[a] X % engine-hour-run
Preventive maintenance costs:		Computations:
For 100% generation		= (Model & no. of engines) X (hours-run/yr) X (maint. costs/hr) = [(4 X G399) $2.00 + (10 X G389) $1.20 + (2 X others) $1.00] 4380 hr = [8 + 12 + 2] 4380 = $96,360.00
	Cost/kWh	= $\frac{\$96,360.00/yr}{14,431,115\ kWh/yr}$ = $0.00667/kWh
For 8.3% generation		= (Above) X 1.5 X 8.3% X engine-hour-run = $96,360.00 X 1.5 X 0.083 = $11,996.82/yr
	Cost/kWh	= $\frac{\$11,996.82/yr}{1,194,560\ kWh/yr}$ = $0.01004/kWh

[a]Factor of 1.5 is used to take into account additional frequent start-and-stop operations.

Overhaul costs/kWh and lube costs/kWh

Basis:	Computations:
(1) 100% generation (i.e., 14,431,115 kWh/yr)	
Hours run between overhaul	= 30,000 hr
Lube oil requirement	= 30,300 l/yr (total) (8000 gal/hr) (estimated)
(2) 8.3% generation (i.e., 1,194,560 kWh/yr)	
Hours run between overhaul	= 10,000 hr
Lube oil requirement	= 11,400 l/yr (total) (3000 gal/yr) (estimated)
(3) 1 liter lube = $0.449 (Mar. 27, 1979) ($1.70 gal)	
(4) Each overhaul cost	= $20,000.00

Overhaul costs ($/kWh):	Computations:
For 100% generation no. of engines requiring overhaul	$= \frac{\text{(no. of engines) X (hr run per year)}}{\text{hours run between overhaul}}$
	$= \frac{16 \times 4380}{30{,}000}$
	= 2.336 no. of engines/yr
Overhaul costs/yr	= 2.336 X $20,000.00/yr
	= $46,720.00/yr
Costs/kWh	$= \frac{\$46{,}720.00/yr}{14{,}431{,}115\ kWh/yr} = \$0.00323/kWh$
For 8.3% generation no. of engines requiring overhaul	$= \frac{16 \times 364}{10{,}000}$
	= 0.5824 no. of engines/yr
Overhaul costs/yr	= 0.5824 X $20,000.00/yr
	= $11,648.00/hr
Costs/kWh	$= \frac{\$11{,}648.00/yr}{1{,}194{,}560/kWh/yr}$
	= $0.00975/kWh

Lube cost ($6Kwh):	Computations:
Cost/kWh	$= \frac{\text{(Cost per gal of lube) X (amount used per yr)}}{\text{generated power per year}}$
For 100% generation	$= \frac{\$0.449 \times 30{,}300\ liters}{14{,}631{,}115\ kWh/yr}$
	= $0.00093/kWh
For 8.3% generation	$= \frac{\$0.449 \times 11{,}400\ liters}{1{,}194{,}560\ kWh/yr}$
	= $0.00428/kWh

Power house operator – delta cost ($/kWh)

Basis:		Computations
(1)	@ 100% generation (i.e., 14,431,115 kWh/yr) requires 3 operators on 3 cycle, 8 hr shifts, total manning	= 12 operators
(2)	@ 8.3% generation (i.e., 1,194,560 kWh/yr) requires 3 operators on 2 cycle, 12 hr shifts, total manning	= 9 operators
(3)	Existing manpower availability	= 9 operators
(4)	Current wages and associated costs per operator/yr	= $28,700.00/yr
(5)	2 cycle, 12 hr shifts, involves operator on overtime on actual time worked @ 1.5 times wages, and associated costs but no additional operators	
(6)	3 cycle, 8 hr shifts, involves no overtime but requires hiring 3 additional operators	

Operator delta costs ($/kWh):	Computations:
For 100% generation	= 3 operators X $28,700.00/yr
	= $86,000.00/hr
Delta cost/kWh	= $\frac{\$86,100.00/yr}{14,431,115\ kWh/yr}$ = $0.00595/kWh
For 8.3% generation	= (9 operators) X (50% overtime X 1.5 rate) X (actual time worked/yr)[a]
	= 9 X $14,350.00/yr X 1.5 X 0.083
	= $16,079.16/yr
Delta cost/kWh	= $\frac{\$16,079.16/hr}{1,194,560\ kWh/hr}$ = $0.01346/kWh

[a]Assuming same as 8.3% generation time.

X-Band Sampling by the Occultation Data Assembly

R. B. Roth
Deep Space Network Operations Section

Extraneous peaks appear in the power spectrum of X-band radio science data from the Occultation Data Assembly after it has been processed. The article shows that the cause lies in an incompatibility between the hardware implementation of Nyquist sampling and the software processing. This also shows that a forthcoming change in the hardware will eliminate the problem.

I. Introduction

The Occultation Data Assembly (ODA) was recently installed in the Deep Space Network's 64-meter subnet. The assembly digitizes and records the received carrier signal, both S- and X-band, from various spacecraft. Several sampled bandwidths and analogue-to-digital converter (ADC) assignments are available, and have been used for the Pioneer Venus Orbiter, and Voyager 1 and 2 Jupiter occultations.

The digitized data provide sufficient information for reconstruction of the carrier for investigation of relatively short-time scale effects. There is, however, an incompatibility between the way the X-band receiver is sampled in a commonly used mode, and the software which is used to reduce the X-band data. The problem is discussed below and examples are presented for the effects on a carrier signal, and on the filter bandpass characteristics.

II. The Problem

The ODA is commonly configured to sample the S-band receiver with one of its four ADCs, and sample the X-band receiver with the remaining three ADCs. The receiver bandwidths have a 1:3, S- to X-band ratio. All ADCs have equal sample frequencies, but the second and third X-band ADCs are staggered in time from the first. This is accomplished by multiplying the sampling frequency by 20, and sampling with the first ADC on the first cycle (simultaneous with the S-band ADC), the second ADC on the eighth cycle and finally, the third ADC on the fifteenth cycle. This gives X-band samples at 7/20, 7/20, 6/20 intervals with respect to the S-band samples. Thus, the X-band receiver is sampled irregularly, giving rise to extraneous power peaks when the data is Fourier transformed; these peaks are purely artifacts of the sampling process. Likewise, the filter bandpass characteristics are distorted by the sampling.

Although the waveform could be correctly reconstructed from the data by modification of the analysis software, the ODA hardware will be modified to sample regularly at S- and X-band. Currently, the X-band power spectrum obtained is, typically, as shown in Fig. 1, for a single-sampled signal. The relative amplitudes and frequencies of the sampling artifacts are computed below.

III. Calculations

The calculations that follow show first, how the signal can be reconstructed from its samples; second, the result if such a reconstruction process were applied to irregularly sampled waveforms; third, the transfer function for the "digital filter" formed by the ODA plus FFT (Fast Fourier Transform) software; and fourth, its effects on a nearly monochromatic input (Fig. 1).

Any waveform that is bandwidth limited may be represented by discrete samples if the samples are taken at a frequency of twice that bandwidth. The open loop receiver bandwidth limits the signal.

If $f(t)$ is the waveform, the sampled waveform can be represented by

$$f_s(t) = f(t) \sum_{n=-\infty}^{\infty} \delta(t - n\Delta T), \Delta T = \text{sample interval}$$

The Fourier transform of this is

$$F_s(\omega) = \frac{1}{2\pi} F(\omega) * \omega_0 \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_0), \quad \omega_0 = \frac{2\pi}{\Delta T}$$

Performing the convolution and substituting for ω_0,

$$F_s(\omega) = \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} F(\omega - n\omega_0)$$

which is just $F(\omega)$ repeated every ω_0 radians. If the bandwidth of the waveform is less than half the sample frequency, there is no overlapping of the terms of $F_s(\omega)$. $F(\omega)$ can be recovered from $F_s(\omega)$ by multiplying by a gate function

$$G_{\omega_0}(\omega) = \begin{cases} 1, & |\omega| \leq \omega_0/2 \\ 0, & |\omega| > \omega_0/2 \end{cases}$$

$$F(\omega) = \Delta T F_s(\omega) G_{\omega_0}(\omega)$$

or, in the time domain,

$$f(t) = \Delta T f_s(t) * g_{\omega_0}(t)$$

$$g_{\omega_0}(t) = \frac{\omega_0}{2\pi} \operatorname{sinc} \frac{\omega_0 t}{2}, \quad \operatorname{sinc} x = \frac{\sin x}{x}$$

Thus $f(t)$ can be reconstructed from $f_s(t)$. Performing the convolution,

$$f(t) = \sum_{n=-\infty}^{\infty} f(n\Delta T) \operatorname{sinc} \frac{\omega_0}{2} (t - n\Delta T)$$

The irregularly sampled function, $\tilde{f}(t)$, can be written in terms of $f(t)$:

$$\begin{aligned} \tilde{f}(t) = \sum_{n=-\infty}^{\infty} \{ & f(3n\Delta T) \operatorname{sinc} [t - 3n\Delta T] \omega_0/2 \\ & + f(3n\Delta T + \epsilon) \operatorname{sinc} [t - (3n+1)\Delta T] \omega_0/2 \\ & + f(3n\Delta T + 2\epsilon) \operatorname{sinc} [t - (3n+2)\Delta T] \omega_0/2 \} \end{aligned}$$

In the X-band case, ϵ is 7 clock counts. Taking the transform,

$$\begin{aligned} \tilde{F}(\omega) = \frac{2\pi}{\omega_0} \sum_{n=-\infty}^{\infty} \{ & f(3n\Delta T) \exp -i3n\Delta T\omega \\ & + f(3n\Delta T + \epsilon) \exp -i(3n+1)\Delta T\omega \\ & + f(3n\Delta T + 2\epsilon) \exp \\ & -i(3n+2)\Delta T\omega \}, \quad |\omega| < \frac{\omega_0}{2} \end{aligned}$$

This may be rewritten in terms of $F(\omega)$. Consider:

$$f_a(t) = \sum_{n=-\infty}^{\infty} f(t + a\epsilon)\, \delta(t - 3n\Delta T)$$

$$F_a(\omega) = \frac{1}{2\pi} F(\omega) \exp ia\epsilon\omega * \frac{2\pi}{3\Delta T} \sum_{n=-\infty}^{\infty} \delta \left(\omega - \frac{2\pi n}{3\Delta T}\right)$$

$$= \sum_{-\infty}^{\infty} \frac{1}{3\Delta T} F \left(\omega - \frac{n\omega_0}{3}\right) \exp ia\epsilon \left(\omega - \frac{n\omega_0}{3}\right)$$

Using an equivalent form of $f_a(t)$,

$$f_a(t) = \sum_{n=-\infty}^{\infty} f(3n\Delta T + a\epsilon)\,\delta\,(t - 3n\Delta T)$$

$$F_a(\omega) = \sum_{n=-\infty}^{\infty} f(3n\Delta T + a\epsilon)\exp -3in\Delta T\omega$$

Multiplying both expressions for $F_a(\omega)$ by $(2\pi/\omega_0)\exp -ia\Delta T\omega$,

$$\frac{2\pi}{\omega_0}\sum_{n=-\infty}^{\infty} f(3n\Delta T + a\epsilon)\exp -i\,(3n + a)\,\Delta T\omega =$$

$$\frac{1}{3}\sum_{n=-\infty}^{\infty} F\left(\omega - \frac{n\omega_0}{3}\right)\exp ia\omega\,(\epsilon - \Delta T)\exp -ia\epsilon\frac{n\omega_0}{3}$$

Taking a to be 0, 1, or 2 gives all the terms of $\tilde{F}(\omega)$:

$$\tilde{F}(\omega) = \frac{1}{3}\sum_{n=-\infty}^{\infty} F\left(\omega - \frac{n\omega_0}{3}\right)(1 + \exp iQ$$

$$+ \exp 2iQ),\quad |\omega| < \frac{\omega_0}{2}$$

where $Q = \omega(\epsilon - \Delta T) - \frac{n}{3}\omega_0\epsilon$. Note that $\tilde{F}(\omega) = F(\omega)$ for $\epsilon = \Delta T$.

IV. Conclusion

A. Effect on a Monochromatic Signal

A typical power spectrum at X-band is shown in Fig. 1. For the purposes of this study, the received signal may be approximated by a delta function.

$$F(\omega) = \delta\left(\omega - \frac{\omega_0}{4}\right)$$

yields

$$\tilde{F}(\omega) = A_0\,\delta\left(\omega - \frac{\omega_0}{4}\right) + A_1\,\delta\left(\omega + \frac{\omega_0}{12}\right)$$

$$+ A_2\,\delta\left(\omega + \frac{5\omega_0}{12}\right)$$

where the amplitudes A_0, A_1, A_2 are determined by

$$A_{-n} = \frac{1}{3}\left\{1 + \cos\left[\omega(\epsilon - \Delta T) - \frac{n}{3}\omega_0\epsilon\right]\right.$$

$$\left. + \cos 2\left[\omega(\epsilon - \Delta T) - \frac{n}{3}\omega_0\epsilon\right]\right\}$$

The A_1 and A_2 terms represent the extraneous power peaks. Physically, the components at frequencies $-\omega_0/12$ and $-5\omega_0/12$ may be considered to be at $\omega_0/12$ and $5\omega_0/12$. This is so because the power at a frequency is proportional to the square of the amplitude of that frequency component. Thus,

$$[A\cos(-\omega)]^2 = A^2\cos^2(-\omega) = (A\cos\omega)^2$$

and

$$[A\sin(-\omega)]^2 = A^2\sin^2(-\omega) = (A\sin\omega)^2$$

so that any component at $-\omega$ may be considered to be at $+\omega$.

B. Effect on Transfer Curve

The initial bandwidth limiting that occurs in the open loop receiver gives a transfer curve similar to that shown in Fig. 2. The equation for $\tilde{F}(\omega)$ has been computed for that transfer function to give the new transfer curve. The change in power in dB is plotted in Fig. 3 for the following case:

$\Delta T = 3.33 \times 10^{-5}$ s

$\omega_0 = 188496$ rad/s = 30 KHz

$\epsilon = 3.5 \times 10^{-5}$ s

power change = $20\log\frac{\tilde{F}(\omega)}{F(\omega)}$

The transfer of the signal across the digital filter is nearly linear (deviations around 1 dB).

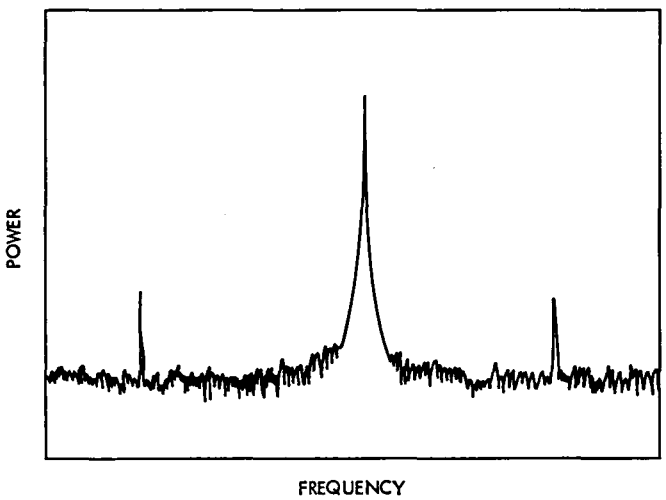


Fig. 1. X-band power spectrum from ODA

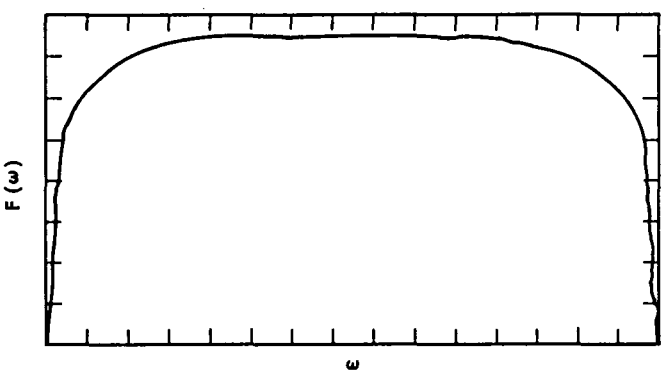


Fig. 2. Open loop receiver transfer curve

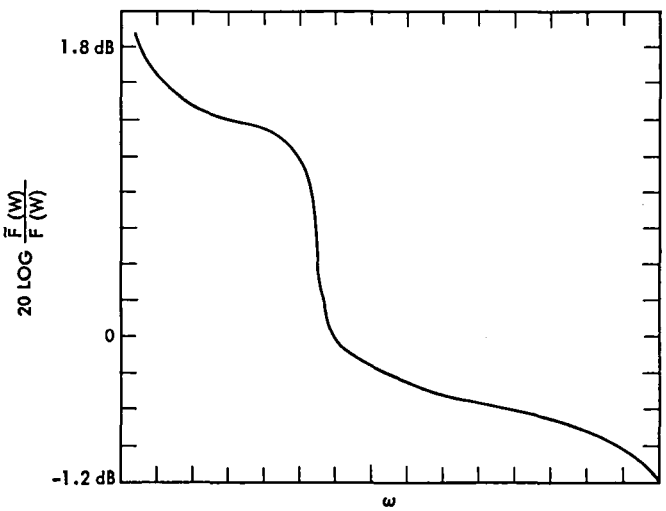


Fig. 3. Difference in signal power between ODA with regular sampling and ODA with irregular sampling

End of Document