

The Telecommunications and Data Acquisition Progress Report 42-113

January–March 1993

E. C. Posner
Editor

(NASA-CR-193144) THE
TELECOMMUNICATIONS AND DATA
ACQUISITION REPORT Progress Report,
Jan. - Mar. 1993 (JPL) 209 p

N93-29586
--THRU--
N93-29603
Unclas

G3/32 0167920

May 15, 1993

NASA

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The Telecommunications and Data Acquisition Progress Report 42-113

January–March 1993

E. C. Posner
Editor

May 15, 1993



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 a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

Preface

This quarterly publication provides archival reports on developments in programs managed by JPL's Office of Telecommunications and Data Acquisition (TDA). In space communications, radio navigation, radio science, and ground-based radio and radar astronomy, it reports on activities of the Deep Space Network (DSN) in planning, supporting research and technology, implementation, and operations. Also included are standards activity at JPL for space data and information systems and reimbursable DSN work performed for other space agencies through NASA. The preceding work is all performed for NASA's Office of Space Communications (OSC). The TDA Office also performs work funded by another NASA program office through and with the cooperation of OSC. This is the Orbital Debris Radar Program with the Office of Space Systems Development.

The TDA Office is directly involved in several tasks that directly support the Office of Space Science (OSS), with OSC funding DSN operational support. In radio science, *The TDA Progress Report* describes the spacecraft radio science program conducted using the DSN. For the High-Resolution Microwave Survey (HRMS), the report covers implementation and operations for searching the microwave spectrum. In solar system radar, it reports on the uses of the Goldstone Solar System Radar for scientific exploration of the planets, their rings and satellites, asteroids, and comets. In radio astronomy, the areas of support include spectroscopy, very long baseline interferometry, and astrometry.

Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund that involve the TDA Office are included.

This and each succeeding issue of *The TDA Progress Report* will present material in some, but not necessarily all, of the following categories:

OSC Tasks:

- DSN Advanced Systems
 - Tracking and Ground-Based Navigation
 - Communications, Spacecraft-Ground
 - Station Control and System Technology
 - Network Data Processing and Productivity
- DSN Systems Implementation
 - Capabilities for Existing Projects
 - Capabilities for New Projects
 - New Initiatives
 - Network Upgrade and Sustaining
- DSN Operations
 - Network Operations and Operations Support
 - Mission Interface and Support
 - TDA Program Management and Analysis
- Ground Communications Implementation and Operations
- Data and Information Systems
- Flight-Ground Advanced Engineering
- Long-Range Program Planning

OSC Cooperative Tasks:

- Orbital Debris Radar Program

Contents

OSC TASKS DSN Advanced Systems TRACKING AND GROUND-BASED NAVIGATION

Improved Linear Ion Trap Physics Package.....	1 - 1
J. D. Prestage NASA Code 310-10-62-15-00	
The Effect of Clock, Media, and Station Location Errors on Doppler Measurement Accuracy.....	7 - 2
J. K. Miller NASA Code 310-10-63-84-02	
Precise Tracking of the Magellan and Pioneer Venus Orbiters by Same-Beam Interferometry—Part II: Orbit Determination Analysis	22 - 3
W. M. Folkner, J. S. Border, S. Nandi, and K. S. Zukor NASA Code 310-10-63-88-01	

COMMUNICATIONS, SPACECRAFT-GROUND

High-Power Ka-Band Amplifier.....	37 - 4
R. Cormier NASA Code 310-20-64-22-00	
Experimental and Modal Verification of an Integral Equation Solution for a Thin-Walled Dichroic Plate With Cross-Shaped Holes	46 - 5
L. W. Epp and P. H. Stanton NASA Code 310-20-64-86-02	
Gravity Referenced Elevation Encoder Development.....	63 - 6
R. E. Goddard NASA Code 310-20-65-67-00	
A Note on the Computation of Antenna Blocking Shadows	74 - 7
R. Levy NASA Code 310-20-65-86-04	
Maximum and Minimum Return Losses From a Passive Two-Port Network Terminated With a Mismatched Load.....	80 - 8
T. Y. Otoshi NASA Code 310-20-65-86-08	
The General Theory of Convolutional Codes.....	89 - 9
R. J. McEliece and R. P. Stanley NASA Code 310-30-71-83-04	
Uncorrectable Sequences and Telecommand	99 - 10
L. Ekroot, R. J. McEliece, S. Dolinar, and L. Swanson NASA Code 310-30-71-83-04	

DSN Systems Implementation

Galileo Post-Gaspra Cruise and Earth 2 Encounter	106 - 11
P. E. Beyer and M. M. Andrews NASA Code 314-40-22-37-01	

CAPABILITIES FOR EXISTING PROJECTS

DSN Support of Mars Observer 118 ⁻¹²
M. R. Traxler
NASA Code 314-40-22-30-01

CAPABILITIES FOR NEW PROJECTS

Signal-to-Noise Ratio Losses in Full Spectrum Combining of Signals With a
Downconverted Subcarrier 123 ⁻¹³
Y. Fera and J. Statman
NASA Code 314-30-61-02-10

NETWORK UPGRADE AND SUSTAINING

DSS-14 Subreflector Actuator Dynamics During the Landers Earthquake 130 ⁻¹⁴
R. Levy and D. Strain
NASA Code 314-30-45-10-21

An Analysis of the 70-Meter Antenna Hydrostatic Bearing by Means of Computer
Simulation 147 ⁻¹⁵
R. D. Bartos
NASA Code 314-30-42-10-87

Flight-Ground Advanced Engineering

Acoustic Charge Transport Technology Investigation for Advanced
Development Transponder 174 ⁻¹⁶
S. Kayalar
NASA Code 315-91-10-13-12

OSS TASKS Radio Science

The Flight Performance of the Galileo Orbiter USO 184 ⁻¹⁷
D. D. Morabito, T. P. Krisher, and S. W. Asmar
NASA Codes 314-40-31-30-26 and 889-54-62-01-00

N93-29587

500683 May 15, 1993

S1-32

167921

P- 6

An Improved Linear Ion Trap Physics Package

J. D. Prestage

Communications Systems Research Section

This article describes an improvement in the architecture of the physics package used in the Linear Ion Trap (LIT)-based frequency standard recently developed at JPL. This new design is based on the observation that ions can be moved along the axis of an LIT by applied dc voltages. The state selection and interrogation region can be separated from the more critical microwave resonance region where the multiplied local oscillator signal is compared with the stable atomic transition. This separation relaxes many of the design constraints of the present units. Improvements include increased frequency stability, and a substantial reduction in size, mass and cost of the final frequency standard.

I. Introduction

The primary reason that ions confined to electromagnetic traps are the basis for stable frequency standards is that in the containerless environment of a trap, processes that equalize atomic state populations and destroy coherence within the state-prepared atomic ensemble are very weak. In past work coherence times have been measured in an ensemble of trapped Hg^+ ions of over 30 sec on the 40.5-GHz transition. Such weak relaxation has permitted a resonance line width $\Delta\nu$ as small as 17 mHz on the 40.5-GHz transition. This line-Q ($\nu/\Delta\nu = 2 \times 10^{12}$) is the highest ever measured on a microwave atomic transition [1,2].

One disadvantage of a trapped ion-based frequency standard is the relatively low density of the confined ions. This leads to a low signal-to-noise ratio (SNR) in the detected atomic resonance and consequently limits clock stability. This situation was greatly improved by the use of the linear ion trap [2] where the number of trapped ions was increased by more than 10 over conventional hyperbolic traps. The stability reached with the improved SNR

of the Linear Ion Trap (LIT) together with the high line-Q has led to a tenfold improvement in clock stability over conventional ion-based clocks. Indeed, the present LIT-based clock shows stability competitive with the best H-masers for averaging times less than 10,000 sec and exceeds H-maser stability beyond 10,000 sec, making it the most stable of all clocks for long-term stability. The LIT clock is, however, relatively recent in its development and its configuration continues to evolve. The purpose of this article is to propose an improvement in the architecture of the current physics package which will lead to improved long-term stability, and a substantial reduction in size, mass and cost of the final frequency standard.

II. Current LIT Operation

The present form of the LIT physics unit for frequency standard operation is shown in Fig. 1. Ions are created inside the linear trap by an electron pulse which ionizes a weak vapor of parent neutral atoms introduced into the vacuum system from a heated HgO isotopic source. They are held in the four-bar linear trap shown inside the region

enclosed by magnetic shields [3]. Before the stable atomic frequency of the trapped ions can be compared with the multiplied output of a local oscillator (LO) or flywheel oscillator, a population difference between the hyperfine levels of the ground state must be created. This is done via optical pumping with ultraviolet (UV) light from a laser or a discharge lamp. As shown in Fig. 1, optical pumping is carried out by a ^{202}Hg lamp whose output light is collected and focused onto the central portion of the ion cloud.

After state preparation, microwave radiation derived from the LO is passed onto the ions to make a frequency comparison with the stable atomic resonance. During this period it is crucial that the atomic resonance frequency not be perturbed by any changes in the trap environment. Such fluctuations will be transferred to the servoed LO, thereby degrading clock stability. Because the state selecting light will shift the atomic clock resonance, it is switched off during the microwave interrogation time. Following the microwave pulse, the lamp is turned on again to determine whether the microwave radiation has changed the population of the hyperfine levels of the atomic ion. Any frequency detuning of the LO from the reference atomic frequency will change the fluorescent light intensity measured when the lamp is turned on. These fluorescence changes are converted to a voltage and fed back to the frequency control port of the LO to keep it on frequency with the atomic oscillator.

It is apparent that the trap region is operated in two modes in the process of steering an LO. In the first mode, the atomic ions are prepared for microwave frequency comparison with the LO. In the second mode, the atomic frequency is compared with the multiplied output of the LO. During the state preparation and interrogation mode, there are no stringent requirements on environmental isolation or regulation, whereas during the resonance mode, great care must be exercised in regulation of the atomic environment to insure stable operation.

III. Improved LIT Design

Figure 2 shows a configuration of an ion trap-based clock where these modes of operation are carried out in two separate regions. The trap shown is a lengthened version of the LIT of Fig. 1. In the new architecture of Fig. 2 the region labeled ion loading and fluorescence is somewhat smaller in size than the trap shown in Fig. 1. A second linear trap has been added for a microwave resonance region.

A dc break forms a junction between these two regions to allow ions to pass from one region to another. The rf

trapping voltage for transverse containment of the ions is continuous across this dc break. When the dc voltage levels of all four trap rods are the same in both the upper and lower regions, thermal motion of the ions will carry them through the junction with no change in axial velocity. When the four trap rods in the upper region are at negative dc voltage with respect to the lower region trap rods, ions within a trap radius or so of the junction will be transported across the junction into the upper region. Only the ions near the junction will experience the electric field forcing them across the gap. Since each ion is in thermal motion along the axis of the trap, it will reach the junction within a trap-length transit time (typically about 1 msec) and then be pulled into the upper region, emptying the lower region of ions. Similarly, when the upper region is dc biased positive with respect to the resonance region, all ions will be transported from the fluorescence region to the ion resonance region. The dc potential along the length of the trap that controls the region the ions occupy is shown in Fig. 3.

The separation of the clock resonance region from the optical and ion loading region relaxes many of the design constraints of the present units. The microwave resonance region can now be designed with no consideration of optical issues. A simple set of cylindrical shields and a solenoid supply a very uniform and stable magnetic environment. In Fig. 2 it is seen that the volume of the resonance region is now reduced by about 100 as the magnetic shields are 10 times smaller in diameter. Since the resonance region requires the most stringent temperature regulation of the entire physics package, the size reduction enables simplification of thermal control of the standard.

Similarly, the state selection optical design is now done with no worry of perturbing the atomic resonance. In practice, this means that the optical components are no longer required to be nonmagnetic as in the previous design. The custom-made copper-titanium flanged UV transmitting windows can be replaced with S-steel flanged windows, which are each about \$2,000 less expensive. The \$4,000 nonmagnetic UV light-collecting horn is replaced by an S-steel elbow connecting the trap region to the vacuum pumps. The light source can be moved much closer to the ion cloud, raising the possibility of using a collimated UV source with no collection mirror, thus further reducing the final package size. Similar modifications could be made in the collection arms to reduce size. These design changes would not allow stable clock operation if the ions underwent microwave clock resonances inside the optical state selection region as in the previous design, because of the large magnetic shifts of the atomic levels induced by the close proximity of the light source and detectors.

Another simplification gained by relaxing the nonmagnetic requirement is in the electron gun assembly and its heater current supply. The electron source for ion creation is now remote from the atomic resonance region, allowing use of the standard (magnetic) filament base and a floatable dc current drive. In the design of Fig. 1, ions are created inside the resonance region and require a floating audio frequency filament heater current to prevent the residual dc magnetic field from the electron gun from shifting the atomic ion frequency.

Additionally, the largest frequency offset and potential instability can be reduced greatly by selecting a somewhat longer trapping length in the atomic resonance region. This frequency shift stems from the finite (i.e., nonzero) ion cloud diameter in which ions spend time in regions of large rf trapping fields where their motion results in frequency pulling via the second-order Doppler or relativistic time dilation effect. The magnitude of this offset depends only on the linear ion density, N/L , and is given by [3]

$$\left(\frac{\Delta f}{f}\right) = -\left(\frac{e^2}{8\pi\epsilon_0 mc^2}\right) \frac{N}{L}$$

An increase in the resonance trapping length to about 200 mm would reduce the clock sensitivity to ion number variations by about a factor of 4 from that in the present system of Fig. 1. Since ion number stabilization to the 0.1-percent level has been demonstrated for 10^{-15} frequency standard operation in that system, an ion number-induced instability noise floor below 2.5×10^{-16} should be achievable in this modified configuration.

IV. Conclusions

An improved architecture for a linear ion trap-based frequency standard is described. Its design is a natural extension of the original LIT and separates the state selection and ion creation region from the atomic resonance region. Many of the design constraints of the present configuration are eliminated and a smaller, cheaper, more stable frequency standard becomes possible.

References

- [1] J. D. Prestage, R. L. Tjoelker, G. J. Dick, and L. Maleki, "Ultra-Stable Hg⁺ Trapped Ion Frequency Standard," *Journal of Modern Optics*, vol. 39, pp. 221–232, 1992.
- [2] J. D. Prestage, R. L. Tjoelker, R. T. Wang, G. J. Dick, and L. Maleki, "Hg⁺ Trapped Ion Standard Performance With the Superconducting Cavity Maser Oscillator as L.O.," *Proc. 1992 IEEE Frequency Control Symposium*, pp. 58–63, 1992.
- [3] J. D. Prestage, G. J. Dick, and L. Maleki, "New Ion Trap for Frequency Standard Applications," *Journal of Applied Physics*, vol. 66, pp. 1013–1017, 1989.

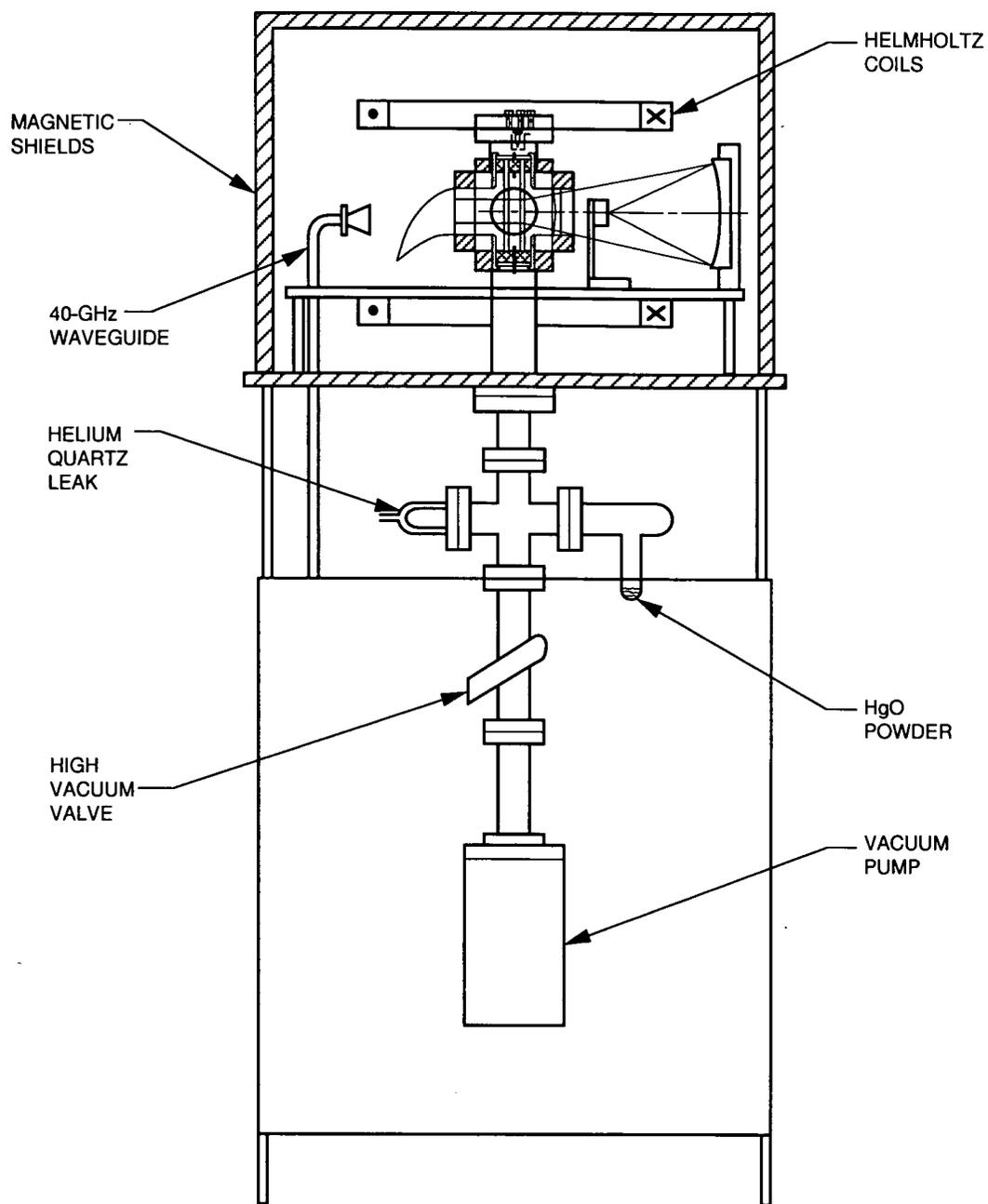


Fig. 1. Present physics unit for the linear ion trap-based frequency standard. Hg ions are state prepared and interrogated with the multiplied output of a local oscillator in the ion trap housed inside the magnetic shielded region. The apparatus is about 2 m high by 1 m on a side.

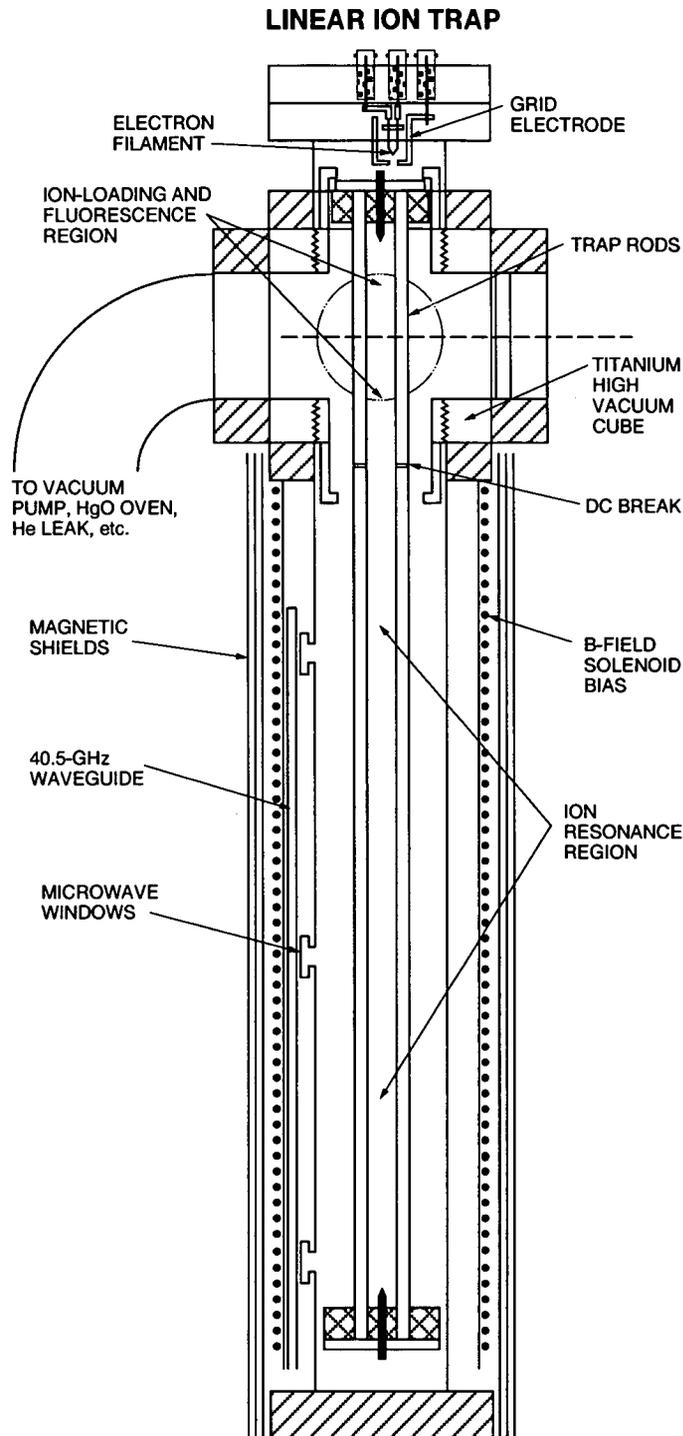


Fig. 2. Improved physics unit for the trapped ion frequency standard. Hg ions are created and state prepared in the upper region of the trap, then electrically moved into the lower region, where the Hg^+ atomic resonance frequency is compared with the local oscillator frequency. The ions are then moved back to the state selection region to check for frequency detuning from the LO during the resonance comparison. Overall dimensions are about 0.5 m by 10 cm.

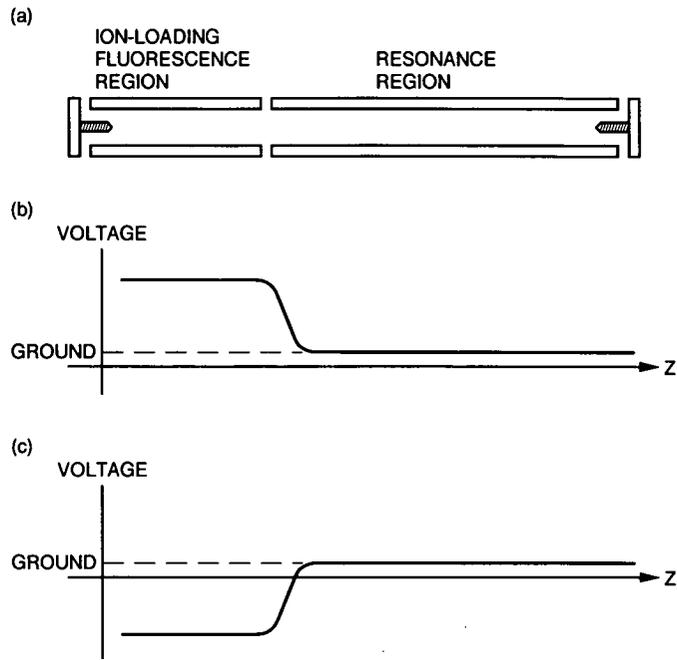


Fig. 3. The improved linear ion trap with separated state selection and microwave resonance regions: (a) the complete ion trap, showing the dc junction; (b) a positive dc bias is applied to the loading and interrogation trap rods forcing the ions into the resonance region; and (c) ions are drawn back into the fluorescence region by applying a negative bias.

S2-04

167922

P. 15

The Effect of Clock, Media, and Station Location Errors On Doppler Measurement Accuracy

J. K. Miller
Navigation Systems Section

Doppler tracking by the DSN is the primary radio metric data type used by navigation to determine the orbit of a spacecraft. The accuracy normally attributed to orbits determined exclusively with Doppler data is about 0.5 microradians in geocentric angle. Recently, the Doppler measurement system has evolved to a high degree of precision primarily because of tracking at X-band frequencies (7.2 to 8.5 GHz). However, the orbit determination system has not been able to fully utilize this improved measurement accuracy because of calibration errors associated with transmission media, the location of tracking stations on the Earth's surface, the orientation of the Earth as an observing platform, and timekeeping. With the introduction of Global Positioning System (GPS) data, it may be possible to remove a significant error associated with the troposphere. In this article, the effect of various calibration errors associated with transmission media, Earth platform parameters, and clocks are examined. With the introduction of GPS calibrations, it is predicted that a Doppler tracking accuracy of 0.05 microradians is achievable.

I. Introduction

The Doppler data type provides a measure of the line-of-sight range rate between a tracking station and a spacecraft. This functional definition is useful for analyzing orbit determination errors that are spacecraft or trajectory dependent, but is of little use for analyzing error sources close to the actual measurement such as media or instrumentation errors. The actual measurement is a count of a signal derived from the signal received from the spacecraft and a frequency standard maintained at the tracking station that is related to the transmitted signal. Thus, a precision model of the Doppler observable would include

a model of the signal path and hardware elements as well as spacecraft dynamics. In practice, the instrumentation errors are small compared to media, station location, and spacecraft dynamics errors and, therefore, a simple functional model of the hardware should suffice.

A model of the Doppler observable is developed below that idealizes some of the instrumentation error sources yet precisely models the external environment. This model is sufficiently precise for computation of the observable and is essentially the model contained in JPL's Orbit Determination Program (ODP) [1]. Of particular interest are models

that are external to the DSN tracking station hardware, yet pertain directly to the signal path. Media effects and the effect of general relativity on the station clocks are examples. Other models, such as station locations and polar motion, though not directly part of the Doppler measurement system, are treated as measurement calibrations for the purposes of analysis and orbit determination.

II. Doppler Measurement Model

The Doppler measurement is simply an electronic count of the number of cycles from a frequency standard (N_c) minus the number of cycles of the spacecraft signal received by the ground station (N_r), scaled by the count time interval (ΔT_c). Thus

$$Z_m = \frac{(N_c - N_r) + n}{\Delta T_c} \quad (1)$$

where n is the measurement noise, which is typically about 1/10 of a cycle. The received frequency and standard frequency need not be counted individually and differenced, but are added together electronically and the beat frequency counted. This is a detail that is dependent on the hardware implementation. The numerical value of Z_m is the number that is recorded on the tracking data file and is sent to the ODP for orbit determination; the units of Z_m are hertz.

In the ODP, one needs to obtain a computed value for Z_m as a function of parameters that are available to the ODP. This function can be derived from the equations of motion and the physical model of the system or can be worked backwards from Eq. (1) for the measurement. Working backwards from the measurement equation is chosen because the starting point is an equation that represents the real physical measurement.

The frequency standard is obtained by scaling the reference oscillator frequency (f_q) to equal the transmitted frequency times the spacecraft turnaround ratio which would nominally be the received frequency if there were no spacecraft Doppler shift or additional delay.

$$N_c = C_3 f_t \Delta T_c \quad (2)$$

where for S-band (2.1 to 2.3 GHz) Doppler,

$$C_3 = \frac{240}{221}$$

$$f_t = 96 f_q$$

$$\Delta T_c = T_{3e} - T_{3s}$$

The count time (ΔT_c) is defined as the difference between the reception time at the start of the count-time interval (T_{3s}) and the reception time at the end of the interval (T_{3e}). For a schematic representation of these times, see Fig. 1. In Eq. (2), all the parameters are constant or arbitrarily specified, including the reception times. The real information content of the measurement is contained within the count N_r . Thus, to obtain a complete equation for the computed measurement, one needs an equation for N_r . It is tempting to differentiate and work in the frequency domain; however, the actual observable is accumulated phase change, and this is a sufficient reason to keep this representation. Therefore, it is necessary to formulate the data type in terms of phase thus bypassing an explicit equation for the received frequency. The equation that relates the measurement to the observable parameters is simply

$$N_r = C_3 N_t \quad (3)$$

where

$$N_t = f_t (T_{1e} - T_{1s})$$

Equation (3) for N_r states that the number of cycles counted at the receiver is equal to the number of cycles transmitted (N_t) times the spacecraft turnaround ratio. This equation is true because the transmitted cycles are effectively the same as the received cycles. Thus, the information content of the measurement is now contained in the transmit times T_{1e} and T_{1s} . Since both of these times are unknown, some additional equations are needed to tie into the observable quantities. At this point in the development, the following equation for the computed measurement is introduced:

$$Z_c = (T_{3e} - T_{3s} - T_{1e} + T_{1s}) \frac{C_3 f_t}{\Delta T_c} \quad (4)$$

Equations are needed for the times in Eq. (4) and these will be developed as functions of ephemeris time t . The equation for the atomic clock at the station is

$$T = t + F(t, x, y) \quad (5)$$

where x refers to state variables and y refers to various constant parameters.

The station time (T) is equal to the ephemeris time t modified by a small correction because of relativity and any other parameter that may affect the running of the clock. The calibration function (F) is a function of time, the solar system gravitational potential, and other constant parameters. The relevant times relating to the Doppler measurement are

$$T_{1s} = t_{1s} + F(t_{1s}, x, y)$$

$$T_{1e} = t_{1e} + F(t_{1e}, x, y)$$

$$T_{3s} = t_{3s} + F(t_{3s}, x, y)$$

$$T_{3e} = t_{3e} + F(t_{3e}, x, y)$$

Making the above substitutions, the equation for the computed observable becomes

$$\begin{aligned} Z_c = & (t_{3e} - t_{3s} - t_{1e} + t_{1s}) \frac{C_3 f_t}{\Delta T_c} \\ & + [F(t_{3e}, x, y) - F(t_{3s}, x, y) - F(t_{1e}, x, y) \\ & + F(t_{1s}, x, y)] \frac{C_3 f_t}{\Delta T_c} \end{aligned} \quad (6)$$

Since the speed of light is constant in any reference frame, one obtains, by integrating along the light path,

$$t_{3e} - t_{1e} = \frac{\rho_{12e} + \rho_{23e}}{c} + \Delta t_{1e}^m + \Delta t_{3e}^m \quad (7)$$

$$t_{3s} - t_{1s} = \frac{\rho_{12s} + \rho_{23s}}{c} + \Delta t_{1s}^m + \Delta t_{3s}^m \quad (8)$$

where the ρ terms represent the integrated distance along the light path and the t^m terms represent the additional delay caused by transmission media. The distances along the light path,

$$\rho_{12s} = \int \int_{t_{1s}}^{t_{2s}} \ddot{\rho} dt dt \quad (9)$$

$$\rho_{23s} = \int \int_{t_{2s}}^{t_{3s}} \ddot{\rho} dt dt \quad (10)$$

$$\rho_{12e} = \int \int_{t_{1e}}^{t_{2e}} \ddot{\rho} dt dt \quad (11)$$

$$\rho_{12e} = \int \int_{t_{2e}}^{t_{3e}} \ddot{\rho} dt dt \quad (12)$$

are obtained by integrating the equations of motion for the spacecraft. These equations are referred to as the light time equations and are solved iteratively for the arguments of integration. The media delay is included in the measurement equation by evaluating the calibration function (G) at the appropriate times:

$$\Delta t^m = G(t, x, y) \quad (13)$$

and

$$\Delta t_{1s}^m = G(t_{1s}, x, y)$$

$$\Delta t_{1e}^m = G(t_{1e}, x, y)$$

$$\Delta t_{3s}^m = G(t_{3s}, x, y)$$

$$\Delta t_{3e}^m = G(t_{3e}, x, y)$$

The final equation [Eq. (14)] for Z_c (the computed measurement) includes the observable equations as well as clock and media calibration functions:

$$\begin{aligned} Z_c = & \frac{\rho_{12e} + \rho_{23e} - \rho_{12s} - \rho_{23s}}{c} \frac{C_3 f_t}{\Delta T_c} \\ & + [F(t_{3e}, x, y) - F(t_{3s}, x, y) - F(t_{1e}, x, y) \\ & + F(t_{1s}, x, y)] \frac{C_3 f_t}{\Delta T_c} + [G(t_{3e}, x, y) - G(t_{3s}, x, y) \\ & + G(t_{1e}, x, y) - G(t_{1s}, x, y)] \frac{C_3 f_t}{\Delta T_c} \end{aligned} \quad (14)$$

III. Analytic Doppler Error Model

Equation (14) for the computed measurement, while useful for actual orbit determination operations, is somewhat cumbersome for error analysis. An error analysis could be performed by combining the sensitivities of each term in Eq. (14). An alternative method would be to develop an approximate analytic representation of the computed measurement which will be pursued below. The range from the tracking station to the spacecraft may be described by the following function obtained by integrating the equations of motion:

$$\rho = R(t, x, y) \quad (15)$$

The velocity of the spacecraft projected along the Earth line of sight is obtained by differentiating the range and is given by

$$\dot{\rho} = \frac{\partial R}{\partial t} \quad (16)$$

which may be approximated by

$$\dot{\rho} \approx \frac{R(t_e, x, y) - R(t_s, x, y)}{\Delta T_c} \quad (17)$$

From the geometry shown on Fig. 1, then

$$R(t_e, x, y) - R(t_s, x, y) = \frac{1}{2} (\rho_{12e} + \rho_{23e} - \rho_{12s} - \rho_{23s}) \quad (18)$$

Substituting Eq. (18) into the equation for the measurement, the following expression is obtained for the geometric part of the computed measurement:

$$Z_c \approx \frac{2C_3 f_t}{c} \frac{\partial R}{\partial t} \quad (19)$$

Consider the following generic function for the clock and media calibration functions:

$$Z = F(t) \quad (20)$$

For small Δt , the second derivative may be approximated by

$$\frac{\partial^2 Z}{\partial t^2} = \frac{F(t + 2\Delta t) - 2F(t + \Delta t) + F(t)}{\Delta t^2} \quad (21)$$

In the limit as Δt approaches zero, this is the definition of the second derivative. Without loss of generality, one can evaluate the second derivative using two separate Δt 's (one for the first derivative and the other for the second derivative)

$$\frac{\partial^2 Z}{\partial t^2} = \frac{F(t + \Delta t_2 + \Delta t_1) - F(t + \Delta t_2)}{\Delta t_1 \Delta t_2} - \frac{F(t + \Delta t_1) - F(t)}{\Delta t_1 \Delta t_2} \quad (22)$$

Solving for the values of the function at the indicated times and rearranging the terms gives

$$F(t + \Delta t_1 + \Delta t_2) - F(t + \Delta t_2) - F(t + \Delta t_1) + F(t) = \frac{\partial^2 Z}{\partial t^2} \Delta t_1 \Delta t_2 \quad (23)$$

If one makes Δt_1 equal to the count time (ΔT_c) and Δt_2 equal to the round-trip light time (Δt_{rtl}), the above calibrations can be approximated by the second derivative of the calibration function, provided this function has frequency components with periods that are long compared to the delta times. Thus, for the clock calibration function evaluations,

$$F(t_{3e}) - F(t_{3s}) - F(t_{1e}) + F(t_{1s}) \approx \frac{\partial^2 F(t)}{\partial t^2} \Delta T_c \Delta t_{rtl} \quad (24)$$

For the media calibration function, one may perform a similar approximation. However, because of the reversal of signs in Eq. (14), the dominant term is given by the first derivative and then

$$G(t_{3e}) - G(t_{3s}) + G(t_{1e}) - G(t_{1s}) \approx 2 \frac{\partial G(t_{1e})}{\partial t} \Delta T_c \quad (25)$$

Substituting into the exact equation for the measurement gives the following approximation:

$$Z_m \approx C_3 f_t \left[\frac{2\dot{\rho}}{c} + \frac{\partial^2 F(t)}{\partial t^2} \Delta t_{rtl} + 2 \frac{\partial G(t_{1e})}{\partial t} \right] \quad (26)$$

However, if the round-trip light time is too long relative to the count time, as would be the case for the outer planets, then

$$Z_m \approx C_3 f_t \left[\frac{2\dot{\rho}}{c} + \frac{\partial F(t_3, x, y)}{\partial t} - \frac{\partial F(t_1, x, y)}{\partial t} + \frac{\partial G(t_3, x, y)}{\partial t} + \frac{\partial G(t_1, x, y)}{\partial t} \right] \quad (27)$$

The evaluation of the partial derivatives is greatly facilitated by breaking up the calibration functions into individual terms and approximating with simple functions (preferably sines, cosines, or exponentials). The sensitivity of the measurement to various y parameters that are used to describe the calibration functions is given by

$$\Delta Z_m \approx C_3 f_t \left[2 \frac{\partial^2 R(t, x, y)}{\partial y \partial t} + 2 \frac{\partial^2 G(t, x, y)}{\partial y \partial t} + \frac{\partial^3 F(t, x, y)}{\partial y \partial t^2} \Delta t_{rtl} \right] \Delta y \quad (28)$$

The y parameters in Eq. (28) can be formed into a covariance matrix which, when pre- and post-multiplied by the measurement sensitivities, gives the measurement variance.

For example, consider the case where the only y parameter is the amplitude of the daily special relativity term that is part of the atomic clock calibration function. This term is defined by

$$F_e = \alpha_e r_s \sin(\omega_e t + \lambda_s + \phi) \quad (29)$$

where α_e is the subject y parameter, r_s is the station spin radius, ω_e is the Earth rotation rate, λ_s is the tracking station longitude, and ϕ is a phase angle needed to bring UT1 into agreement with ephemeris time. Typical nominal values of these parameters are given by

$$\alpha_e = 3.17679 \times 10^{-10} \text{ sec/km}$$

$$r_s = 5,204 \text{ km}$$

$$\omega_e = 7.292 \times 10^{-5} \text{ rad/sec}$$

Taking the required first and second partial derivatives, one obtains the following equation for the measurement error as a function of time:

$$\sigma_z \approx |C_3 f_t \Delta t_{rtl} r_s \omega_e^2 \sin(\omega_e t + \lambda_s + \phi)| \sigma_{\alpha_e} \quad (30)$$

where, at X-band and for a typical round-trip light time,

$$T_{rtl} = 1,512 \text{ sec}$$

$$C_3 = \frac{880}{749}$$

$$f_t = 32 f_q + 0.65 \times 10^{10} \text{ Hz}$$

$$f_q = 20.98 \times 10^6 \text{ Hz}$$

$$\sigma_{\alpha_e} = 3.17679 \times 10^{-10} \text{ sec/km}$$

This equation gives a peak amplitude of 70 mHz for a 100-percent error in the amplitude (α_e). Thus, to obtain a 1-mHz accuracy in the measurement, the calibration error must be less than 2 percent.

IV. Data Noise Evaluation

Approximately two orbits of Magellan data were processed from February 7, 1991, 19:29:02 (ET) to February 8, 1991, 00:10:33 (ET). The postfit residuals are shown on Fig. 2, which involved solving for the spacecraft state, a constant nongravitational acceleration, and a fourth-degree and -order spherical harmonic gravity field. The first 10,000 sec of data shown on Fig. 2 are from Deep Space Station (DSS) 15. The gap around 6,000 sec is data unavailable near periapsis. Data after 12,000 sec are from DSS 45.

Recall from Eq. (1) that the measurement noise is

$$Z_m = \frac{N_c - N_r}{\Delta T_c} + \frac{n}{\Delta T_c} \quad (31)$$

The data noise is approximately 1/10 of a cycle count, and is independent of frequency. For a 60-sec count time, the Doppler data noise is therefore about 1.66 mHz. This value agrees with Fig. 2, which shows noise of this magnitude from DSS 15 and appears to be uniformly distributed but with some structure. The data from DSS 45 are a bit more noisy and appear more Gaussian.

Doppler data are scaled by the count time to make the recorded measurement proportional to range rate. The

Doppler measurement sensitivity to line-of-sight velocity is given by Eq. (19):

$$Z_{\dot{\rho}} \approx \frac{2C_3 f_t}{c} \dot{\rho} \quad (32)$$

At S-band, typical values for the constants in the above equation are:

$$C_3 = \frac{240}{221}$$

$$f_t = 96 f_q$$

$$f_q = 22 \times 10^6 \text{ Hz}$$

$$c = 299792.458 \text{ km/sec}$$

and solving Eq. (32) for $\dot{\rho}$ gives 0.108 mm/sec for $Z_{\dot{\rho}} = 1.66$ mHz data noise. One can also obtain from Eq. (32) the well-known result that for $\dot{\rho} = 1$ mm/sec the measurement noise is $Z_{\dot{\rho}} = 15.3$ mHz. At X-band, the measurement noise $Z_{\dot{\rho}} = 1.66$ mHz corresponds to 0.03 mm/sec velocity.

V. Troposphere Calibration

A radio signal passing through the Earth's troposphere will be delayed depending on the dielectric constant of the media and path length. A functional formula for this delay has been defined above.

$$\Delta t' = G_t(t, x, y) \quad (33)$$

The sensitivity of the Doppler measurement to media is functionally given by Eq. (28) and one may extract the part that pertains to the troposphere.

$$\Delta Z_m \approx 2C_3 f_t \frac{\partial^2 G_t(t, x, y)}{\partial y \partial t} \Delta y \quad (34)$$

The troposphere delay has been separated into wet and dry components that are functions of delay at zenith (z) and elevation angle (γ):

$$G_t(t, x, y) = R_d + R_w \quad (35)$$

$$R_d = f_d(z_d, \gamma) \quad (36)$$

$$z_d = f_{z_d}(t, y) \quad (37)$$

$$R_w = f_w(z_w, \gamma) \quad (38)$$

$$z_w = f_{z_w}(t, y) \quad (39)$$

$$\gamma = f_\gamma(t) \quad (40)$$

In taking the required partial derivatives with respect to time, one assumes R_w and R_d are linear in z_w and z_d and γ is a function of only time:

$$\frac{\partial G_t}{\partial t} = \frac{\partial R_d}{\partial \gamma} \dot{\gamma} + \frac{\partial R_d}{\partial z_d} \frac{\partial z_d}{\partial t} + \frac{\partial R_w}{\partial \gamma} \dot{\gamma} + \frac{\partial R_w}{\partial z_w} \frac{\partial z_w}{\partial t} \quad (41)$$

The expression for the measurement sensitivity becomes

$$\Delta Z_m \approx 2C_3 f_t \left[\frac{\partial^2 R_d}{\partial y \partial \gamma} \dot{\gamma} + \frac{\partial R_d}{\partial z_d} \frac{\partial^2 z_d}{\partial y \partial t} + \frac{\partial^2 R_w}{\partial y \partial \gamma} \dot{\gamma} + \frac{\partial R_w}{\partial z_w} \frac{\partial^2 z_w}{\partial y \partial t} \right] \Delta y \quad (42)$$

The first term in the above equation represents the non-linearity of the dry troposphere mapping function, and the second term represents the variation in the dry troposphere z height due to local weather. The next two terms are the same quantities for the wet troposphere. The troposphere wet and dry mapping functions are tabulated as delay as a function of spacecraft elevation angle. Empirical formulas for these mapping functions are given by [2,3]

$$R_d = \frac{1}{C} \frac{z_d}{\sin \gamma + \frac{A_d}{B_d + \tan \gamma}} \quad (43)$$

$$R_w = \frac{1}{C} \frac{z_w}{\sin \gamma + \frac{A_w}{B_w + \tan \gamma}} \quad (44)$$

where

$$\sin \gamma = \cos \delta \cos \lambda \cos \phi + \sin \lambda \cos \phi + \sin \phi \sin \delta \quad (45)$$

$$\lambda = \omega_e t + \lambda_s - \alpha \quad (46)$$

The dry component of the troposphere (R_d) is a function of the delay at zenith (z_d), the elevation angle (γ), and constants A_d and B_d that are provided to model the bending at low-elevation angles. The wet component (R_w) is similarly defined. The elevation angle (γ) is computed as a function of the latitude of the tracking station (ϕ), the declination of the spacecraft (δ), and the local hour angle with respect to the spacecraft (λ). The local hour angle is zero when the spacecraft is at zenith and is a function of the Greenwich hour angle ($\omega_e t$), the station longitude (λ_s), and the right ascension of the spacecraft (α).

The dry component of the troposphere is assumed to be stable and most of the variability is associated with the wet component. The variation in the wet component may be modeled as a periodic variation in the z height (z_w). The actual variation obtained by tracking the GPS satellites in early 1991 is shown on Fig. 3. The hourly variation in the wet component of the troposphere appears as a random walk that would require a high-order Fourier series to represent analytically. For the purpose of error analysis, the variation may be modeled as a simple sinusoid with amplitude and frequency selected to be representative of Fig. 3.

$$z_w = z_{w_0} + z_{w_1} \sin(\omega_{w_1} t) \quad (47)$$

For the troposphere, the y parameters that are of interest are the constant coefficient of the wet z height function (z_{w_0}) and the periodic coefficient (z_{w_1}). In order to obtain these sensitivities, it is necessary to evaluate the partial derivatives contained in the above measurement sensitivity function [Eq. (59)] that pertain to the wet troposphere component. For elevation angles above the horizon by a few degrees, one can approximate the mapping function by

$$R_w \approx \frac{1}{c} \frac{z_w}{\sin \gamma} \quad (48)$$

and

$$\frac{\partial R_w}{\partial \gamma} = \frac{1}{c} \frac{-z_w \cos \gamma}{\sin^2 \gamma} \quad (49)$$

$$\frac{\partial R_w}{\partial z_w} = \frac{1}{c} \frac{1}{\sin \gamma} \quad (50)$$

The partial derivatives of γ and z_w with respect to time are given by

$$\dot{\gamma} = \frac{\cos \phi (\cos \lambda - \cos \delta \sin \lambda)}{\cos \gamma} \dot{\lambda} \quad (51)$$

$$\dot{\lambda} = \omega_e \quad (52)$$

$$\dot{z}_w = z_{w_1} \omega_{w_1} \cos(\omega_{w_1} t) \quad (53)$$

One is interested in the sensitivity of the Doppler measurement to the constant and periodic terms of the expression for the wet troposphere z height:

$$\Delta Z_m \approx \frac{2C_3 f_t}{c} \frac{\partial^2 R_w}{\partial y \partial \gamma} \dot{\gamma} \Delta y \quad (54)$$

Substituting the above partial derivatives into this equation yields

$$\Delta Z_{z_0} \approx \frac{2C_3 f_t}{c} \frac{\cos \gamma \dot{\gamma}}{\sin^2 \gamma} \Delta z_{w_0} \quad (55)$$

and so

$$\begin{aligned} \Delta Z_{z_1} &\approx \frac{2C_3 f_t}{c} \\ &\times \left\{ \frac{1}{\sin \gamma} \omega_{w_1} \cos \omega_{w_1} t - \frac{\cos \gamma \dot{\gamma}}{\sin^2 \gamma} \sin \omega_{w_1} t \right\} \Delta z_{w_1} \end{aligned} \quad (56)$$

An example of the effect of the troposphere on X-band Doppler is the application of GPS troposphere calibration data to the Magellan spacecraft. The GPS data were obtained at a time when the spacecraft was in orbit about Venus. It would be preferable to have cruise data, where the effect of the troposphere is more pronounced, rather than orbiting data where separation of the tropospheric effect from the gravity field modelling effects is difficult. Recall the two orbits of Magellan Doppler residuals shown on Fig. 2. At the time of the first data point, the parameters in the above Doppler sensitivity equations had the following values:

$$C_3 = \frac{880}{749}$$

$$\gamma = 37.49 \text{ deg}$$

$$f_t = 32f_q + 0.65 \times 10^{10} \text{ Hz}$$

$$\dot{\gamma} = 3.818 \times 10^{-5} \text{ rad/sec}$$

$$f_q = 20.98 \times 10^6 \text{ Hz}$$

$$z_{w_0} = 5 \times 10^{-5} \text{ km}$$

$$z_{w_1} = 5 \times 10^{-5} \text{ km}$$

$$\omega_{w_1} = 1.454 \times 10^{-4} \text{ rad/sec}$$

The above analytic Doppler sensitivity equations reveal a Doppler measurement error sensitivity of 0.044 mHz per centimeter of constant wet troposphere z height (z_{w_0}) error and a periodic Doppler measurement sensitivity of 0.044 mHz in quadrature with 0.134 mHz per centimeter of variable wet troposphere z height amplitude (z_{w_1}). Figure 4 shows the prefit residuals for the periodic wet troposphere z height calibration defined above. A Doppler shift of about 0.67 mHz is apparent early on for the data from DSS 15; at the end of Fig. 4, the Doppler shift is about 2 mHz for the low elevation data from DSS 45. The data in Fig. 4 indicate that the error contribution of the troposphere, which may amount to several mHz, could be reduced or eliminated with GPS calibration data.

VI. Ionosphere Calibration

A radio signal passing through the ionosphere experiences a reduction in group velocity and an equal increase in phase velocity that is a function of the frequency and the number of charged particles along the signal path. The Doppler measurement is dependent on the phase velocity and the advance of the signal is functionally defined by

$$\Delta t^i = G_i(t, x, y) \quad (57)$$

The sensitivity of the Doppler measurement to media is functionally given by Eq. (28) and may be adapted to the ionosphere:

$$\Delta Z_m \approx 2C_3 f_t \frac{\partial^2 G_i(t, x, y)}{\partial y \partial t} \Delta y \quad (58)$$

An empirical formula for the effect of the ionosphere on the Doppler measurement is given by [4]

$$G_i = \frac{-1}{c} \sum_{j=0}^n k C_j X^j \quad (59)$$

$$X = 2 \left(\frac{t - t_a}{t_b - t_a} \right) - 1 \quad (60)$$

where the C_j 's are coefficients of a polynomial in time (t) from t_a to t_b normalized over the interval of -1 to $+1$ and k is a proportionality factor introduced for the purpose of computing error sensitivity. The first partial derivative of G_i with respect to t is needed:

$$\frac{\partial G_i}{\partial t} = \frac{2k}{t_b - t_a} \sum_{j=1}^n j C_j X^{j-1} \quad (61)$$

The sensitivity of the Doppler measurement to the constant of proportionality k is given by

$$\Delta Z_k \approx \frac{4C_3 f_t}{c} \frac{1}{t_b - t_a} \sum_{j=1}^n j C_j X^{j-1} \Delta k \quad (62)$$

For the first data point shown on Fig. 2, the following values for the constant parameters are needed to compute the Doppler sensitivity in addition to some of the parameters given above for the troposphere:

$$t = -280,772,197 \text{ sec}$$

$$t_a = -280,785,442 \text{ sec}$$

$$t_b = -280,745,040 \text{ sec}$$

$$C_j = 5.8432, -1.2655, 5.8368, -1.0912,$$

$$-0.6015, 0.5551 \text{ m}$$

Assuming a 10-percent error in the ionosphere model ($\Delta k = 0.1$), the error in the Doppler measurement attributable to the ionosphere is 0.143 mHz.

VII. Station Location Errors

The Doppler measurement of the distance from a tracking station antenna to a spacecraft enables the orbit determination system to estimate spacecraft state, station locations, planetary ephemerides, and many other parameters that are of interest for navigation. If the primary purpose of the estimation process is determining the orbit of a spacecraft, then the station locations and the orientation of the Earth's crust in inertial space are often regarded as part of the measurement system. Station locations, timing and polar motion, continental drift, and solid Earth tides, to name a few quantities, are calibrated out of the data.

Regarding the station locations as part of the measurement system, the following function may be used for error analysis and approximates the effect of the tracking station when the spacecraft is a great distance from the Earth:

$$R_s = \frac{1}{c} r_s \cos \delta \sin(\omega_e t + \lambda_s) \quad (63)$$

This function can be regarded as a calibration of the data and the Doppler measurement sensitivity is functionally the same as for the troposphere and ionosphere. The range is premultiplied by one over c to give this function the units of time consistent with the other calibration functions.

$$\Delta Z_m \approx 2C_3 f_t \frac{\partial^2 R_s(t, x, y)}{\partial y \partial t} \Delta y \quad (64)$$

The sensitivity of the Doppler measurement to the distance of the tracking station from the Earth's spin axis (r_s) is:

$$\Delta Z_{r_s} \approx \frac{2C_3 f_t}{c} \omega_e \cos \delta \cos(\omega_e t + \lambda_s) \Delta r_s \quad (65)$$

For the first data point shown on Fig. 2, the following values for the constant parameters are needed to compute the Doppler sensitivity in addition to the parameters given above for the troposphere and ionosphere:

$$r_s = 5,204 \text{ km}$$

$$\omega_e = 7.292 \times 10^{-5} \text{ rad/sec}$$

$$\delta = 8.39 \text{ deg}$$

At present, the DSN station locations have been determined in the terrestrial (Earth-fixed) frame with an accuracy of about 0.1 m. Assuming a 0.1-m error in the tracking station spin radius (r_s), the error in the Doppler measurement attributable to station locations is about 0.4 mHz.

VIII. Clock Calibration

According to the theory of relativity, a clock running in a frame of reference that is moving with respect to an observer's frame of reference or in a gravitational potential field will appear to run slower by an observer that is assumed stationary and is removed from the potential field. Therefore, a distant observer who is stationary with respect to the solar system will see the atomic clocks at the DSN stations running slower than his or her hypothetical clock. The observer's clock records coordinate time, which is called post-Newtonian time (PNT) [5,6].

The relationship between PNT and the proper time measured by an atomic clock is given by the metric. For a particle moving in an orbit around the Sun, the metric in isotopic Schwarzschild coordinates is given by

$$ds^2 = \left(1 - \frac{2U}{c^2}\right) c^2 dt^2 - v^2 dt^2 + O(c^{-2})$$

where

$$v^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2 \quad (66)$$

Solving for proper time ($ds^2 = c^2 d\tau^2$) yields

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{2U}{c^2} - \left(\frac{v}{c}\right)^2} \quad (67)$$

which can be further approximated by

$$\frac{d\tau}{dt} = 1 - \frac{\mu_s}{c^2 r} - \frac{1}{2} \frac{v^2}{c^2} - \frac{\mu_e}{c^2 r_e} \quad (68)$$

where the Earth's gravitational potential is separated from the Sun. The atomic clock time (τ) is obtained as a function of t by integrating the metric in conjunction with the equations of motion:

$$\tau = \int_{t_0}^t (1 - L) dt \quad (69)$$

where

$$L = \frac{\mu_s}{c^2 r} + \frac{1}{2} \frac{v^2}{c^2} + \frac{\mu_e}{c^2 r_e} \quad (70)$$

The function L can be separated into a constant term (L_0), secular terms that grow with time (L_s), and periodic terms (L_p). Thus

$$L = L_0 + L_s + L_p \quad (71)$$

The constant term (L_0) is obtained by averaging L over all time and can be represented by

$$L_0 = \frac{1}{c^2} \left(\frac{\mu_s}{r_0} + \frac{1}{2} v_0^2 \right) + \frac{\mu_e}{c^2 r_e} \quad (72)$$

where r_0 and v_0 are constants that give the correct value of L_0 in the above equation. For the Earth's orbit about the Sun, r_0 is approximately the semimajor axis of the orbit and v_0 is approximately the mean orbital velocity. Since the orbit is nearly an ellipse,

$$\frac{\mu_s}{a} = \frac{2\mu_s}{r} - v^2 \quad (73)$$

and for $r = a$,

$$L_0 \approx \frac{3\mu_s}{2c^2 a} + \frac{\mu_e}{c^2 r_e} \quad (74)$$

The secular terms L_s are assumed to be zero because of conservation of energy and momentum. This leaves the periodic terms and these are given by

$$L_p = \frac{1}{c^2} \left(\frac{\mu_s}{r} - \frac{\mu_s}{r_0} + \frac{1}{2} v^2 - \frac{1}{2} v_0^2 \right) \quad (75)$$

Recall that the station clock calibration function is defined by

$$\tau = t + F(t, x, y) \quad (76)$$

which can be evaluated directly by numerical integration, giving

$$\tau = t + \int_{t_0}^t -L_0 - \frac{1}{c^2} \left(\frac{\mu_s}{r} - \frac{\mu_s}{r_0} + \frac{1}{2} v^2 - \frac{1}{2} v_0^2 \right) dt \quad (77)$$

This equation will be referred to as F_i .

Moyer [1] provides an approximate analytic formula for the periodic terms, attributing much of the original work to Brooks Thomas:

$$\begin{aligned} \tau \approx t - L_0 (t - t_0) &- \frac{2}{c^2} (\dot{\mathbf{r}}_b^s \cdot \mathbf{r}_b^s) - \frac{1}{c^2} (\dot{\mathbf{r}}_b^e \cdot \mathbf{r}_e^b) \\ &- \frac{1}{c^2} (\dot{\mathbf{r}}_e^c \cdot \mathbf{r}_e^e) - \frac{1}{c^2} (\dot{\mathbf{r}}_s^c \cdot \mathbf{r}_b^s) - \frac{\mu_j}{c^2 (\mu_j + \mu_s)} (\dot{\mathbf{r}}_j^s \cdot \mathbf{r}_j^s) \\ &- \frac{\mu_{sa}}{c^2 (\mu_{sa} + \mu_s)} (\dot{\mathbf{r}}_{sa}^s \cdot \mathbf{r}_{sa}^s) \end{aligned} \quad (78)$$

In the notation used above, the position of the body identified by the subscript is with respect to the body identified by the superscript, where c is the solar system barycenter, s is the Sun, b is the Earth-Moon barycenter, e is the Earth, j is Jupiter, and sa is Saturn. This equation will be referred to as the vector function and is designated by F_v .

Another formula is provided in part 2 of [1] that describes the periodic terms as functions of sines and cosines. The dominant terms are given by

$$\begin{aligned} \tau \approx t - L_0 (t - t_0) &- 1.658 \times 10^{-3} \sin E \\ &- 1.548 \times 10^{-6} \sin D \\ &- 3.17679 \times 10^{-10} r_s \sin(UT1 + \lambda) - \dots \end{aligned} \quad (79)$$

where E , D , and $UT1$ are the angles describing the Earth's orbit about the Sun, the Moon's orbit about the Earth, and the rotation of the Earth about its spin axis, respectively, and λ is the ground longitude. If one replaces the sines and cosines in the complete expression for τ by a power series and truncates, there is a polynomial in t . The above equation will be referred to as the polynomial function and is designated by F_p .

Since there are three functions that purportedly give the same result, it may be instructive to tabulate the differences as a function of time. One can then apply the error sensitivity given by Eq. (28) to determine the error in the Doppler measurement. This procedure will yield the error in the differences between the functions but will not reveal which function is correct. Indeed, the correct function may be some unknown fourth function.

As a basis for this comparison, the function (F_i) obtained by numerical integration is selected. Figure 5 shows a plot of UTC minus ephemeris time (ET) as a function of ET. The ET is defined as PNT time with the constant rate term L_0 removed. A detailed discussion of this subject is provided by Hellings [6]. Dropping the L_0 term from the definition of ET results in the length of the ET second being shortened. Therefore, in order to compensate, the masses of all the bodies in the solar system are scaled. UTC is simply an atomic clock time τ adjusted by a constant that includes accumulated leap seconds. The 1991 value of this constant is -58.184 sec.

In order to perform the integration to obtain F_i , a value for L_0 and the constant of integration must be obtained. An initial guess for L_0 is obtained from Eq. (41) and the constant term is obtained from the vector function F_v [Eq. (46)]. The equations of motion are integrated for 1 year and a mean value of L_0 is obtained which constrains the function F_i to equal the vector function F_v at the end points of the interval. The linear term L_0 obtained in this manner has a value of 1.55035×10^{-8} , which compares quite favorably with the linear term, given in [2], of 1.55052×10^{-8} . The difference may be attributed to the averaging interval or the effect of Jupiter and Saturn on the solar system barycenter. Next, the integrated vector and polynomial clock functions over the same 1-year time interval are evaluated, and the differences $F_v - F_i$ and $F_p - F_i$ are plotted on Fig. 6. Inspection of Fig. 6 reveals an annual error term with an amplitude of about 1.5×10^{-6} sec in both the vector and polynomial functions, a biweekly term of about 1.0×10^{-7} sec in only the polynomial function, and a daily term that is less than 1.0×10^{-12} sec.

The error functions shown on Fig. 6 are quite small compared to the actual clock calibration function shown on Fig. 5. As previously noted, the error functions arise from differences in the methods for computing the function and as a result provide very little insight into the error source. The error source may be high-order terms that have been neglected, approximations that have been made, or simple programming errors in implementation of the comparisons. However, the biweekly oscillation in the

polynomial function is highly suspect because this term has been omitted from the expansion. Without judging the quality of the approximations, one proceeds to compute the effect of these error terms on the Doppler measurement.

An error function can be defined that includes all of the suspect terms and this is given by

$$F_e = \tau_e + L_e(t - t_0) + A_d \sin(\omega_d t) + A_{bw} \sin(2\omega_{bw} t) + A_y \sin(\omega_y t) \quad (80)$$

where A_d , A_{bw} , and A_y are the amplitudes of the daily, biweekly, and annual terms identified above, respectively. The clock offset error is τ_e and the clock rate error is L_e . Recall from Eq. (27) that the sensitivity of the Doppler measurement to errors in the amplitude (A) is given by

$$\Delta Z_m \approx C_3 f_t \frac{\partial^3 F_e(t, x, y)}{\partial A \partial t^2} \Delta t_{rtt} \Delta A$$

and the resulting error is given by

$$\sigma_z \approx |C_3 f_t \Delta t_{rtt} \omega^2 \sin(\omega t)| \sigma_A \quad (81)$$

Observe that the constant and linear terms do not contribute to the measurement error. Numerical values for daily, biweekly, and annual terms are given in Table 1. The Doppler measurement errors associated with the daily and annual terms are negligible. The biweekly error of 0.0345 mHz, though small for Doppler measurements, may present a problem for VLBI measurements.

IX. Summary

A survey of the X-band Doppler measurement system has supported the generally accepted notion that troposphere calibrations are the dominant error source. Using only a seasonal troposphere model, calibration errors in the troposphere contribute 1 or 2 mHz to the measurement error. With the GPS calibration of the troposphere, one may be able to obtain a factor of five reduction in the troposphere error. With station location errors at about the 10-cm level, the inherent Doppler measurement accuracy of 1 to 2 mHz may be fully utilized. A 1-mHz Doppler measurement error would translate into about a 0.02 mm/sec error in the line-of-sight velocity and, according to recent studies [7], perhaps 50 nrad in angular measurement accuracy.

References

- [1] T. D. Moyer, *Mathematical Formulation of the Double-Precision Orbit Determination Program (DPODP)*, JPL Technical Report 32-1527, Jet Propulsion Laboratory, Pasadena, California, May 15, 1971.
- [2] C. C. Chao, "New Tropospheric Range Corrections with Seasonal Adjustment," *Deep Space Network Progress Report 32-1526*, vol. 6, Jet Propulsion Laboratory, Pasadena, California, pp. 67-73, December 15, 1971.
- [3] C. C. Chao, "A New Method to Predict Wet Zenith Range Correction From Surface Measurements," *Deep Space Network Progress Report 32-1526*, vol. 14, Jet Propulsion Laboratory, Pasadena, California, pp. 33-41, April 15, 1973.
- [4] H. N. Royden, D. W. Green, and G. R. Walson, "Use of Faraday-Rotation Data from Beacon Satellites to Determine Ionosphere Corrections for Interplanetary Spacecraft Navigation," *Proc. Satellite Beacon Symposium*, Warszawa, Poland, May 19-23, 1980, pp. 345-355, 1981.
- [5] T. D. Moyer, "Transformation from Proper Time on Earth to Coordinate Time in Solar System Barycentric Space-Time Frames of Reference: Parts 1 and 2," *Celestial Mechanics*, vol. 23, pp. 33-68, January 1981.
- [6] R. W. Hellings, "Relativistic Effects in Astronomical Timing Measurements," *The Astronomical Journal*, vol. 91, no. 3, pp. 650-659, March 1986.
- [7] S. W. Thurman and J. A. Estefan, "Radio Doppler Navigation of Interplanetary Spacecraft Using Different Data Processing Modes," *Advances in the Astronautical Sciences*, Paper AAS 93-163, AAS/AIAA Spaceflight Mechanics Meeting, Pasadena, California, February 22-24, 1993 (to be published).

Table 1. Doppler error sensitivities to clock error terms.

Error term	Amplitude, sec	Frequency, rad/sec	Doppler error, Hz
Daily	1.0×10^{-12}	7.292×10^{-5}	6.77×10^{-8}
Biweekly	1.0×10^{-7}	5.209×10^{-6}	3.45×10^{-5}
Annual	1.5×10^{-6}	1.991×10^{-7}	7.57×10^{-7}

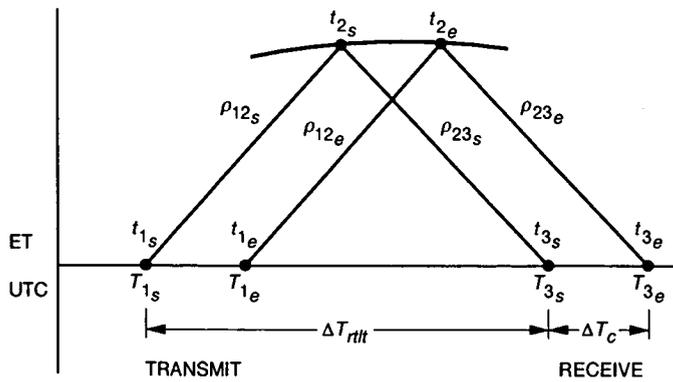


Fig. 1. Doppler observable schematic diagram.

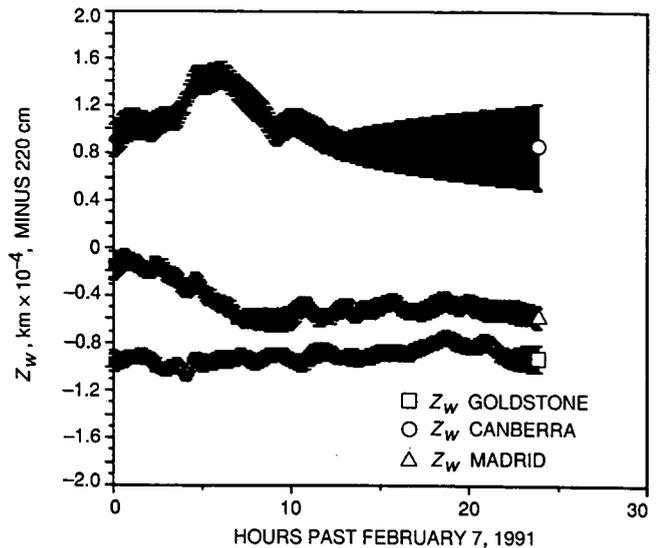


Fig. 3. GPS troposphere calibrations for February 7, 1991.

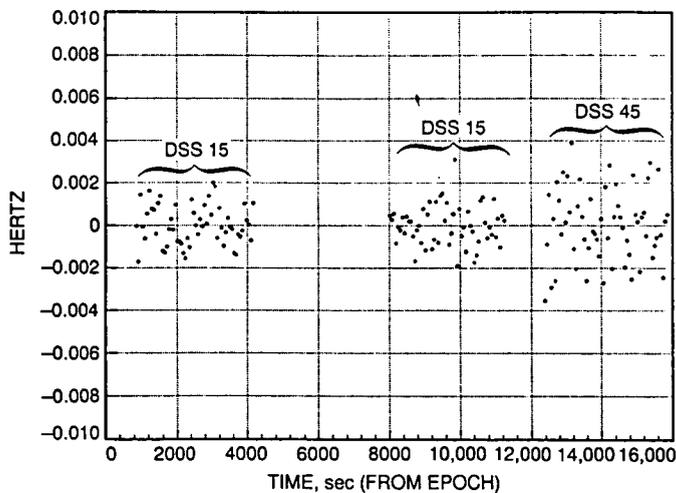


Fig. 2. Magellan Doppler residuals starting from February 7, 1991 19:29:02 (ET).

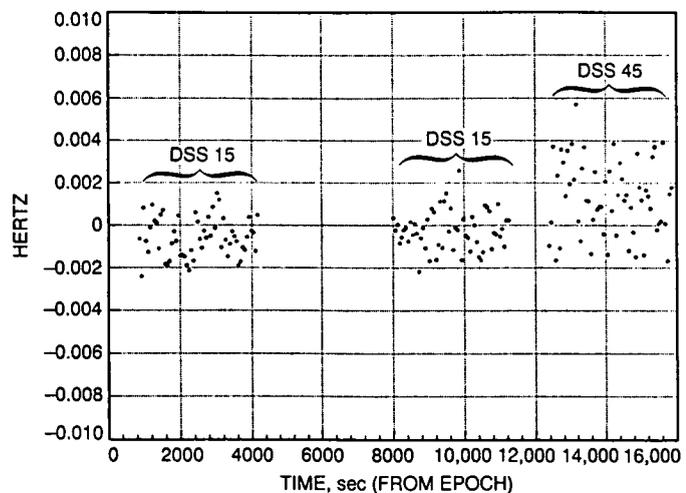


Fig. 4. Magellan Doppler residuals with simulated GPS troposphere calibrations.

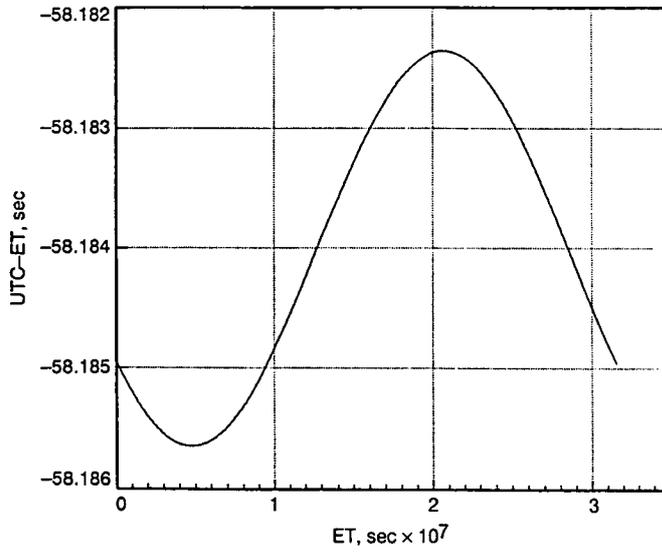


Fig. 5. Integrated UTC minus ET as a function of ET.

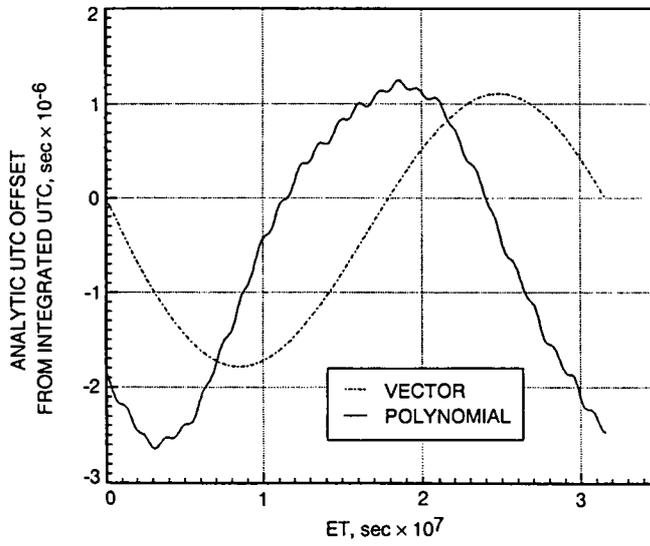


Fig. 6. Analytic UTC minus integrated UTC as a function of ET.

53-13

167923

TDA Progress Report 42-113

P - 15

N93-29589

500692

May 15, 1993

Precise Tracking of the Magellan and Pioneer Venus Orbiters by Same-Beam Interferometry— Part II: Orbit Determination Analysis

W. M. Folkner, J. S. Border, S. Nandi, and K. S. Zukor
Tracking Systems and Applications Section

A new radio metric positioning technique has demonstrated improved orbit determination accuracy for the Magellan and Pioneer Venus Orbiter orbiters. The new technique, known as Same-Beam Interferometry (SBI), is applicable to the positioning of multiple planetary rovers, landers, and orbiters which may simultaneously be observed in the same beamwidth of Earth-based radio antennas. Measurements of carrier phase are differenced between spacecraft and between receiving stations to determine the plane-of-sky components of the separation vector(s) between the spacecraft. The SBI measurements complement the information contained in line-of-sight Doppler measurements, leading to improved orbit determination accuracy. Orbit determination solutions have been obtained for a number of 48-hour data arcs using combinations of Doppler, differenced-Doppler, and SBI data acquired in the spring of 1991. Orbit determination accuracy is assessed by comparing orbit solutions from adjacent data arcs. The orbit solution differences are shown to agree with expected orbit determination uncertainties. The results from this demonstration show that the orbit determination accuracy for Magellan obtained by using Doppler plus SBI data is better than the accuracy achieved using Doppler plus differenced-Doppler by a factor of four and better than the accuracy achieved using only Doppler by a factor of eighteen. The orbit determination accuracy for Pioneer Venus Orbiter using Doppler plus SBI data is better than the accuracy using only Doppler data by 30 percent.

I. Introduction

Remote reconnaissance of planets in our solar system is conducted by NASA using unmanned space probes. A hyperbolic flyby of a planetary system may provide a few snapshots of geologic, atmospheric, and electromagnetic phenomena, which then reveal, through analyses, some un-

derstanding of the underlying physical processes which are taking place. A spacecraft placed in orbit about a distant planet, on the other hand, will provide a much longer time history of measurements of various phenomena, leading to more comprehensive physical understandings. Navigation is one of the many critical engineering functions necessary to support the planning and operations of space flight

missions. This article presents results of a flight demonstration of a new technique for improving navigation for planetary orbiters.

Radio antennas in the DSN provide communication links with distant spacecraft. Measurements of the microwave signal used for commanding the spacecraft and for relaying telemetry data from the spacecraft to Earth provide the basis for radio navigation. Any change in range between a Deep Space Station and a spacecraft affects the Doppler shift of the transmitted radio signal. Though many techniques, including ranging, radio interferometry, and onboard optical imaging, are used for interplanetary navigation, orbit determination for planetary orbiters has relied primarily upon Doppler data. The motion of an orbiter about a planet, induced by gravity, places a strong signature in the Doppler data received at Earth. Dynamic models allow the state of the orbiter relative to the central body to be estimated from a time history of the Doppler shift.

The accuracy of navigation solutions and the ability to project the spacecraft trajectory forward may directly impact the quality of the science return. Pointing, scheduling, and configuration of onboard instruments rely upon predictions of the spacecraft trajectory. Interpretation and registration of images and other measurements rely upon reconstruction of the spacecraft trajectory. Determination of harmonic coefficients of the planet's gravity field depends directly on the orbit determination accuracy. Improvements to navigation, such as reducing the volume of tracking time necessary to maintain a specified level of orbital accuracy, predicting a trajectory further ahead within a specified error tolerance, or improving the accuracy of the final reconstructed trajectory solution, can simplify operations and enhance the science return.

For a short-period (2-24 hr) planetary orbiter, the orientation of the orbit plane is the trajectory component least well determined by line-of-sight Doppler measurements. Doppler data acquired simultaneously at two widely spaced DSN stations, and then differenced, provide sensitivity to the orientation of the orbit plane [1]. Differenced-Doppler has been used operationally during the orbit phase of the Magellan mission to help meet stringent navigation requirements [2,3]. For the case when two spacecraft are in orbit about the same planet, an observable formed from Doppler measurements, differenced between stations and differenced between spacecraft, is expected to provide further improvements to navigation [4]. A demonstration of this technique using the Magellan and Pioneer Venus orbiters at Venus took place in the spring of 1991. A detailed discussion of the data acquisition and measurement error analysis has been given earlier [5].

In 1991 the Pioneer Venus Orbiter (PVO) spacecraft, launched in 1978, was in a highly eccentric orbit about Venus with a period of about 24 hours. The Magellan spacecraft joined PVO in orbit around Venus on August 10, 1990. During 1991, Magellan was in a less eccentric orbit with a period of about 3.26 hours. Same-Beam Interferometry (SBI) data sets were acquired in February and April 1991. Orbit determination solutions from these data sets have been obtained using various combinations of Doppler, differenced-Doppler, and SBI data. Formal errors associated with the solutions and solution comparisons for adjacent data arcs are examined to assess orbit determination accuracy. An overview of the simultaneous tracking technique is presented below, followed by discussions of the data scheduling, orbit determination strategy, and orbit determination results.

II. Radio Metric Measurements

Three types of radio metric measurements were included in this demonstration: two-way Doppler, differenced-Doppler, and SBI. Two-way Doppler is collected for a single spacecraft from a single Deep Space Station. Differenced-Doppler is collected by two widely separated Deep Space Stations for a single spacecraft. SBI is collected for two spacecraft simultaneously at two widely separated Deep Space Stations. The DSN Deep Space Stations used are located in California, Australia, and Spain.

Two-way Doppler (referred to below as Doppler) is collected when the Deep Space Station sends a stable carrier signal to the spacecraft and the spacecraft replies with a signal phase-locked to the uplinked signal. The frequency shift of the signal received by the Deep Space Station compared to the transmitted signal provides a measure of the rate of change of range to the spacecraft. The DSN currently tracks planetary orbiters at either S-band (2.3 GHz) or at X-band (8.4 GHz). During the time of interest for this demonstration, most of the Magellan Doppler data acquired were derived from the station transmitting and receiving signals at X-band while PVO data were derived from a station transmitting and receiving a signal at S-band. For Doppler at S-band, the intrinsic data accuracy is limited by solar charged particle fluctuations to about 1.0 mm/sec for the inferred range-rate for a 60-sec averaging time. For data taken at X-band, the accuracy is also limited by solar charged particle fluctuations, but at a reduced level. The X-band Doppler intrinsic data accuracy is typically about 0.1 mm/sec for the inferred range-rate for a 60-sec averaging time.

Differenced-Doppler data are collected when the spacecraft carrier signal is measured at two Deep Space Stations.

The difference in the received carrier frequencies provides a measure of the difference in the range-rate from each Deep Space Station to the spacecraft. One component of the spacecraft velocity in the plane normal to the line-of-sight (plane of the sky) is inferred from this difference in line-of-sight range-rate, namely the component in the direction of the vector separating the two Deep Space Stations projected onto the plane of the sky. The intrinsic accuracy of X-band differenced-Doppler is typically 0.05 mm/sec for the inferred differenced line-of-sight range-rate for a 60-sec averaging time. Accuracy is improved relative to Doppler because station differencing reduces the effect of solar plasma fluctuations by removing fluctuations common to the two downlink ray paths. The accuracy with which the plane of sky velocity component is inferred is approximately the differenced range-rate accuracy times the ratio of the Earth-spacecraft distance to the distance between the two Deep Space Stations.

The SBI measurement of two spacecraft is depicted in Fig. 1. The two spacecraft in orbit about the same planet are so close angularly, as seen from Earth, that they may be observed in the same beamwidth of an Earth-based radio antenna. Each spacecraft carrier signal phase is recorded by two widely separated Deep Space Stations. Differencing the received carrier phases, first between stations and then between spacecraft, gives a measure of the separation of the two spacecraft in the plane of the sky along the projected baseline. The phase difference can be ambiguous by an integer number of cycles; the ambiguity must be resolved by a priori information (such as a sufficiently precise Doppler-only orbit) or by estimating a phase bias parameter for each SBI data arc. SBI data were taken at S-band for this orbit determination demonstration since PVO was tracked at S-band and Magellan was transmitting low-rate data at S-band in addition to the primary X-band signal. The SBI data accuracy corresponded to a doubly-differenced range accuracy of 1.5 mm for 5-min integration times [5]. From this doubly-differenced phase the separation of the two spacecraft in the plane of the sky can be inferred with an angular accuracy of 180 prad for a baseline length of 8000 km (which is an average length of the separation vector between antennas from different DSN complexes projected onto the plane normal to the Earth-spacecraft direction). At a distance of 1.5 astronomical units (AU's), the SBI data accuracy corresponds to a spacecraft-separation measurement accuracy of 40 m. X-band data are expected to be more accurate by an order of magnitude due to reduced sensitivity to solar charged particle fluctuations.

All of the radio signals are affected by delays due to Earth ionosphere and troposphere as well as delays due to solar plasma. Calibrations for the troposphere were ap-

plied based on a seasonal model [6]. Calibrations for the Earth's ionosphere were applied based on daily measurements from Earth-orbiting beacon satellites [7].

III. Estimation Models

The spacecraft trajectory was integrated from initial position and velocity conditions (epoch state) using models for the dynamic forces on the spacecraft. The largest force was due to the gravitational field of Venus, which was modeled as a point mass (gravitational mass [GM]) term and potential field composed of spherical harmonic terms to degree and order 21 estimated from several years of radio metric data for PVO and Magellan [8]. Other significant forces were due to solar pressure, the solar point mass perturbation and, for Magellan, atmospheric drag and momentum wheel desaturation thrusts which occur twice daily. The right ascension and declination of the Venus spin axis and the rotation period were derived from Magellan radar images of surface features [9]. The rotation angle of Venus about the spin axis at a reference epoch was estimated for this demonstration from a 10-day arc of PVO and Magellan Doppler data.

The Deep Space Station locations were mapped from Earth-fixed locations to inertial space using models for precession, nutation, and solid Earth tides, and calibrations for polar motion and length of day variations. Computed values for measurements were derived from nominal values for the spacecraft epoch state, force models, and inertial Deep Space Station locations. A least-squares fit to the observations minus the computed measurements was made to estimate model parameters. For this demonstration, gravity field parameters were not adjusted since, for short data arcs, epoch state errors can be aliased into gravity field parameters. The uncertainties in the spacecraft trajectory caused by imperfect unadjusted model parameters were included through the use of consider analysis [10]. The derived uncertainty in the trajectories depends on the formal error covariance for the solved-for parameters (computed) and on the uncertainty assumed a priori of the unadjusted (considered) parameters.

For short data arcs, the spacecraft trajectory uncertainty is usually dominated by the uncertainty in the unadjusted gravity field. Because of this, the optimal orbit determination solution may not be achieved by weighting all of the data at its intrinsic accuracy since the estimated epoch state is derived by neglecting the considered parameters. By neglecting the gravity field, the estimation filter will produce a solution based on an over-optimistic estimate for the spacecraft plane-of-sky velocity based on the Doppler data. Without taking this into account in some

manner the differential data types may not fully influence the solution. The effect of mismodeling can be reduced by deweighting the Doppler data and including differential data, weighted at its intrinsic accuracy, in the estimation. This strategy has been used in the operational navigation for Magellan [2,3].

IV. Data Arcs

During the spring of 1991, Magellan was conducting radar mapping operations. Magellan typically performed radar mapping for one hour of each orbit, during which there was no signal transmitted to Earth. Radio metric data could be collected for Magellan during the two hours of telemetry playback each orbit. For the same time period, PVO was in an orbit with a period of 24 hours. S-band Doppler data from PVO were collected for approximately 6 hours per day centered roughly about periapsis, which occurred during the California–Australia visibility period. SBI data were acquired over an eight-day period beginning February 16, 1991, and over a ten-day period beginning April 6, 1991. Figures 2 and 3 show the orbits of the two spacecraft as viewed from Earth for these two time periods. Nominal orbital elements for the spacecraft are given with respect to the plane of the sky in Tables 1 and 2.

For the SBI demonstration, data could be acquired only when both Magellan and PVO were transmitting to Earth and when stations were allocated at two DSN complexes. Because Magellan used differenced-Doppler operationally, stations from different DSN complexes were scheduled to simultaneously track Magellan for about five hours during each 48-hour period. SBI data could then be acquired at those stations, on a non-interference basis, when PVO was also transmitting. This scheduling resulted in an average of two hours of SBI data acquired every other day. One SBI bias parameter was needed for each hour of SBI data since the Magellan signal was interrupted by either a mapping cycle or an attitude calibration after each hour of telemetry.

Because of the sparseness of the SBI and PVO Doppler data, 48-hour non-overlapping data arcs were chosen for orbit determination solutions. A typical data schedule for a 48-hour data arc is shown in Fig. 4. Each of the data arcs contained approximately 13 orbits of Doppler data from Magellan, about 5 hours of differenced-Doppler for Magellan, and Doppler from two PVO orbits, each with about 6 hours of data centered about periapsis.

Four two-day data arcs were formed for the period February 14 to February 22, 1991, as summarized in Table 3. Since no Doppler data were collected from PVO for

the orbit beginning on February 14, 1991, Doppler data from the previous orbit were included to allow each solution to contain data from two PVO orbits. PVO Doppler data within 1 hour of periapsis were excluded to reduce sensitivity to gravity field mismodeling. SBI data were acquired on the California–Australia baseline near the time of PVO periapsis. Five data arcs were formed from data acquired from April 6 to April 16, 1991, as summarized in Table 4. Most of the SBI data acquired in April were during the California–Australia overlap period with some data also acquired from the California–Spain baseline.

V. Orbit Determination Strategy

For this demonstration, orbit solutions for each spacecraft were formed for each data arc using different combinations of data. Three combinations of data were used for Magellan: Doppler only, Doppler plus differenced-Doppler, and Doppler plus SBI. Solutions for PVO were formed using only Doppler data and using Doppler plus SBI data. Orbit determination accuracy was assessed by comparing the orbit solutions for adjacent data arcs. To do this, the orbit solution from each data arc was propagated forward to the first orbit in the succeeding data arc and differenced with the succeeding solution trajectory. This solution-to-solution consistency provides one measure of orbit determination accuracy for post-fit data analysis. Orbit prediction, while of interest for mission operations, is not addressed here because neither experiment (February 1991 or April 1991) was long enough to provide more than one or two orbit prediction comparisons for prediction times of approximately one week (which is the typical period of interest).

The quantities estimated for each data arc were six epoch-state parameters for each spacecraft, an atmospheric drag coefficient for Magellan, and phase biases for the SBI data. The epoch for each spacecraft was chosen to be an apoapsis near the beginning of the 48-hour data arc. A priori uncertainties for the estimated state and phase bias parameters were very large so as not to significantly constrain the solution. The a priori uncertainty for the Magellan atmospheric drag was 100 percent of its nominal value; this uncertainty is consistent with variations in the Venus atmospheric density above 100 km [11]. (Because the gravity field is mismodeled and no gravity field corrections were estimated, the estimated atmospheric drag tended to absorb gravity field mismodeling and hence not represent the actual atmospheric drag. This estimation of atmospheric drag is used here to allow comparison with other Magellan orbit determination solutions [2,3].)

The data weights used in the solutions varied depending on which combinations of data were used. The intrinsic

accuracies of the data for a 60-sec sampling time were assumed to be 0.1 mm/sec for the Magellan X-band Doppler data, 1 mm/sec for the S-band PVO Doppler data, and 0.05 mm/sec for the Magellan X-band differenced-Doppler data. The accuracy of the SBI data was 1.5 mm for 5-min averaging times. When fitting only Doppler data for Magellan and PVO, the Doppler data were weighted at their intrinsic accuracy. When fitting Magellan using Doppler and differenced-Doppler data, first the Doppler data were fit. Next, the differenced-Doppler data were included, weighted at their intrinsic accuracy, and the Doppler data deweighted by an increasing factor until the root-mean-square Doppler residual increased by 10 percent over the Doppler-only case. The typical deweighting factor for the Doppler data was 10–20. This empirically derived procedure allowed the differenced-Doppler data to influence the solution without unduly weakening the Doppler data [2,3]. When fitting Doppler and SBI data for Magellan and PVO, the SBI and PVO Doppler data were weighted at their intrinsic accuracies while it was found necessary to deweight the Magellan Doppler data by a factor of 100 to allow the post-fit SBI residuals to be minimized.

The orbit solutions for Magellan using Doppler plus differenced-Doppler data were similar to the operational orbit solutions. Operational orbit determination is performed using data arcs covering twelve orbits, using X-band Doppler and differenced-Doppler data. Consecutive operational Magellan orbit solutions use overlapping data arcs with four orbits of data in common between solutions. This has provided the sub-kilometer solution-to-solution consistency needed by the radar mapping instrument [2,3].

VI. Orbit Determination Covariance

In addition to comparing successive orbit solutions to measure orbit determination accuracy, the solution-to-solution differences are compared below to a nominal orbit covariance. This orbit covariance was formed using a priori uncertainties for a number of consider parameters. Table 5 lists the a priori uncertainties assumed.

The gravity field uncertainty was a major error source for all solutions and dominated the orbit uncertainty for solutions using only Doppler data. Due to computational limitations, the considered gravity field covariance was of degree and order 6 rather than the covariance of the field of degree and order 21. A diagonal covariance of degree and order 6 was taken from a previous gravity field determination [12] scaled by an empirically determined factor of 1.5. With this assumed gravity field uncertainty, the observed solution-to-solution variations for Magellan solutions using only Doppler data approximately agreed with

the considered gravity field uncertainty. The uncertainty in Venus' GM was similarly taken to be a value which gave approximately the observed variation in determination of the Magellan and PVO semi-major axes from solutions using only Doppler data. These assumptions for the gravity field uncertainty were adopted only to give an appropriate spacecraft trajectory uncertainty for solution-to-solution differences for this demonstration.

Solar pressure forces were considered with an uncertainty of 10 percent of their nominal value. The zenith ionosphere uncertainty was taken to be 10^{17} electrons/m² which is a typical uncertainty in the daily ionosphere calibration [7]. The zenith troposphere uncertainty was taken to correspond to 4 cm of path delay due to observed variation in water vapor content compared with the seasonal model employed.¹

Small thruster firings occurred twice daily for Magellan to desaturate momentum wheels used to control the spacecraft attitude. The effect of these thruster firings on the spacecraft trajectory was modeled as an impulsive maneuver. The magnitude of the velocity imparted to the spacecraft from each maneuver was typically ~ 3 mm/sec. Calibrations for the thruster firings are provided on the spacecraft telemetry from which the magnitude of the maneuver can be determined to a few percent [13]. The uncertainty in each maneuver was assumed to be 0.1 mm/sec.

Uncertainties in station frequency and timing standards affect station-differenced data types more strongly than single-station Doppler data. The uncertainty in station frequency calibrations was important for solutions containing differenced-Doppler data. The station frequency calibration uncertainty was assumed to be 5×10^{-14} sec/sec [14]. Uncertainties in station clock epoch were important for solutions containing SBI data. The effect of an unknown offset in the time-tags for the SBI data at the two Earth receivers is discussed in [5]. For analysis of the SBI data, nominal values for the difference in station clocks between the DSN stations were taken from Very Long Baseline Interferometry measurements made routinely for maintaining knowledge of Earth orientation.² The uncertainty of this determination of the station-differenced clock epoch uncertainty was 0.2 μ sec.

¹ S. E. Robinson, "Errors in Surface Model Estimates of Zenith Wet Path Delays Near DSN Stations," JPL Interoffice Memorandum 335.4-594 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 3, 1986.

² S. H. Oliveau, L. Sung, and J. A. Steppe, "TEMPO Group Clock Synchronization and Syntonization Report from the DSN VLBI Mark IV-85 System," JPL Engineering Memorandum 335-192 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 25, 1991.

Figures 5 and 6 show the root-sum-square (rss) position covariance for Magellan and PVO at apoapsis using several combinations of data types. The rss position uncertainty is usually largest at apoapsis due to the fact that the uncertainty in determination of the longitude of the orbit ascending node with respect to the plane of the sky dominates the uncertainty in spacecraft position determination. The position uncertainty for both spacecraft when only Doppler data are included is dominated by the considered gravity field uncertainty. This is in contrast to the case for PVO when using a one-day data arc where the data noise dominated the position uncertainty and the position determination uncertainty was much larger [15]. The position determination uncertainty for Magellan using Doppler plus differenced-Doppler data contains nearly equal contributions from data noise, troposphere, clock rate, and gravity field uncertainties. The position uncertainties for solutions containing SBI data are dominated by gravity field uncertainty but at a reduced level. With the 48-hour Doppler data arc the position improvement for PVO when SBI data are included is much less than if a one-day data arc is used [15].

These orbit determination covariances are nominal only for this demonstration period, especially for data types other than SBI. No attempt has been made to optimize orbit determination performance by altering data scheduling, elevation cutoff, or different data weighting algorithms [16]. The orbit covariances are used primarily to check that the observed solution-to-solution differences are understood in terms of known mismodeled parameters.

VII. Orbit Determination Results

Figure 7 shows the rss position differences between solutions for Magellan from adjacent data arcs using only Doppler data, plotted over one orbit. The expected differences are also shown. The expected position difference curves are derived by assuming each solution is an independent sample from a distribution characterized by the formal covariance given in the previous section. Thus the expected difference between two solutions that use similar data schedules is just the position uncertainty for either solution times the square root of two. The Doppler-only solution statistics are dominated by the (considered) gravity field uncertainty. The gravity field uncertainty was determined in such a way as to get approximate agreement between the expected position difference and the actual solution differences for Doppler-only solutions for Magellan.

Figure 8 shows the rss position differences and the expected position difference for Magellan orbit solutions from

adjacent data arcs using Doppler plus differenced-Doppler data. The Magellan mission requirement is for adjacent solutions to differ by less than ~ 1.4 km (0.15 km radial, 1 km cross-track, and 1 km down track) over the mapping period of the orbit, which is approximately the central 1-hour period shown in Figs. 7, 8, and 9.³ It can be seen that Doppler data alone would not satisfy the mission requirements. The Doppler plus differenced-Doppler solutions generally satisfy the mission requirements (and could be improved by using overlapping data arcs as is done operationally). Note that the trajectory differences for the Doppler plus differenced-Doppler solutions are consistently less than the value expected from the covariance analysis. This implies that considering a constant zenith troposphere uncertainty of 4 cm overestimates the effect on the trajectory, possibly because the deviations from the calibrations for the two sample time periods were smaller than normal.

Figure 9 shows the orbit solution differences for Magellan solutions using Doppler plus SBI data. The Doppler plus SBI solutions are seen to be significantly better than the Doppler-only or Doppler plus differenced-Doppler solutions as expected. This is true even though the SBI data were acquired at S-band while the differenced-Doppler data were acquired at X-band. SBI data acquired at X-band are expected to be more accurate by about one order of magnitude [5].

Figure 10 shows the rss position differences and the expected position difference for PVO orbit solutions from adjacent data arcs using only Doppler data. Figure 11 shows the orbit solution differences for PVO solutions using Doppler plus SBI data. The PVO Doppler-only solution differences are much smaller than the Magellan solution differences for solutions with either Doppler-only or Doppler plus differenced-Doppler. This is due to the PVO periapsis altitude being much higher than Magellan's periapsis altitude, which makes the PVO orbit determination much less sensitive to gravity field mismodeling. The addition of SBI data only slightly improves the PVO orbit solutions because, with a two-day data arc and low sensitivity to gravity field mismodeling, the Doppler data determine the longitude of the ascending node of PVO's orbit with accuracy comparable to the SBI data.

Table 6 lists the time-averaged orbit position difference and an overall average for Magellan and PVO for each combination of data types studied. This figure of merit is introduced to quantitatively compare the orbit determination performance. Using the seven solution-to-solution

³ S. N. Mohan, *Magellan Navigation Plan*, JPL Document 630-51, Rev. B (internal document), Jet Propulsion Laboratory, Pasadena, California, March 23, 1988.

comparisons possible for this demonstration, Table 6 indicates that orbit solutions using SBI data are significantly more accurate for Magellan and slightly more accurate for PVO. The quantitative ratios will, in general, depend on orbit geometry, data arcs, and estimation strategy.

VIII. Conclusion

Orbit determination results have been obtained for Pioneer Venus Orbiter and Magellan using same-beam interferometry data, which is a new data type for plane-

tary orbiter navigation. The orbit determination accuracy using this data type, based on solution-to-solution consistency, has been explained in terms of nominal error models. For the particular orbit determination strategy and observational geometry used for this limited data set, the orbit determination accuracy for Magellan using SBI in combination with two-way Doppler data is better by a factor of four than orbit determination accuracy using two-way Doppler plus differenced-Doppler data and better by a factor of eighteen than orbit determination accuracy using Doppler alone. This new data type should find much application in the future as more missions with multiple orbiters and/or landers are flown to Mars.

Acknowledgments

The authors thank Doug Engelhardt of the Magellan navigation team and Neil Mottinger of the Pioneer Venus Orbiter navigation team for their assistance in spacecraft modeling.

References

- [1] S. R. Poole, M. P. Ananda, and C. E. Hildebrand, "Radio Interferometric Measurements for Accurate Planetary Orbiter Navigation," in *AAS Advances in the Astronautical Sciences*, vol. 40, part 1, pp. 93-111, San Diego, California: Univelt Inc., 1980.
- [2] D. B. Engelhardt, J. B. McNamee, S. K. Wong, F. G. Bonneau, E. J. Graat, R. J. Haw, G. R. Kronschnabl, and M. S. Ryne, "Determination and Prediction of Magellan's Orbit," paper AAS-91-180, presented at the AAS/AIAA Spaceflight Mechanics Meeting, Houston, Texas, February 11-13, 1991.
- [3] J. D. Giorgini, E. J. Graat, T.-H. You, M. S. Ryne, S. K. Wong, and J. B. McNamee, "Magellan Navigation Using X-Band Differenced Doppler During Venus Mapping Phase," paper AIAA-92-4521, presented at the AIAA/AAS Astrodynamics Conference, Hilton Head, South Carolina, August 10-12, 1992.
- [4] W. M. Folkner and J. S. Border, "Orbiter-Orbiter and Orbiter-Lander Tracking Using Same-Beam Interferometry," *The Telecommunications and Data Acquisition Progress Report 42-109*, vol. January-March 1992, Jet Propulsion Laboratory, Pasadena, California, pp. 74-86, May 15, 1992.
- [5] J. S. Border, W. M. Folkner, R. D. Kahn, and K. S. Zukor, "Precise Tracking of the Magellan and Pioneer Venus Orbiters by Same-Beam Interferometry, Part I: Data Accuracy Analysis," *The Telecommunications and Data Acquisition Progress Report 42-110*, vol. April-June 1992, Jet Propulsion Laboratory, Pasadena, California, pp. 1-20, August 15, 1992.

- [6] C. C. Chao, *The Troposphere Calibration Model for Mariner Mars 1971*, JPL Technical Report 32-1587, Jet Propulsion Laboratory, Pasadena, California, pp. 61-76, March 1974.
- [7] H. N. Royden, D. W. Green, and G. R. Walson, "Use of Faraday-Rotation Data from Beacon Satellites to Determine Ionospheric Corrections for Interplanetary Spacecraft Navigation," *Proceedings of the Satellite Beacon Symposium*, edited by A. W. Wernik, Warszawa, Poland, pp. 345-355, May 1980.
- [8] J. B. McNamee, G. R. Kronschnabl, S. K. Wong, and J. E. Ekelund, "A Gravity Field to Support Magellan Navigation and Science," *J. Astron. Sci.*, vol. 40, pp. 107-134, 1992.
- [9] M. E. Davies, T. R. Colvin, P. G. Rogers, P. W. Chodas, W. L. Sjogren, E. L. Akim, V. A. Stepanyantz, Z. P. Vlasova, and A. I. Zakharov, "The Rotation Period, Direction of the North Pole, and Geodetic Control Network of Venus," *J. Geophys. Res.*, vol. 97, pp. 13,141-13,151, 1992.
- [10] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, San Diego, California: Academic Press, 1977.
- [11] G. M. Keating, J. L. Bertaux, S. W. Bougher, T. E. Cravens, R. E. Dickinson, A. E. Hedin, V. A. Krasnopolsky, A. F. Nagy, J. Y. Nicholson III, L. J. Paxton, and U. von Zahn, "Model of Venus Neutral Upper Atmosphere: Structure and Composition," *Adv. Space Res.*, vol. 5, pp. 117-171, 1985.
- [12] N. A. Mottinger, W. L. Sjogren, and B. G. Bills, "Venus Gravity: A Harmonic Analysis and Geophysical Implications," *J. Geophys. Res.*, vol. 90, supplement, pp. C739-C756, February 15, 1985.
- [13] D. B. Engelhardt and S. N. Mohan, "Deterministic Errors in the Magellan Orbit Due to Attitude Control Thruster Activity," paper AIAA-89-0349, presented at the 27th Aerospace Sciences Meeting, Reno, Nevada, January 9-12, 1989.
- [14] P. A. Clements, A. Kirk, and R. Unglaub, "Results of Using the Global Positioning System to Maintain the Time and Frequency Synchronization in the Deep Space Network," *The Telecommunications and Data Acquisition Progress Report 42-89*, vol. January-March 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 67-72, May 15, 1987.
- [15] W. M. Folkner, D. B. Engelhardt, J. S. Border, and N. A. Mottinger, "Orbit Determination for Magellan and Pioneer 12 Using Same-Beam Interferometry," paper AAS 91-393, presented at the AAS/AIAA Astrodynamics Specialist Conference, Durango, Colorado, August 19-22, 1991.
- [16] J. S. Ulvestad, "Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories, Part II: 8.4 GHz Performance and Data-Weighting Strategies," *The Telecommunications and Data Acquisition Progress Report 42-108*, vol. October-December 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 49-65, February 15, 1992.

Table 1. Spacecraft orbital elements with respect to the plane-of-sky on February 16, 1991.

Element	Magellan	PVO
Semimajor axis, km	10425.0	39453.4
Eccentricity	0.39287	0.81985
Inclination, deg	38.301	122.99
Argument of perigee, deg	74.797	88.513
Longitude of ascending node, deg	-15.387	-41.412
Time past periapsis, sec	-5830.6	-43194
Epoch	12:06:00	13:36:04
Periapsis altitude, km	278	1056

Table 2. Spacecraft orbital elements with respect to the plane-of-sky on April 6, 1991.

Element	Magellan	PVO
Semimajor axis, km	10425.0	39450.8
Eccentricity	0.39307	0.82352
Inclination, deg	22.541	65.005
Argument of perigee, deg	-83.371	103.64
Longitude of ascending node, deg	147.50	-30.623
Time past periapsis, sec	-5866.8	-43183
Epoch	12:45:15	13:25:13
Periapsis altitude, km	278	1056

Table 3. Data coverage for SBI demonstration.

Solution	February 1991			
	13-16	16-18	18-20	20-22
PVO start time	Feb. 13, 13:36	Feb. 16, 13:36	Feb. 18, 13:36	Feb. 20, 13:36
Magellan start time	Feb. 14, 14:27	Feb. 16, 12:05	Feb. 18, 12:59	Feb. 20, 13:52
Data stop time	Feb. 16, 12:05	Feb. 18, 12:59	Feb. 20, 13:36	Feb. 22, 11:42
PVO Doppler, hr	11.3	13.4	16.3	11.9
Magellan Doppler, hr	24.8	26.0	26.0	25.2
Magellan differenced- Doppler, hr	6.2	6.0	5.0	5.0
SBI, hr	1.0	2.6	2.2	1.0

Table 4. Data coverage for SBI demonstration.

Solution	April 1991				
	6-8	8-10	10-12	12-14	14-16
PVO start time	April 6, 13:25	April 8, 13:24	April 10, 13:24	April 12, 13:23	April 14, 13:22
Magellan start time	April 6, 12:45	April 8, 10:23	April 10, 11:16	April 12, 12:10	April 14, 13:04
Data stop time	April 8, 10:23	April 10, 11:16	April 12, 12:10	April 14, 13:04	April 16, 13:30
PVO Doppler, hr	23.7	22.6	16.6	13.9	20.1
Magellan Doppler, hr	21.1	25.6	25.3	26.2	23.6
Magellan differenced- Doppler, hr	3.2	6.3	6.8	8.9	8.4
SBI, hr	1.4	1.7	3.9	3.0	1.7

Table 5. Assumptions for orbit determination covariance analysis.

Unadjusted parameters	A priori uncertainty
Solar pressure	10 percent of nominal value
Magellan thruster firing	0.1 mm/sec
Venus gravitational mass	$0.07 \text{ km}^3 \text{ sec}^{-2}$
Venus gravity field	Diagonal covariance for terms to degree and order 6 from [12] scaled by 1.5
Zenith troposphere	4 cm
Zenith ionosphere	$10^{17} \text{ electrons/m}^2$
Station clock rate	5×10^{-14}
Station-differenced clock epoch	0.2 μsec

Table 6. Time-averaged position differences for Magellan and PVO using only Doppler data, Doppler plus differenced-Doppler data, or Doppler plus SBI data.

Data	RSS position difference for Magellan, km	RSS position difference for PVO, km
Doppler only	4.35	0.32
Doppler plus differenced-Doppler	0.97	-
Doppler plus SBI	0.24	0.23

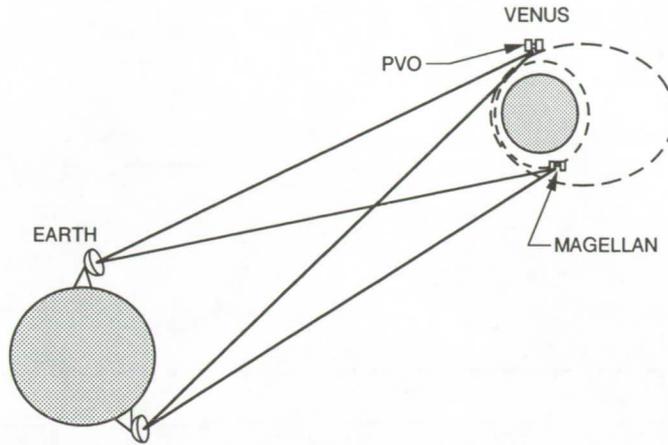


Fig. 1. Same-beam interferometry measurement technique.

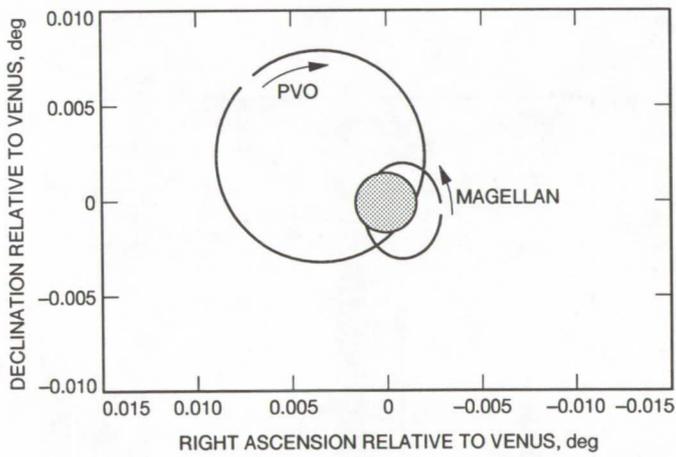


Fig. 2. Magellan and PVO orbits about Venus as seen from Earth on February 16, 1991.

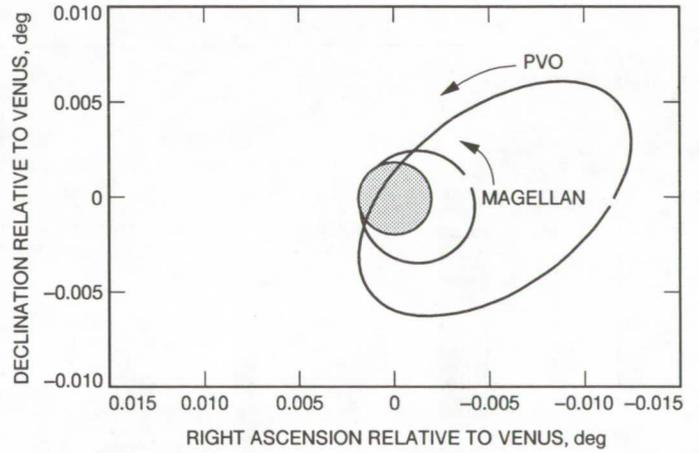


Fig. 3. Magellan and PVO orbits about Venus as seen from Earth on April 6, 1991.

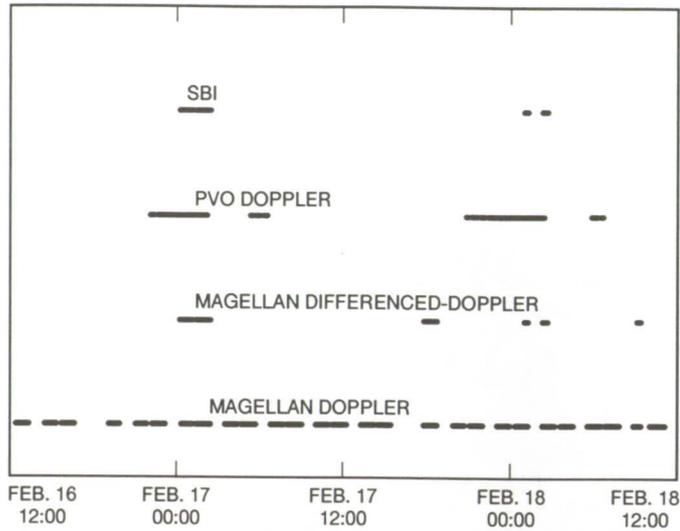


Fig. 4. Typical 48-hour data schedule for Magellan and PVO.

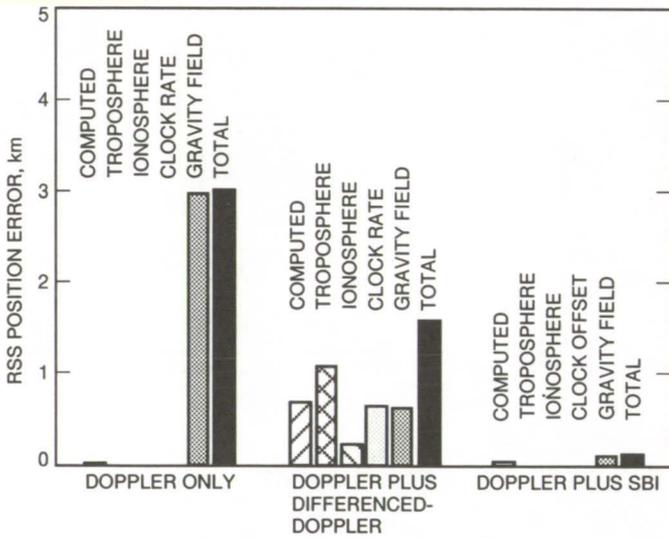


Fig. 5. Magellan position uncertainty at apoapsis using various combinations of data. Contributions from solar pressure and maneuver uncertainties are negligible and are not shown.

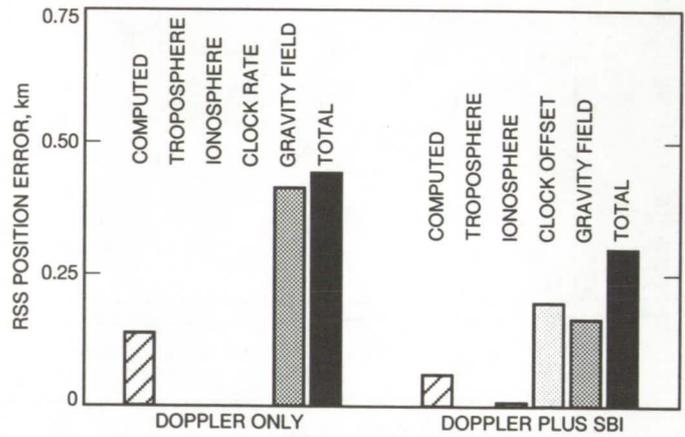


Fig. 6. PVO position uncertainty at apoapsis using various combinations of data. Contribution from solar pressure uncertainty is negligible and is not shown.

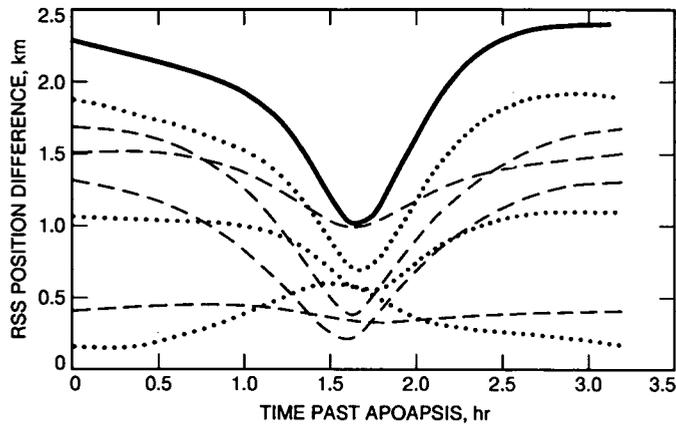


Fig. 7. Magellan solution-to-solution trajectory differences using only Doppler data. The dashed curves are for solutions in February 1991. The dotted curves are for solutions in April 1991. The dark solid curve is the expected trajectory difference based on the covariance analysis.

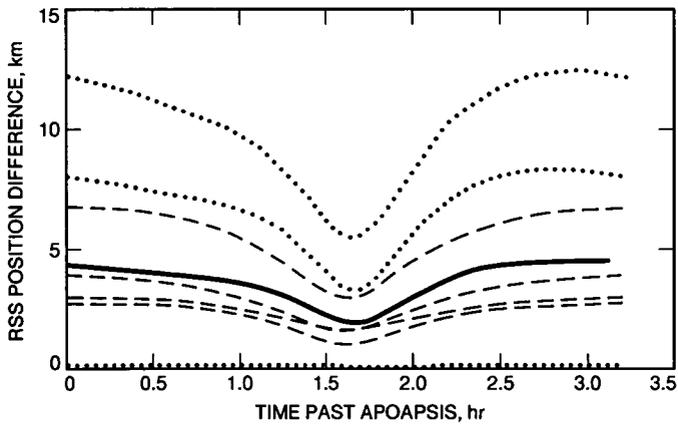


Fig. 8. Magellan solution-to-solution trajectory differences using Doppler plus differenced-Doppler data. The dashed curves are for solutions in February 1991. The dotted curves are for solutions in April 1991. The dark solid curve is the expected trajectory difference based on the covariance analysis.

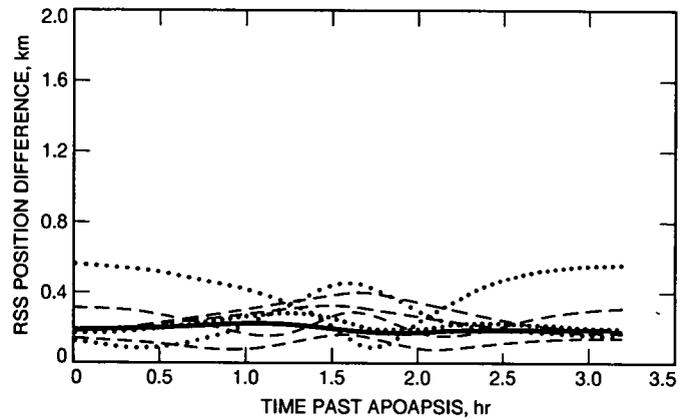


Fig. 9. Magellan solution-to-solution trajectory differences using Doppler plus SBI data. The dashed curves are for solutions in February 1991. The dotted curves are for solutions in April 1991. The dark solid curve is the expected trajectory difference based on the covariance analysis.

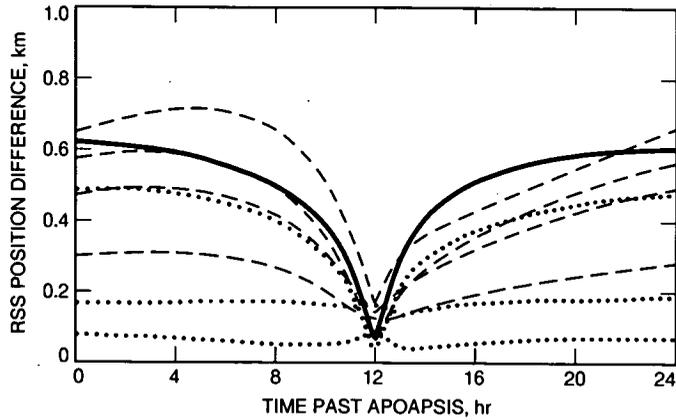


Fig. 10. PVO solution-to-solution trajectory differences using only Doppler data. The dashed curves are for solutions in February 1991. The dotted curves are for solutions in April 1991. The dark solid curve is the expected trajectory difference based on the covariance analysis.

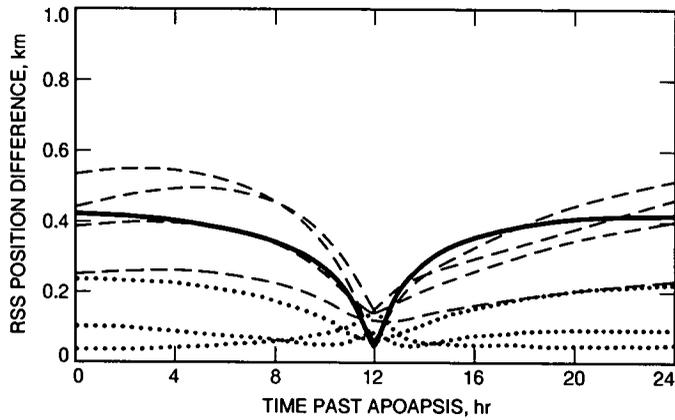


Fig. 11. PVO solution-to-solution trajectory differences using Doppler plus SBI data. The dashed curves are for solutions in February 1991. The dotted curves are for solutions in April 1991. The dark solid curve is the expected trajectory difference based on the covariance analysis.

54-33

167924

p-9

High-Power Ka-Band Amplifier

R. Cormier

Radio Frequency and Microwave Subsystems Section

Development of a high-power tube suitable to power a Ka-band (34.5-GHz) antenna transmitter located at the Goldstone, California, tracking station is continuing. The University of Maryland Laboratory for Plasma Research and JPL are conducting a joint effort to test the feasibility of phase locking a second-harmonic gyrotron both by direct injection at the output cavity and by using a priming cavity to bunch the electrons in the beam. This article describes several design options and the results of computer simulation testing.

I. Introduction

For the past 7 years, there has been continuing effort in the design and development of a high-power tube to generate 200 to 400 kW of continuous wave (CW) power at 34.5 GHz (Ka-band) for a high-power antenna transmitter to be located at Goldstone, California. Varian Associates completed the design of a 400-kW gyroklystron in 1986 [1]. Subsequently, several studies were made of the feasibility of phase locking a gyrotron oscillator [2,3]. Two methods of phase locking are possible: adding a priming cavity to bunch the beam or direct injection at the output. A recently developed harmonic gyrotron is a promising new device for this transmitter.

The original plan was to generate detailed electrical, mechanical, and thermal designs that would form the basis for the construction and testing of a prototype gyroklystron. Due to lack of funding, construction of the prototype never occurred [1]. Nonetheless, the findings from

the design of the gyroklystron are applicable to the design of a gyrotron with a priming cavity; both devices are similar in performance and in construction.

II. Gyroklystron Amplifier

Table 1 is a list of important design specifications, and Fig. 1 is a schematic diagram of the proposed gyroklystron. The assembly consists of a magnetron injection gun (MIG), two buncher cavities, an output cavity, a collector and RF windows, and a mode converter. Not shown is the superconducting magnet assembly. The total assembly length, including cathode oil tank and mode converter at the output, is 400 cm. The design incorporates the following important technical and conceptual advances:

- (1) Space-charge effects cause increased beam velocity spread in the drift space that connects the gun to the circuit. Nevertheless, beam quality can be main-

tained by compensating for this effect and by minimizing the distance from the gun to the circuit.

- (2) Sufficient loading can be provided by buncher cavities made of graphite, replacing external cavity loads. This requires a circuit design that allows no more than 5 kW of heat dissipation in each buncher cavity.
- (3) Drift sections that are weakly cut off affect buncher cavity designs, appreciably altering beam loading and cavity resonant frequency. Linear stability requires unconventional buncher cavity geometries.
- (4) The amplifier can operate as a mode-locked oscillator, with output phase and frequency controlled at saturation even if the output cavity is unstable in the linear mode. Stability constraints can be relaxed while still allowing the device to perform as required.
- (5) The magnetic field profile in the taper section connecting the output cavity to the collector affects linear gain and nonlinear behavior (hysteresis effects of the amplifier).

A maximum design voltage gradient well below 100 kV/cm was chosen for reliable and long-life CW operation. After considerable effort had been expended to reduce the gradient on the high-voltage seal using previously designed guns, it was concluded that a major redesign was necessary. As a result, a new gun was designed with a larger outer diameter and new internal dimensions; this design resulted in reduction of the surface gradient at the gun anode tip to levels below 75 kV/cm.

Velocity spread, a most important parameter, affects the efficiency of gyrotrons. A MIG cathode usually operates in the temperature-limited mode, and the perpendicular velocity spread of electrons is assumed to be independent of the current as the beam drifts from the gun acceleration region to the first cavity. This assumption of laminar flow required a higher cathode angle, resulting in a cathode loading in excess of 10 A/cm². Because of high loading, this approach was abandoned.

Space-charge effects cause increasing perpendicular velocity spread in the gun drift space. This was a previously unexplored phenomenon in gun design. Electrons emitted from the upper part of the cathode continuously accelerate while those emitted from the lower part decelerate. For electrons with intercepting paths, the result is nonuniform acceleration. Many techniques can be used to reduce the effects of increased velocity spread. Reducing gun drift space length, for example, significantly reduces

the velocity spread; this reduction occurs because velocity spread enhancement is a cumulative effect. The maximum current loading tolerated in the magnet dictates the minimum length between gun and circuit. The addition of a bucking coil reduced velocity spread by 50 percent by modifying the field profile in the gun acceleration region.

The interactive circuit consists of two buncher cavities and an output cavity. Buncher cavities use the dominant TE_{11} mode while the output cavity uses the overmode TE_{12} . Between cavities, drift sections prevent internal feedback. A third buncher cavity to fully bunch the beam before entering the output cavity required large power dissipation, estimated at between 10 and 20 kW for a fundamental mode TE_{11} cavity, and consequently was not included. Final bunching of the electrons occurs in the output cavity to develop the required gain. An overmoded output cavity keeps ohmic wall dissipation to the reasonable levels of less than 1 kW/cm² and provides sufficient stored energy compatible with 400-kW output.

A special design feature consists of manufacturing the TE_{112} mode buncher cavities out of graphite in place of external loads. A cold test of the closed graphite cavity gave an ohmic Q of about 150. The drift sections shift resonance down by about 2 GHz and modify the mode profile. The external Q also approximately doubled with the addition of drift sections. The solution to the problem presented by weakly cutoff drift sections was to operate the buncher cavities in the TE_{112} mode. This solution provided stable operation and no mode competition with TE_{111} . Axial beam velocity spread did affect gain and output power. A 23-percent reduction in saturated output power occurred for a 15-percent axial velocity spread. Limiting axial velocity spread to less than 10 percent is necessary to maintain acceptable gain and output power.

A normalized magnetic field of 0.937 provides optimum performance, which is also close to the stability boundary. At low drive power levels, computer simulations predict a constant linear gain, up to between 3.5 and 4 watts of drive, where a large nonlinear increase in power occurs. The jump in power was 3 dB and the phase decreased by 60 deg. At this higher output power level, the phase again became constant, indicating amplifier operation. When reducing drive, the output did not follow the same path as when drive was increasing—the sudden jump in output power occurred at a lower drive power level. One conclusion drawn was that when operating under CW-saturated conditions, a well-defined line dividing stable and unstable operation does not exist. When the magnetic field was increased past the stability boundary, sufficient drive power locked the output, causing output phase and frequency

to follow input drive within a narrow bandwidth. For all practical purposes, the device performed as a phase-locked saturated amplifier.

If beam wave interaction occurs in the up-taper region of the output cavity (as it surely must unless the magnetic field suddenly decreases right outside the output cavity), the sudden rise in gain and the hysteresis effects of output power with drive power previously seen in a computer simulation disappear. Allowing interaction in the up-taper did not change saturated output power.

Table 2 summarizes sensitivities to key operating parameters. Computer simulations provided the values. Two parameters that are difficult to predict due to incomplete theory are harmonic emissions and the noise figure. Based on results of measurements made on other gyrotrons, second harmonic emission should be less than -20 dBc. As for the noise figure, an estimate of 83 dB due to shot noise in a temperature-limited beam is calculated. The noise contribution of electrostatic modes was not calculated.

III. Two-Cavity Phase-Locked Gyrotron

If operating at one power level is acceptable, then an oscillator could be phase locked to provide communication and radar signals. Victor Granatstein, Peter Latham, and Yuval Carmel made two feasibility studies of gyrotrons with two cavities, one developing a new gyrotron and the other modifying the Varian VGA-8003 by adding a priming cavity [2,3]. Use of a priming cavity allows bunching of the electrons in the beam. With the beam bunched, frequency and phase of the output can be controlled. Except for a decrease in output power from 400 kW to 200 kW and linear amplifier operation, the requirements remained practically unchanged from those previously used for the gyrokystron design. Both designs meet minimum requirements for this transmitter. The modified VGA-8003 meets the minimum requirement of 25 dB for carrier suppression but not the goal of 60 dB. Table 3 gives a comparison of results of these two feasibility studies. Table 4 summarizes the sensitivity to variations in operating parameters.

In the first feasibility study, the output cavity operates in the circularly polarized TE_{11} mode. With an operating output power capability of 330 kW, the calculated peak cavity surface dissipation was 2.5 kW/cm^2 , which is excessive. By limiting the output power to 260 kW, the peak wall loss equals 2 kW/cm^2 and average losses equal 1.2 kW/cm^2 . These preliminary values are quite high but actual values will probably be somewhat lower in the final design. Other major requirements are an efficiency of at least 30 percent with a bandwidth of better than

30 MHz. To operate from commercially available generators requires 35 dB of gain. Calculations show a buncher cavity with a Q of 22, a drift length of 8 cm, and an output cavity with a minimum diffractive Q of 44—all values that are easily achievable in a practical device. This design meets the major requirements.

The addition of a priming cavity and drift space to the Varian VGA-8003 free-running gyrotron oscillator is an alternative to the previous two-cavity phase-locked gyrotron. The VGA-8003 operates in the TE_{01} mode, which has much lower cavity wall losses. For the VGA-8003, an efficiency of 37 percent compared favorably to 26 percent estimated for the previous design. The VGA-8003 requires two mode converters, one to convert TE_{01} to TE_{11} and another to convert linear to right-hand polarization. The addition of a priming cavity and drift section lengthens the VGA-8003, which affects magnet design and efficiency as a result of increased velocity spread. TE_{01} mode cavities are overmoded and allow mode competition. Whatever the approach, a full design remains to be made. The conclusion reached is that use of a priming cavity is a possible means of phase locking a gyrotron to meet the requirements for a DSN Ka-band transmitter.

IV. Single-Cavity Phase-Locked Gyrotron

Use of a developed gyrotron without any modifications is particularly attractive because it minimizes tube development cost and risk. The only available gyrotron, the VGA-8003 manufactured for Lawrence Livermore by Varian Associates, requires a small shift in frequency from 35 to 34.5 GHz. The present tube is not tunable and the output frequency may be anywhere in a 100-MHz bandwidth for optimum performance. For DSN application it would be necessary to optimize performance close to 34.5 GHz.

Phase control in a single cavity gyrotron can be achieved by feeding a signal from a microwave reference source back into the gyrotron's output cavity. The relationship between phase-locking bandwidth and gain, which is exact for a single cavity, is

$$\frac{\delta\omega}{\omega_0} = \frac{1}{Q} \sqrt{\frac{1}{1+G}}$$

where ω_0 and Q are the cavity resonant frequency and quality factor, respectively. Here $\delta\omega$ is the locking bandwidth, which is 1/2 the total bandwidth, and G is the gain. When $G \gg 1$, the expression reduces to Adler's relation [4]

$$\frac{\delta\omega}{\omega_0} = \frac{1}{Q} \sqrt{\frac{P_i}{P_o}}$$

Typical values for Q and output power P_o from Table 5 require 3.3 kW of drive power P_i into the output cavity of the gyrotron for a locking bandwidth of 15 MHz. Allowing for transmission losses, the driver requires an output of 5 kW, but no tubes are available for this power level. Available drivers are 100 W for a travelling wave tube and 1 kW for a klystron. This reduces locking bandwidth to about ± 7 MHz for a 1-kW driver and ± 2.3 MHz for the 100-W driver. To obtain additional bandwidth the center frequency of oscillation may be shifted by varying the beam voltage.

The dependence of resonant frequency on beam voltage was derived by Victor Granatstein, Peter Latham, and Yuval Carmel in a report prepared for JPL [5]. The equation relating pitch angle to beam voltage is

$$\frac{V_b}{\alpha} \frac{d\alpha}{dV_b} = \frac{1 + \alpha^2}{2}$$

Using typical values for the VGA-8003 from Table 5,

$$\frac{d\omega}{\omega} = 0.026 \frac{dV_b}{V_b}$$

a change of 2.8 kV corresponds to a frequency change of 30 MHz. The stability limit of the DSN beam supply is 0.01 percent, which maintains resonant frequency to ± 100 kHz. The addition of a phase detector that compares the gyrotron's output with a reference signal could be used to generate an error signal to control the beam supply with a high-power pass tube. For example, frequency hopping of 8 MHz in less than 10 μ sec and frequency ramping of 78.4×10^6 MHz/sec requires the beam voltage to shift the resonant frequency for the injected signal to reacquire locking. Although this appears feasible, detailed design analysis needs to be completed. The frequency response of the closed-loop phase detector and modulator is a major limitation to controlling beam voltage.

The cathode voltage is the only practical parameter that can be used to shift the gyrotron frequency. The magnetic field controls gyrotron output frequency, but a fundamental problem exists in modulating this parameter. Magnetic field changes take too much time to diffuse through the conductive copper walls of the microwave cavity. A changing magnetic field induces eddy currents in the walls, and the flux inside the cavity is not fully changed

until the eddy currents have decayed. The time for the eddy currents to decay is roughly 10 msec.

The use of a developed gyrotron shifts the costs and risks to the development of a high-power circulator and a high-voltage pass-tube modulator. The circulator isolates the locking signal from the high-power signal out of the gyrotron. Preliminary calculations suggest that such a circulator is realizable in a beam waveguide transmission line [7]. A quasi-optical circulator placed in the beam waveguide directs the gyrotron signal to the antenna while directing the driver locking signal back to the output cavity of the gyrotron. Design of this component is critical to carrying out this phase-locking approach.

V. Second Harmonic Gyrotron

The University of Maryland Laboratory of Plasma Research recently acquired a second harmonic free-running gyrotron oscillator from the People's Republic of China. That gyrotron has produced over 200-kW of output power at 35 GHz with an efficiency of 35 percent. Its success is largely due to the use of a special complex cavity [6]. Dr. H. Z. Guo, who invented this cavity, is now working at the University of Maryland. Phase-locking of this gyrotron is feasible using the same techniques previously discussed in this article for the VGA-8003. Some advantages of this device are especially attractive. Use of the second harmonic reduces the cyclotron frequency by a factor of two; consequently, the magnetic field strength is also reduced by two. The magnetic field for a 35-GHz gyrotron requires nominally 13,000 gauss, which can be generated only with a superconducting magnet. For a second harmonic gyrotron, a magnetic field of 6500 gauss is easily achieved with room temperature solenoid magnets or even permanent magnets.

When phase locking a second-harmonic gyrotron by means of a priming cavity, the locking signal frequency is 17.5 GHz. Drivers at this lower frequency are available with higher output power. Phase locking directly into the output cavity requires a signal at the fundamental cyclotron frequency of 35 GHz. Still, the advantage of a lower magnetic field is preserved. Figure 2 is a schematic of the proposed second harmonic gyrotron with a priming cavity. The output propagates in the TE_{03} mode, and conversion to TE_{11} can be done efficiently.

VI. Future Work

The feasibility tests of a phase-locked second-harmonic gyrotron are continuing under a joint effort by the University of Maryland Laboratory for Plasma Research and

JPL. Two methods are still being considered: direct injection at the output cavity and bunching the electrons in the beam. JPL has responsibility for developing a quasi-optical circulator. The University of Maryland has

a second-harmonic gyrotron from China and is developing a modulator to pulse this tube. The goal is to show the capability of phase locking a gyrotron and to generate specifications for a Ka-band transmitter.

References

- [1] M. Caplan, J. Nelson, A. Salop, H. Jory, R. Bier, B. Stockwell, P. Reysner, R. Garcia, A. Nordquist, and M. Mizuhara, *Design of a 400 KW Ka Band Gyroklystron Amplifier and Magnet Assembly, Final Report - Phase II*, Palo Alto, California: Varian Associates, Inc., July 1986.
- [2] P. E. Latham, V. L. Granatstein, and Y. Carmel, *Design of a High-Gain, Wide-Bandwidth, Phase-Locked Ka-Band, 200 kW Gyrotron for Use in a Deep Space Communication and Radar System*, Silver Spring, Maryland: V. L. Granatstein, December 1989.
- [3] P. E. Latham, V. L. Granatstein, and Y. Carmel, *Study of Phase Locking of the Varian VGA-8003 Gyrotron Oscillator by Adding a Priming Cavity*, Silver Spring, Maryland: V. L. Granatstein, November 1990.
- [4] R. Adler, "A Study of Locking Phenomena in Oscillators," *Proceedings of the IEEE*, vol. 61, pp. 1380-1385, October 1973.
- [5] P. E. Latham, V. L. Granatstein, and Y. Carmel, *Study of Phase-locking of a Single-Cavity Gyrotron Oscillator (Varian Type VGA 8003) for Application in the Deep Space Network*, Silver Spring, Maryland: V. L. Granatstein, October 1990.
- [6] H. Guo, D. S. Wu, G. Liu, H. Miao, S. Z. Qian, and W. Z. Qin, "Special Complex Open-Cavity and Low-Magnetic-Field High-Power Gyrotron," *IEEE Trans. Plasma Sci.*, vol. 18, no. 3, pp. 326-333, June 1990.
- [7] T. Veruttipong, J. R. Withington, V. Galindo-Israel, W. Imbriale, and D. A. Bathker, "Design Considerations for Beam Waveguide in the NASA Deep Space Network," *IEEE Transactions on Antennas and Propagation*, vol. 36, no. 12, pp. 1779-1787, December 1988.

Table 1. Design specifications.

Parameter	Requirement
Center frequency, GHz	34.5
Instantaneous bandwidth, MHz	34 (30 for gyrotron)
Saturated gain, dB	50 (minimum 35)
RF saturated output power, kW	400 CW (minimum 200)
Harmonic power	TBD
Phase switching, 180 deg, nsec	400
Out of band power, percent	<1 at 0.1 Hz from carrier
Amplitude modulation, dBc	-40 (1 Hz to 1000 Hz)
Group delay dispersion, psec	1.2 (± 6 MHz)
Frequency hopping, MHz	8 in 10 μ sec
Frequency ramping, Hz/nsec	0.04
Carrier suppression, dB	25 with 60 goal
RF output mode	TE_{11} circular polarized
Extraneous modes, dB	-15
Ellipticity, dB	<1
Stability, output power, dB	0.1
DC beam power supply, MW	1.1 maximum
DC beam voltage, kV	100 maximum
Efficiency, percent	40 (minimum 30)
Duty	CW (up to 4-hr pulse)

Table 2. Gyrokystron pushing factors.

Parameter	Phase	Power
Cathode voltage	0.01 deg/volt	0.0013 dB/volt
Filament voltage	90 deg/volt	1.45 dB/volt
Linear drive power	8 deg/dB	0.98 dB/dB
Saturated drive power	5 deg/dB	0.029 dB/dB
Circuit field	0.98 deg/gauss	0.00615 dB/gauss
Gun field	26 deg/gauss	0.74 dB/gauss
Body coolant temperature	0.7 deg/deg C	0.003 dB/deg C
Collector coolant temperature	1 deg/deg C	~ 0

Table 3. Two-cavity gyrotron expected transmitter performance.

Parameter	New design	Modified VGA-8003
Center frequency, GHz	34.5	35
Instantaneous bandwidth, MHz	200	50
Duty	Continuous	Continuous
Output power, kW	260	230
RF output mode	TE_{11}	TE_{01}
Harmonics, dB	< -20	< -20 dB
Drive power, W	~100	~100
Efficiency, percent	>26	~37

Table 4. Two-cavity gyrotron pushing factors.

Parameter	New design	Modified VGA-8003
<u>Change in output power, percent</u>	0.5	0.4
<u>Change in beam current, percent</u>		
<u>Change in output power, percent</u>	0.8	1.0
<u>Change in beam voltage, percent</u>		
<u>Change in phase, deg</u>	3.3	1.8
<u>Change in beam current, percent</u>		
<u>Change in phase, deg</u>	20	41
<u>Change in beam voltage, percent</u>		
<u>Change in output power, dB</u>	0.2	0.3
<u>Change in injector power, dB</u>		
<u>Change in output phase, deg</u>	—	1.0
<u>Change in input power, dB</u>		
<u>Change in beam current, percent</u>	0.083 ^a	0.083 ^a
<u>Change in filament voltage, percent</u>		
<u>Change in output power, percent</u>	0.013	0.02
<u>Change in cavity temperature, deg C</u>		
<u>Change in output power, percent</u>	~0	~0
<u>Change in collector temperature, deg C</u>		
<u>Change in phase, percent</u>	0.08	0.03
<u>Change in cavity temperature, deg C</u>		
<u>Change in phase, percent</u>	1.5	1.5
<u>Change in collector temperature, deg C</u>		
<u>Change in phase, percent</u>	1.8	1.8
<u>Change in waveguide temperature, deg C</u>		
<u>Change in output power, percent</u>	~0	~0
<u>Change in waveguide temperature, deg C</u>		

^a Based on Varian Associates, Inc., estimates for VGT-8195 TWT Gun.

Table 5. Specifications for Varian gyrotron oscillator VGA-8003.

Parameter	Minimum	Maximum	Typical
Output cavity mode	—	—	TE_{01}
Frequency, GHz	34.5	35.5	35.0
Q	250	300	275
Beam voltage, kV	70	90	85
Beam current, A	4	10	7
$\alpha = v_{t0}/v_{z0}$	1.5	2.0	1.75
Magnetic field, kG	—	—	13.6
Beam radius, cm	—	—	0.254

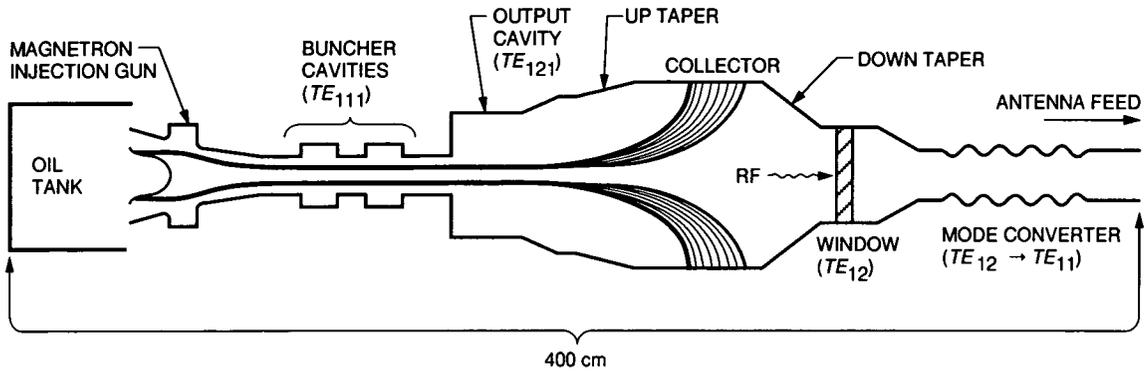


Fig. 1. Schematic of proposed gyrokystron amplifier.

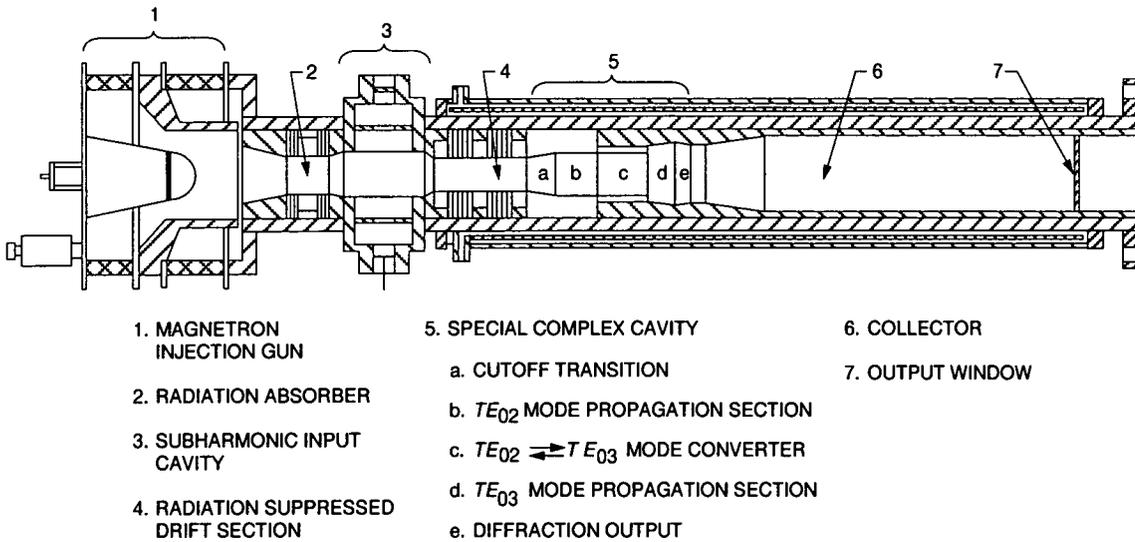


Fig. 2. Schematic design of a Ka-band, compact, phased-locked gyrotron with subharmonic injection.

500701

J5-33

167925

R-17

Experimental and Modal Verification of an Integral Equation Solution for a Thin-Walled Dichroic Plate With Cross-Shaped Holes

L. W. Epp

Radio Frequency and Microwave Subsystems Section

P. H. Stanton

Ground Antennas and Facilities Engineering Section

In order to add the capability of an X-band uplink onto the 70-m antenna, a new dichroic plate is needed to replace the Pyle-guide-shaped dichroic plate currently in use. The replacement dichroic plate must exhibit an additional passband at the new uplink frequency of 7.165 GHz, while still maintaining a passband at the existing downlink frequency of 8.425 GHz. Because of the wide frequency separation of these two passbands, conventional methods of designing air-filled dichroic plates exhibit grating lobe problems. A new method of solving this problem by using a dichroic plate with cross-shaped holes is presented and verified experimentally.

Two checks of the integral equation solution are described here. One is the comparison to a modal analysis for the limiting cross shape of a square hole. As a final check, a prototype dichroic plate with cross-shaped holes was built and measured.

I. Introduction

The unit cell shape of a thick frequency selective surface, or dichroic plate, is dependent on its frequency requirements. One aperture shape may be chosen to give wider bandwidths, and another may be chosen for sharper frequency roll-off. This is analogous to circuits where the need for differing frequency responses determines the circuit topology.

Acting as spatial frequency filters, dichroic plates are a critical component in providing the DSN spacecraft command and control uplinks as well as spacecraft downlinks.

Currently these dichroic plates separate the S-band at 2.0–2.32 GHz from the X-band at 8.4–8.45 GHz. But new spacecraft communication requirements are also calling for an uplink frequency at 7.165 GHz. Even more frequency demands are being placed on the dichroic plates being designed and built for the new beam-waveguide antennas. Future spacecraft will require these dichroic plates to effectively separate Ka-band frequencies in the 31–35 GHz range.

The requirements for these dichroic plate surfaces are low transmission loss of <0.1 dB at high power levels, and a minimal relative phase shift between polarizations for

circular polarization transmission. Even in the past, these demanding requirements caused the dichroic plate designs to change from circular apertures to Pyle apertures.¹ Currently the dichroic plate used on the 70-m antenna in the DSN consists of Pyle-shaped apertures/waveguides, see Fig. 1. Originally this plate began with a design for normal incidence using circular holes. When the circular-hole plate was measured at the desired 30-deg incident angle, the resonant frequencies for the two orthogonal polarizations had been shifted. As a first correction, the calculated electrical lengths of a Pyle guide at 8.425 GHz were set equal to the electrical lengths of the circular guide at the two measured resonance frequencies. The plate was then rebuilt with Pyle-shaped apertures.

More current work has shown the successful demonstration of design techniques for straight, rectangular apertures at an incident angle of 30 deg [1]. Rectangular apertures offer a slight design advantage in terms of packing density over the Pyle aperture plate, and therefore may give improved performance at wider angles of incidence. Wide-angle performance is a consideration in practice where the dichroic plate is excited at angles other than the nominal 30 deg due to feed horn effects. For example, a 22-dB horn pattern can have considerable power in other plane waves incident in the range of 30 deg \pm 15 deg [2].

Grating lobes are the excitation of higher-order Floquet harmonics that transmit and reflect power at angles other than the incident angle. They become a consideration when the bandwidth required to include the new frequency of 7.165 GHz conflicts with the desired incident angle of 30 deg and the operation at the higher frequency band of about 8.425 GHz. In order to design a suitable dichroic plate with rectangular or Pyle-guide apertures, the aperture dimensions must be large enough to be above cutoff for the lower frequency passband. Even when using minimal wall thickness, this creates a unit cell capable of supporting grating lobes at the higher frequency passband. In addition, although dual passbands can be demonstrated with a rectangular aperture, there is not enough independent control to align the transverse electric/transverse magnetic (TE/TM) passbands as shown in Fig. 2. In contrast, Fig. 2 shows that the cross shape's more independent passband control makes dual passbands possible.

Because of the cross shape's increased packing density, grating lobes are pushed out to wider angles of incidence.

¹P. D. Potter, *Improved Dichroic Reflector Design for the 64-Meter Antenna S- and X-Band Feed Systems*, JPL Technical Report 32-1526 (internal document), vol. 19, Jet Propulsion Laboratory, Pasadena, California, pp. 55-62, February 15, 1974.

One method, proposed by Otoshi and Franco [3], of eliminating the grating lobe problems of the Pyle-shaped apertures was to fill them with Teflon dielectric material. The dielectric constant of 2.04 would allow the holes to be smaller, and therefore to be packed more closely together. Somewhat difficult to manufacture, the Teflon dielectric plugs were placed in liquid nitrogen and then allowed to expand into the holes of the dichroic plate, with an unknown effect on the dielectric constant. Their computer analysis showed that the dielectric constant needed to be known to an accuracy of better than 5 percent. Any change in the dielectric constant affects the desired plates thickness.

Prototypes for this type of dielectric-filled dichroic plate showed that if fabrication difficulties could be overcome, the loss tangent of the Teflon contributed an additional 1.3 K of noise temperature. With uplink frequency powers in excess of 100 kW, the question of arcing at improperly filled interfaces due to the high power levels was raised by Otoshi and Franco [3]. Lossy dielectric materials could also pose a power dissipation problem. Thus it is clearly desirable to find an air-filled dichroic plate solution that would be superior in power dissipation, power handling, and noise temperature considerations.

II. The Dichroic Plate With Cross-Shaped Holes

In order to diplex the S- and X-bands (the downlink at 8.425 GHz and the new uplink of 7.165 GHz), cross-shaped apertures in a dichroic plate have been investigated. Figure 3 shows the geometry and gives the prototype dimensions. The analysis consisted of an integral equation method of moments solution [4,5]. This method solves for the currents along the walls inside the cross, including those along the hole depth in the z direction.

Integral equation methods can be extended to handle wall thickness, but the model used here consists of a thin-walled cross on a skewed grid. Validation for using the thin-wall model comes from the need to minimize grating lobe behavior at the higher frequency band of operation. Therefore, the unit cell size must be kept to a minimum, which implies that the wall thickness too should be minimized.

A. Modal Analysis of the Limiting Shape of the Square Hole

The unit cell of the dichroic plate with cross-shaped holes on an irregular grid consists of six vertical plates as shown in Fig. 4. The current continuity between unit cells is enforced by the currents J_1 and J_2 shown. The

flexibility of this analysis is that any other shape that can be constructed of vertical plates can also be analyzed, such as the honeycomb [5].

In the modal analysis of a dichroic plate problem, this flexibility is not present, since the modes of the waveguides (repeated periodically in the dichroic plate) change as the waveguide shape changes. Modal analysis is further complicated for waveguide shapes such as the cross-shaped waveguide, where simple analytical formulations of the waveguide modes do not exist. By expansion in suitable eigenfunctions and numerical solution using the Ritz-Galerkin method, Lin solved for the modes of the cross-shaped waveguide [6]. Another method for finding the modes of the cross includes the finite element method as demonstrated by Stalzer et al. [7]. Using the modes found by Lin [6], a modal analysis of a dichroic plate with cross-shaped holes on a regular grid was developed by Compton et al. [8].

To check the accuracy of the integral equation solution in this work for an irregular or skewed grid, a limiting case of the cross waveguide was chosen. By setting $b_x = 0$, the analysis becomes that of a square hole, which can be compared to the modal analysis of Chen [1]. A good estimate of the error from the thin-wall approximation can be made from the modal analysis of the square hole.

The integral equation solution uses a first-order spatial domain Green's function evaluation [4,5,9]. The Green's function for the electric field integral equation (EFIE) is given by

$$G_p = (G_p - G^a) + F^{-1} (\tilde{G}^a) \quad (1)$$

where the spatial domain Green's function is given by

$$G_p(\vec{r}_0, \vec{r}') = \sum_{m=-\infty}^{m=+\infty} \sum_{n=-\infty}^{n=+\infty} e^{-j\vec{k}_0 \cdot \vec{\rho}_{mn}} \frac{e^{-jk_0 \sqrt{|\vec{\rho}_0 - \vec{\rho}' - \vec{\rho}_{mn}|^2 + (z_0 - z')^2}}}{4\pi \sqrt{|\vec{\rho}_0 - \vec{\rho}' - \vec{\rho}_{mn}|^2 + (z_0 - z')^2}} \quad (2)$$

The spatial lattice translation vector needed is defined by

$$\vec{\rho}_{mn} = (nd \cos \Omega) \hat{x} + (mc + nd \sin \Omega) \hat{y} \quad (3)$$

and Fig. 4 defines the distances c and d as well as the skew angle Ω . The corresponding asymptotic Green's function used was

$$G_p(\vec{r}_0, \vec{r}') = \sum_{m=-\infty}^{m=+\infty} \sum_{n=-\infty}^{n=+\infty} e^{-j\vec{k}_0 \cdot \vec{\rho}_{mn}} \frac{e^{-jk_0 \sqrt{|\vec{\rho}_0 - \vec{\rho}' - \vec{\rho}_{mn}|^2 + (|z_0 - z'| + cCA)^2}}}{4\pi \sqrt{|\vec{\rho}_0 - \vec{\rho}' - \vec{\rho}_{mn}|^2 + (|z_0 - z'| + cCA)^2}} \quad (4)$$

The cell area needed in the asymptotic form of spectral domain Green's function

$$G_p(\vec{r}_0, \vec{r}') = \frac{1}{\text{cell area}} \sum_{m=-\infty}^{m=+\infty} \sum_{n=-\infty}^{n=+\infty} \frac{e^{-j\gamma(|z_0 - z'| + cCA)} e^{j(\vec{K}_{mn} - \vec{k}_0) \cdot (\vec{\rho}_0 - \vec{\rho}')}}{2j\gamma} \quad (5)$$

is the area enclosed inside the unit cell shown in Fig. 4 where

$$\vec{K}_{mn} = 2\pi \left[\frac{n}{d \cos \Omega} - \frac{m \sin \Omega}{c \cos \Omega} \right] \hat{x} + \frac{2\pi m}{c} \hat{y} \quad (6)$$

When the basis and testing functions used in the EFIE method of moments (MOM) solution do not overlap in the z direction, the spatial acceleration shown in Eq. (1) is not used. Instead, it is replaced by the spectral domain Green's function found from Eq. (5) by letting $cCA = 0$.

From the spectral domain Green's function, it is seen that the Floquet harmonics are described by plane wave spectral components propagating according to

$$\gamma = \begin{cases} \sqrt{k_0^2 - \beta_{xmn}^2 - \beta_{ym}^2} & k_0^2 > \beta_{xmn}^2 + \beta_{ym}^2 \\ -j\sqrt{\beta_{xmn}^2 + \beta_{ym}^2 - k_0^2} & \beta_{xmn}^2 + \beta_{ym}^2 > k_0^2 \end{cases} \quad (7)$$

and

$$\beta_{xmn} = (\vec{K}_{mn} - \vec{k}_0) \cdot \hat{x} = 2\pi \left[\frac{n}{d \cos \Omega} - \frac{m \sin \Omega}{c \cos \Omega} \right] - k_x \quad (8)$$

$$\beta_{ym} = (\vec{K}_{mn} - \vec{k}_0) \cdot \hat{y} = \frac{2\pi m}{c} - k_y \quad (9)$$

The phase shift from cell to cell is determined by the phase shifts of the incident plane wave

$$\begin{aligned} k_x &= -k_0 \sin \theta_i \cos \phi_i \\ k_y &= -k_0 \sin \theta_i \sin \phi_i \\ k_z &= -k_0 \cos \theta_i \end{aligned} \quad (10)$$

The modal analysis uses rectangular waveguide modes inside the rectangular holes, matched to Floquet modes at the apertures. Formulations for the Floquet modes can be found in [2] and will not be included here. But of importance is the number of Floquet modes chosen in relationship to the number of rectangular waveguide modes. In comparison to the integral equation method, it was necessary for convergence to increase the number of waveguide modes. If $mmax$ and $nmax$ represent the highest waveguide mode indices used, the Floquet modes m and n were chosen so that

$$\begin{aligned} \beta_{xmn} &\approx \frac{(mmax)\pi}{H_x} \\ \beta_{ym} &\approx \frac{(nmax)\pi}{H_y} \end{aligned} \quad (11)$$

and then incremented by three.

Figure 5 shows the geometry of the square-hole test case. For initial comparison to the thin-wall integral equation solution, the dimensions were $H_x = H_y = S_x = S_y = 2.59$ cm. The definitions of the two incident polarizations used are shown in Fig. 6. Figures 7 and 8 show the comparison of the two solutions for normal incidence. The TE polarization shows a maximum error of 0.66 percent and the TM polarization shows twice as much error, or 1.33 percent. For completeness, the phase comparisons of these two test cases are shown in Figs. 9 and 10.

The modal solutions convergence is affected by the number of rectangular waveguide modes used. In the analysis of Figs. 7–10, the number of rectangular waveguide modes used was 220. The convergence of the resonant locations for the modal analysis is shown in Fig. 11. The number of basis functions used along the depth of the holes for the integral equation solution was 30. It too is sensitive to the number of basis functions used, as shown in Fig. 12. Besides the discretization error, other possible sources of error exist, and it should be noted that no attempt was made to match the proper edge behavior for a thin wall in the modal solution.

For plate rigidity, the prototype plate was designed with a wall thickness of 0.041 cm. Therefore, an inherent error in this analysis is the effect of this wall thickness. A good estimate of this error can be made from the modal analysis of the square hole. Figures 13 and 14 show the modal analysis of the square hole where the wall thickness of the prototype is changed from a thin wall to one of a width of 0.041 cm. Both Figs. 13 and 14 were made for normal incidence, which differs from the 30-deg incident angle of the prototype plate. Note that Figs. 13 and 14 show the analysis when the wall thickness has been added in such a way as not to change the periodicity of the unit cell, $H_x = H_y = 2.55$ cm and $S_x = S_y = 2.59$ cm. The thin wall solution has dimensions $H_x = H_y = S_x = S_y = 2.59$ cm. The error for the TE polarization is 0.44 percent and the error for the TM polarization is 0.55 percent.

When the wall thickness is added in such a way as to change the periodicity of the structure, a more dramatic shift in the resonant location of each polarization occurs. Figures 15 and 16 show the results when $H_x = H_y = 2.59$ cm and $S_x = S_y = 2.63$ cm. Here the error increases from 0.44 percent to 1.33 percent for the TE polarization and from 0.55 percent to 1.44 percent for the TM polarization. This implies that to minimize the effect of the wall thickness in the integral equation solution for the cross shape, it should be added in such a manner as not to change the periodicity. Note here that changing the periodicity of the

square-hole shape significantly lowered the frequency of the resonant locations.

B. Integral Equation Design of Prototype S-/X-Band Dichroic Plate With Cross-Shaped Holes

For the prescribed angle of incidence $\theta = 30$ deg, minimum grating lobes are exhibited by the cross shape when $\phi = 0.0$ deg. The TE polarization is more affected by the cross-shape dimensions of a_x and b_x . Generally, the cross shape can be thought of as two orthogonal rectangular waveguides, one for each polarization. So the x arm of the cross can be thought of as a rectangular waveguide that transmits the TE polarization. The TM polarization is more affected by the y arm of the cross shape. Of course, because of the exact shape of the cross waveguide, one polarization will couple to the other. But because the two polarizations are only weakly coupled, a useful design that optimizes the performance of both bands can be made.

The results for an air-filled, cross-shaped aperture with dimensions $a_x = 2.65$ cm, $b_x = 0.6$ cm, $a_y = 2.76$ cm, $b_y = 0.7125$ cm, an aperture depth of 3.5 cm, a skewed grid angle of $\Omega = 33.95$ deg, and an incident angle of $\theta = 30$ deg are shown in Fig. 17. Transmission results are shown in Fig. 18 and both results are summarized in Table 1. These results and dimensions were used as a basis for the prototype dimensions. The number of basis functions used in discretizing the depth was 12.

Because cross-shaped apertures pack more tightly than rectangular apertures, it is possible to obtain a dichroic plate with the desired dual X-band passbands grating lobe free. In contrast, this could not be done with the air-filled rectangular shape without the introduction of grating lobes because the minimum unit cell size is the size of the aperture. The more tightly packed cross-shaped elements overlap in the x direction, reducing the unit cell size in the x direction to less than the length of the cross arm a_x .

This design also gives a minimum phase shift between the two transmitted polarizations at the more critical downlink frequency of 8.425 GHz. The phase difference at the downlink frequency is small, 2.93 deg, but the trade-off is a phase difference at the uplink of 16.9 deg, as shown in Fig. 19. If more phase shift is allowed at the uplink frequency, which could be compensated for by a polarizer, a slightly different design may allow for slightly more bandwidth.

III. Grating Lobe Study

For a dichroic plate, a grating lobe is due to the propagation of a higher-order Floquet harmonic other than the

zeroth order $m = n = 0$ harmonic. The zeroth-order Floquet harmonic is usually of interest because it satisfies Snell's law for reflection and transmission. In other words, the angles of transmission and reflection are equal to the angle of incidence. The current design dimensions give a plate that will be grating lobe free to $\theta = 44$ deg at $\phi = 0$ deg (i.e., from the more tightly packed x direction) and frequency of 8.625 GHz. Figure 20 shows the results of the preliminary cross-shape design with the incident angles of a 22-dB feed horn transformed to the coordinates of the dichroic plate as done by Chen [2]. The curve marked maximum theta (Θ) is the maximum incident angle that a plane wave can strike this dichroic plate before grating lobes are produced.

IV. Experimental Verification and Analytical Improvements

The aluminum test plate was constructed using wire electrical discharge machining from an aluminum sample of 39.4 cm by 39.4 cm. The overall plate consisted of 213 holes, covering an internal area of approximately 36.2 cm by 36.2 cm. Figure 21 shows the mechanical drawing for the prototype plate. Note that the dimensions are slight modifications of the above preliminary design to allow for an addition of 0.041 cm of wall thickness. Figure 22 shows the partially completed prototype, and Fig. 23 shows a close-up of the cross-shaped holes.

The measured interior dimensions of the cross-shaped waveguide composing the prototype were measured to be $a_x' = 2.64744$ cm, $b_x' = 0.60008$ cm, $a_y' = 2.75742$ cm, $b_y' = 0.70021$ cm, and the hole depth was 3.35 cm. A comparison of these measured values with those given in Fig. 21 shows that a tight manufacturing tolerance was held as all internal cross dimensions were decreased consistently by 0.0254 cm. This very slight shortening of all interior dimensions was reflected in wall thickness and not in the distance between the centers of the holes. The only fabrication error was a slight shift in the centers of one row of cross-shaped holes of approximately 0.0254 cm.

Figure 24 shows the experimental results, while Figs. 25 and 26 show the initial computed results before making improvements. The computed results use the interior dimensions of the cross-shaped waveguide and the incident angle of $\theta = 30$ deg and $\phi = 0.0$ deg. The number of basis functions used to discretize the depth was 12, which was clearly insufficient.

As mentioned in the comparison of the modal analysis, it is necessary to use the exterior cross dimensions so

that the periodicity of the theoretical model is the same as in the prototype. Adjusting for the wall thickness in this manner, the measured values give $a_x = 2.69047$ cm, $b_x = 0.60008$ cm, $a_y = 2.80004$ cm, and $b_y = 0.70021$ cm. The skew grid angle of the theoretical calculations becomes the correct $\Omega = 33.81$ deg of the prototype. After making accuracy improvements to the theoretical calculations along with this adjustment of dimensions, the comparison to computed results is shown in Figs. 27 and 28.

Improvements to the theoretical analysis consisted of two parts, one geometric and one analytic. The geometric improvements came from realizing that there is a correct way to adjust the theoretical dimensions that were determined by comparison with the limiting case of the square-hole shape. This was demonstrated in Figs. 15 and 16.

One analytical improvement consisted of increasing the sampling density of the currents along the wall. This allows the current to be modeled more accurately in rapidly varying areas like the edges of the plate and around the 90-deg bends on the inside of the cross. This proved to be more important in discretizing the depth of the dichroic plate, as the current discretization uses a pulse approximation to triangle basis functions that are sampled at a point just inside the length of the dichroic plate. This leads to a small length error which can be compensated for, and decreases as the sampling density is increased.

Increasing the sampling density also allows the resonant frequencies to converge to their true value. This was demonstrated by Fig. 12, where it was seen that the resonant frequency locations were seen to be still converging for 27 basis functions per wavelength. In the case of the cross shape, the number of basis functions along the length of the cross was varied until the resonant frequencies changed to less than 0.01 GHz. For the results of Figs. 27 and 28, a total of 30 basis functions was used to discretize the depth, except immediately around the resonant loca-

tions where, for increased accuracy, 36 basis functions were used.

The second set of analytical improvements came from replacing the approximation to the periodic spatial domain Green's function being used in the first term of Eq. (1). Unfortunately this approximation was implemented in order to make the solution more computer efficient [5]. This approximation amounts to limiting the mutual coupling effects of each cross hole to its nearest neighbors and sums the contributions for $m, n = -1, 0, 1$. This was replaced with a summation that was then summed to the same specified accuracy as the spectral Green's function term. Again these parameters were increased until the resonant frequencies changed to less than 0.01 GHz. This doubled the computation time.

A summary of the error for the cross-shaped test plate is given in Table 2.

The conductivity loss has not yet been calculated. An advantage of this method is that the conductivity loss is present in the theoretical model. In the past, conductivity loss has been estimated based on the assumption of a dominant TE_{10} (square) waveguide mode [2]. This loss can be calculated as a function of frequency, and its impact on the transmission loss can be found.

V. Conclusions

A detailed analysis of a cross-shaped dichroic plate has been presented that shows dual passbands in the transmission region. As a check on the analysis used on the cross shape, it was compared to the modal analysis for the special case of the square hole. The modal analysis of the square hole was used to determine the effect of the modelling error of the thin-wall approximation used in the cross shape. As a final check on the analysis for the cross shape, the construction and testing of a cross-shaped dichroic plate prototype have been described.

Acknowledgments

The authors thank Jacqueline C. Chen for providing the modal analysis programs and Harry Reilly, Jr., for taking care of all the manufacturing details of the prototype plate. The authors also acknowledge the integral equation programs provided by Roy E. Jorgenson of Sandia National Laboratories and Professor Raj Mittra of the University of Illinois, which formed the basis of the cross analysis, making this work possible.

References

- [1] J. C. Chen, "Analysis of a Thick Dichroic Plate With Rectangular Holes at Arbitrary Angles of Incidence," *The Telecommunications and Data Acquisition Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 9–16, February 15, 1991.
- [2] J. C. Chen, "X-/Ka-Band Dichroic Plate Design and Grating Lobe Study," *The Telecommunications and Data Acquisition Progress Report 42-105*, vol. January–March 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 21–30, May 15, 1991.
- [3] T. Y. Otoshi and M. M. Franco, "Dual Passband Dichroic Plate for X-Band," *The Telecommunications and Data Acquisition Progress Report 42-94*, vol. April–June 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 110–134, August 15, 1988.
- [4] R. E. Jorgenson, "Electromagnetic Scattering From a Structured Slab Comprised of Periodically Placed Resistive Cards," Ph.D. dissertation, University of Illinois, Urbana, Illinois, 1989.
- [5] R. E. Jorgenson and R. Mittra, "Scattering from structured slabs having two-dimensional periodicity," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 2, pp. 151–157, February 1991.
- [6] F. C. Lin, "Modal characteristics of crossed rectangular waveguides," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-25, no. 9, pp. 756–763, September 1977.
- [7] H. J. Stalzer, Jr., M. D. Greenman, and F. G. Willwerth, "Modes of crossed rectangular waveguide," *IEEE Transactions on Antennas and Propagation*, vol. AP-24, pp. 220–223, March 1976.
- [8] R. C. Compton, R. C. McPhedran, G. H. Derrick, and L. C. Botten, "Diffraction properties of a bandpass grid," *Infrared Phys.*, vol. 23, no. 5, pp. 239–245, 1983.
- [9] R. E. Jorgenson and R. Mittra, "Efficient calculation of the free-space periodic Green's function," *IEEE Transactions on Antennas and Propagation*, vol. 38, no. 5, pp. 633–642, May 1990.

Table 1. Preliminary results for a dichroic plate with cross-shaped holes.

Frequency, GHz	Reflection, dB		Transmission, dB	
	TE polarization	TM polarization	TE polarization	TM polarization
7.065	-19.96	-18.38	-0.04	-0.09
7.165 (uplink)	-28.33	-29.19	-0.003	-0.03
7.265	-16.53	-28.72	-0.09	-0.03
8.325	-23.55	-26.09	-0.02	-0.07
8.425 (downlink)	-30.68	-43.33	-0.003	-0.06
8.525	-19.49	-25.14	-0.05	-0.08
2.1	-0.00	-0.00	-40.23	-35.24
2.3	-0.00	-0.00	-39.08	-34.06

Table 2. Summary of error for the cross-shaped prototype.

Resonance location: Polarization frequency, GHz	Initial error: Computation and measurement, %	Final error: Computation and measurement, %	Total analytical improvement: Analysis and geometry, %
TE, 7.075	1.27	0.85	0.42
TM, 7.119	0.65	1.04	-0.39
TE, 8.197	2.71	0.22	2.49
TM, 8.186	2.92	0.48	2.44

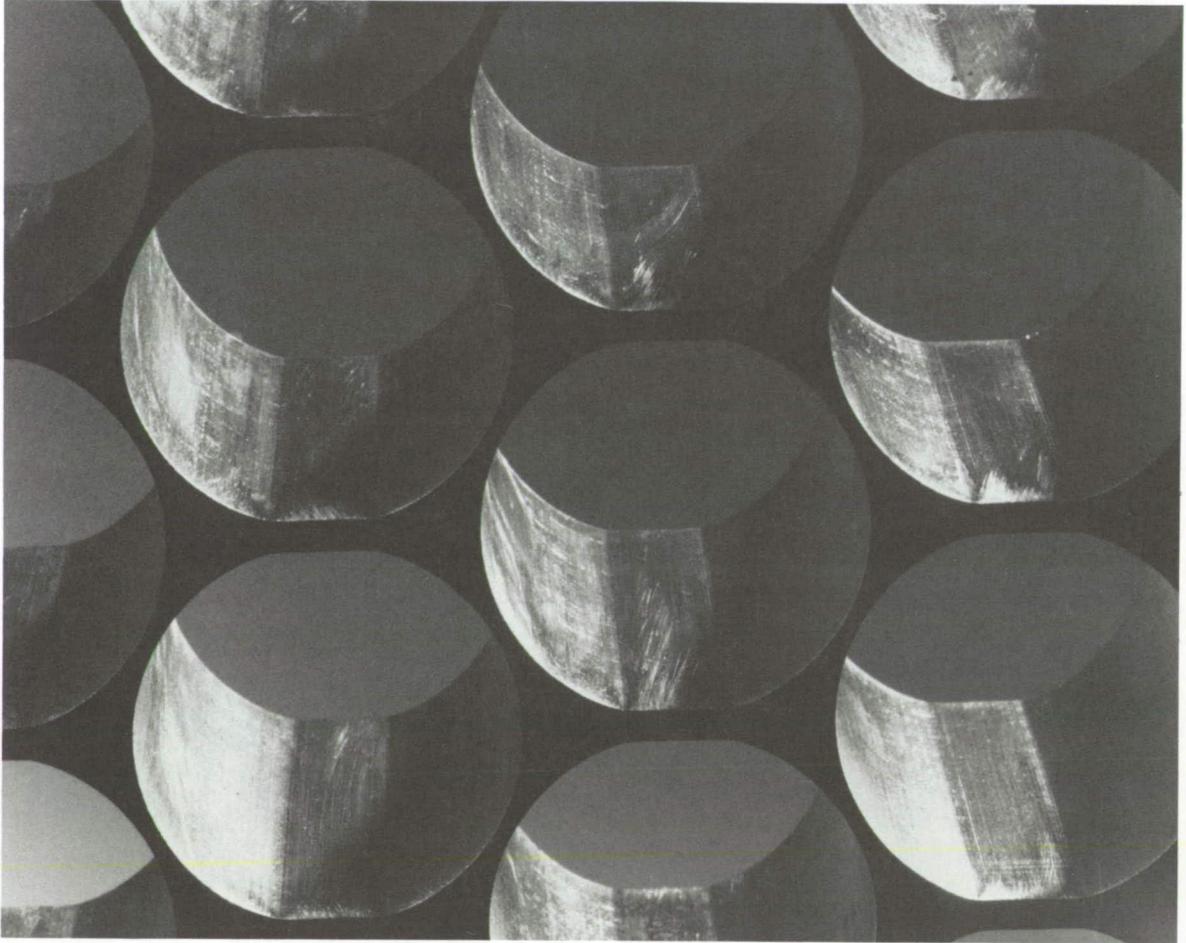


Fig. 1. Close-up of a dichroic plate consisting of Pyle apertures.

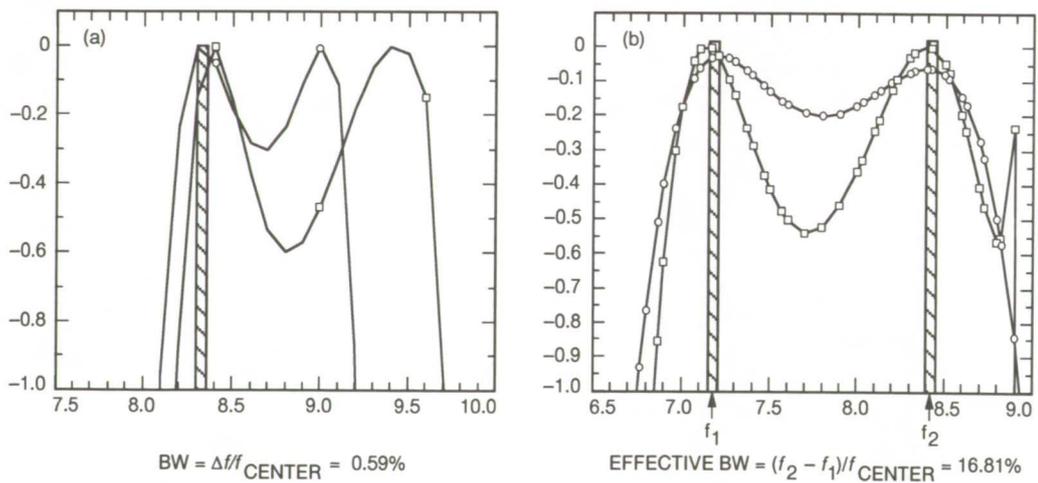
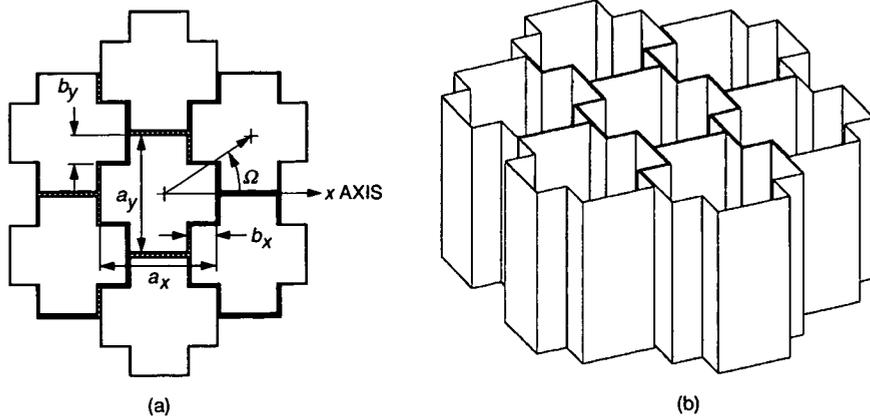


Fig. 2. Passbands of: (a) rectangular aperture, and (b) cross-shaped aperture.



PROTOTYPE DIMENSIONS: $a_x = 2.64744$ cm, $b_x = 0.60008$ cm, $a_y = 2.75742$ cm
 $b_y = 0.70021$, HOLE DEPTH = 3.35 cm
 $a_x + \text{WALL THICKNESS} = 2.69047$ cm
 $a_y + \text{WALL THICKNESS} = 2.80004$ cm

Fig. 3. Cross-shaped S/X-band dichroic plate—prototype dimensions: (a) top view, and (b) side view.

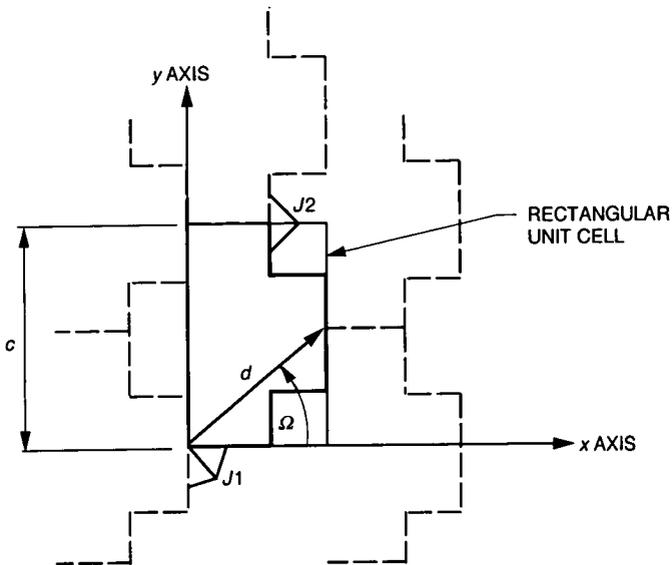


Fig. 4. Unit cell for the cross-shaped dichroic plate.

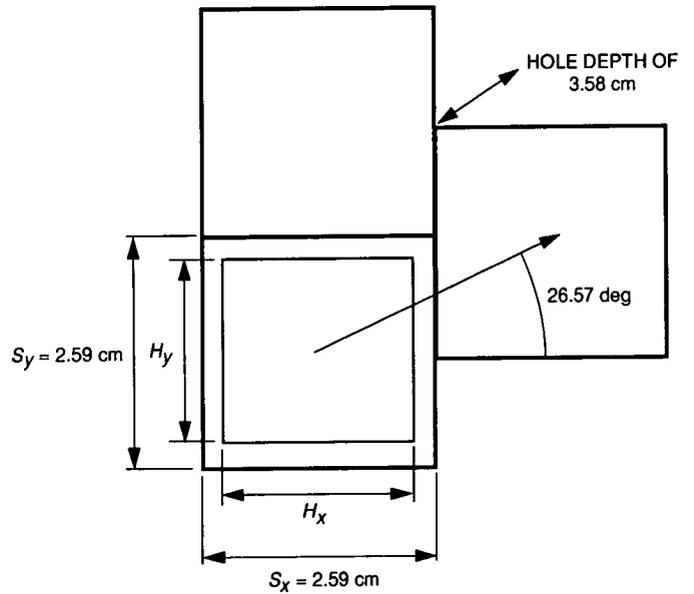


Fig. 5. Square-hole geometry.

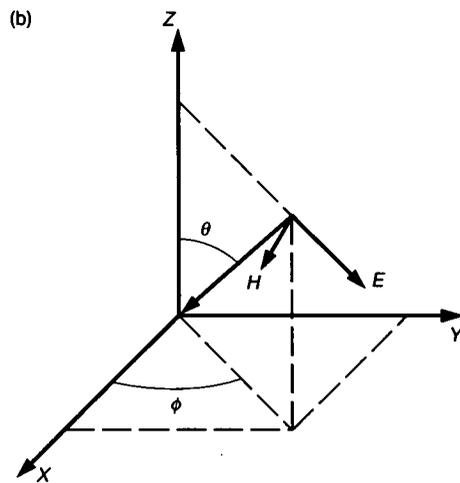
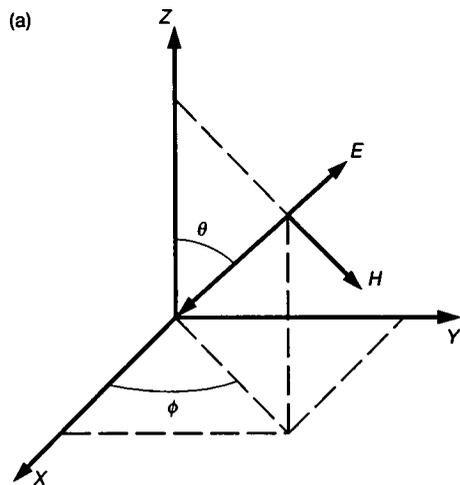


Fig. 6. Definition of two incident plane waves striking a surface in the x - y plane at the angle of θ and ϕ : (a) TE, and (b) TM.

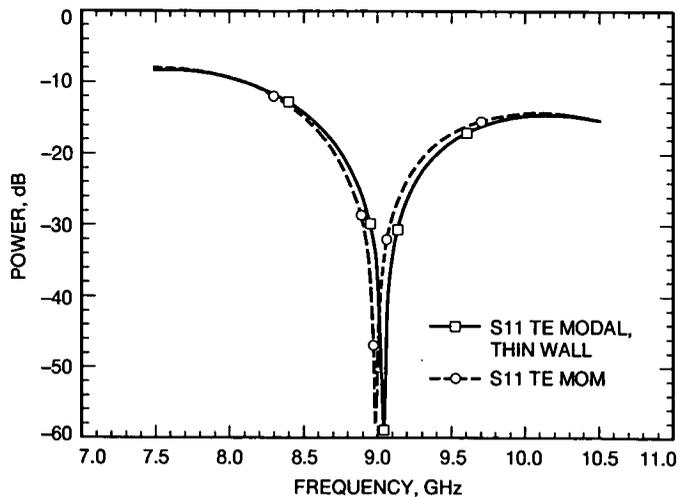


Fig. 7. Comparison of modal analysis to integral equation solution, TE polarization.

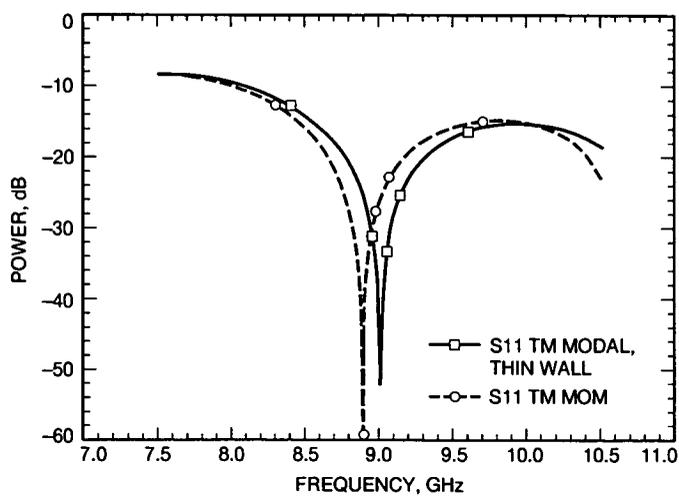


Fig. 8. Comparison of modal analysis to integral equation solution, TM polarization.

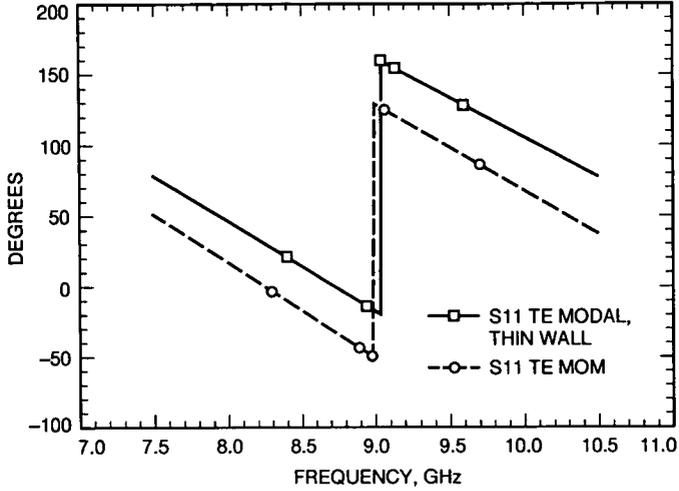


Fig. 9. Comparison of modal analysis to integral equation solution, reflected phase TE.

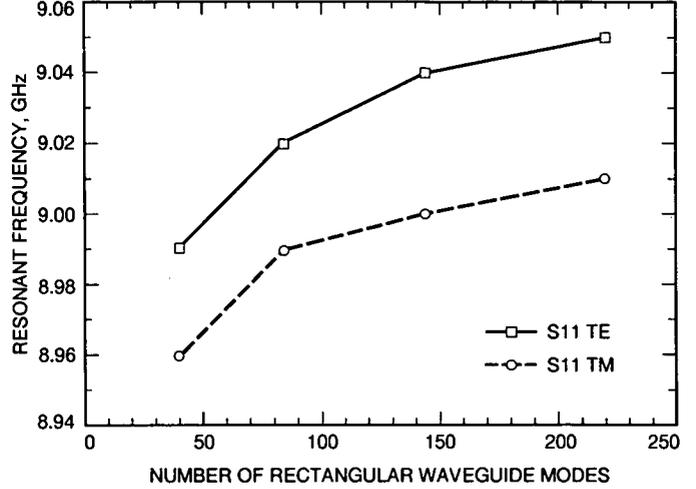


Fig. 11. Resonant frequency convergence in the modal solution of the square hole with thin walls.

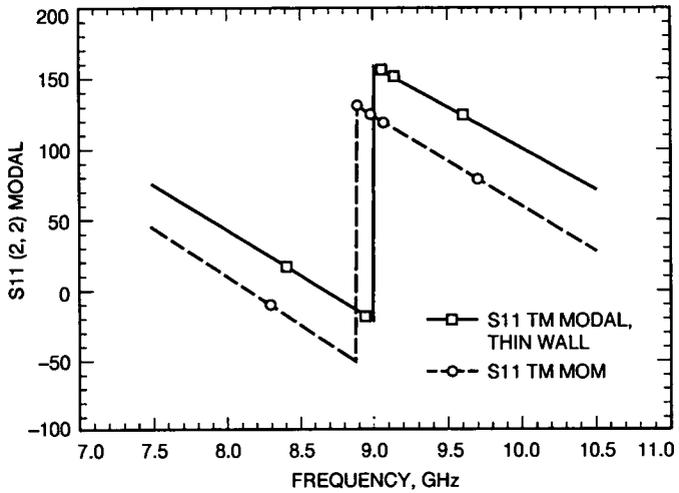


Fig. 10. Comparison of modal analysis to integral equation solution, reflected phase TM.

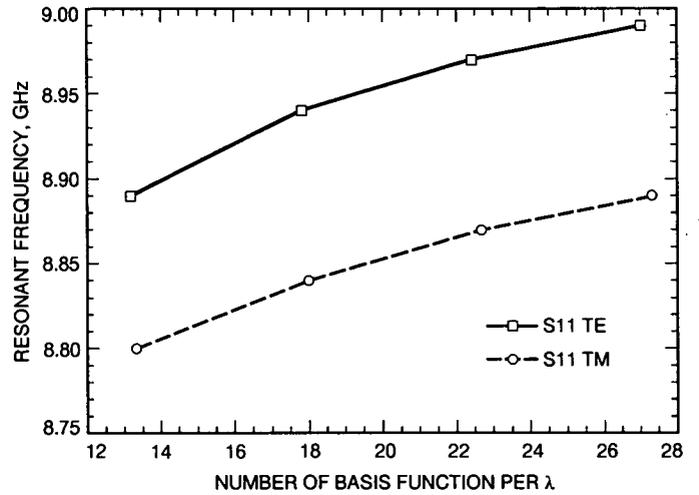


Fig. 12. Resonant frequency convergence in the method of moments solution of the square hole with thin walls.

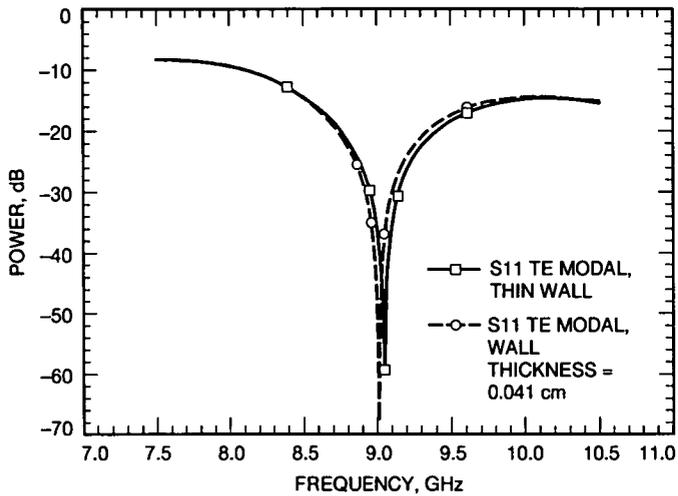


Fig. 13. Comparison of effect of wall thickness on the TE reflection coefficient.

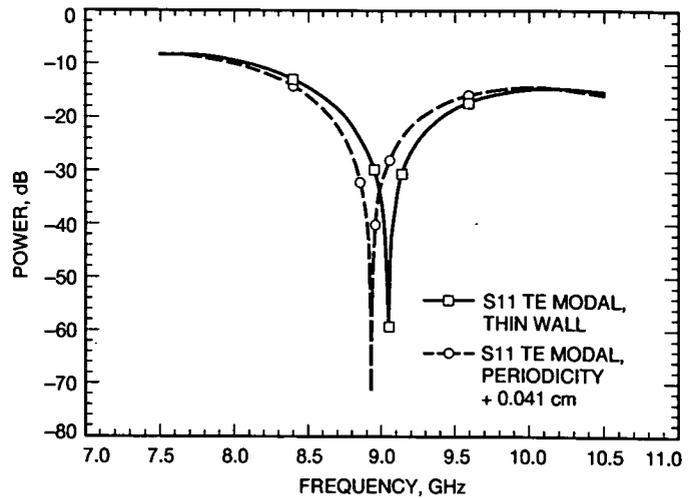


Fig. 15. Effect of periodicity change on the TE reflection coefficient.

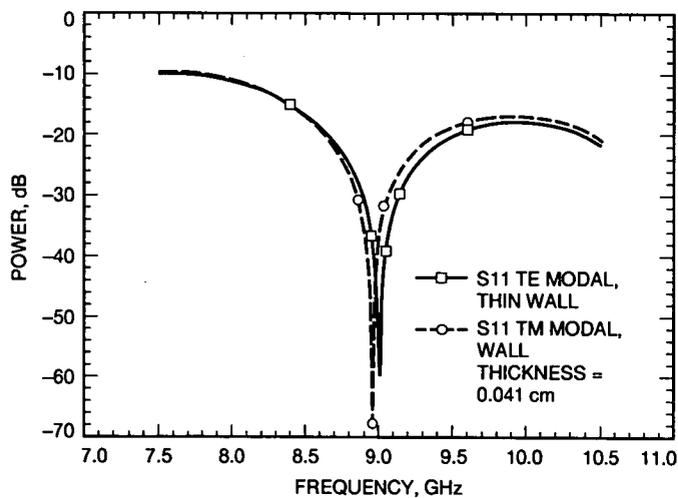


Fig. 14. Comparison of effect of wall thickness on the TM reflection coefficient.

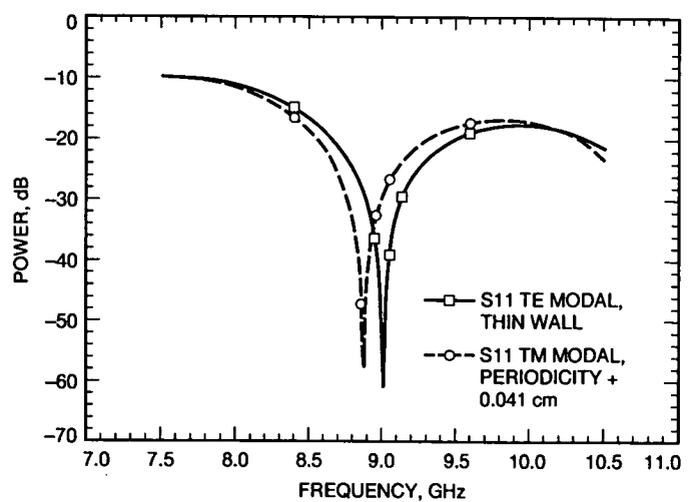


Fig. 16. Effect of periodicity change on the TM reflection coefficient.

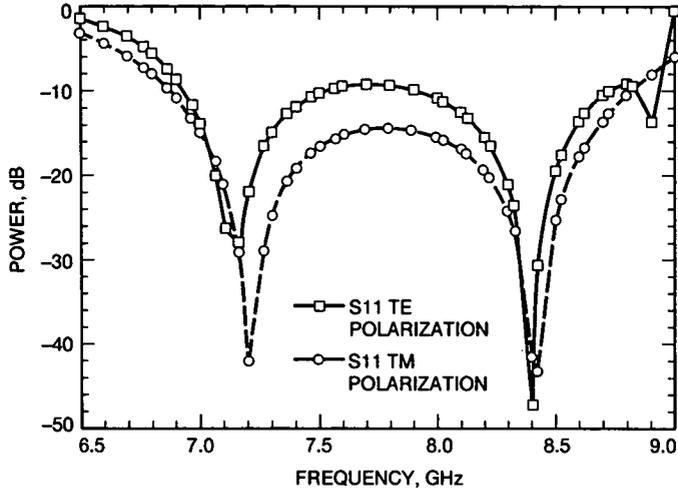


Fig. 17. Power reflection coefficient for a dichroic plate with cross-shaped holes ($a_x = 2.65$ cm, $b_x = 0.6$ cm, $a_y = 2.76$ cm, $b_y = 0.7125$ cm, depth = 3.35 cm).

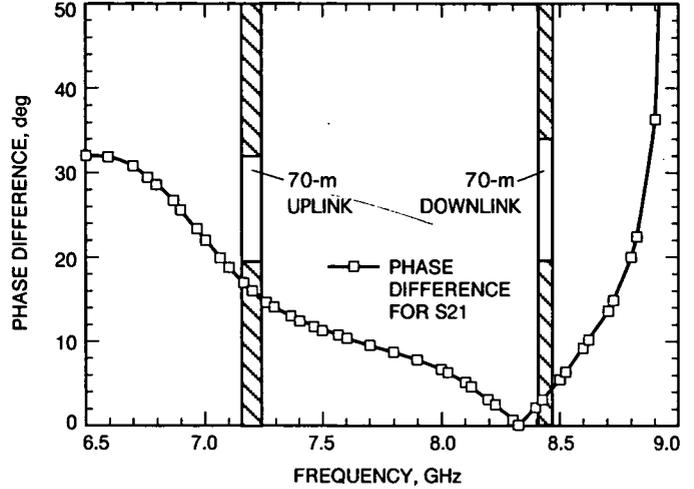


Fig. 19. Phase difference between polarizations for a dichroic plate with cross-shaped holes ($a_x = 2.65$ cm, $b_x = 0.6$ cm, $a_y = 2.76$ cm, $b_y = 0.7125$ cm, depth = 3.35 cm).

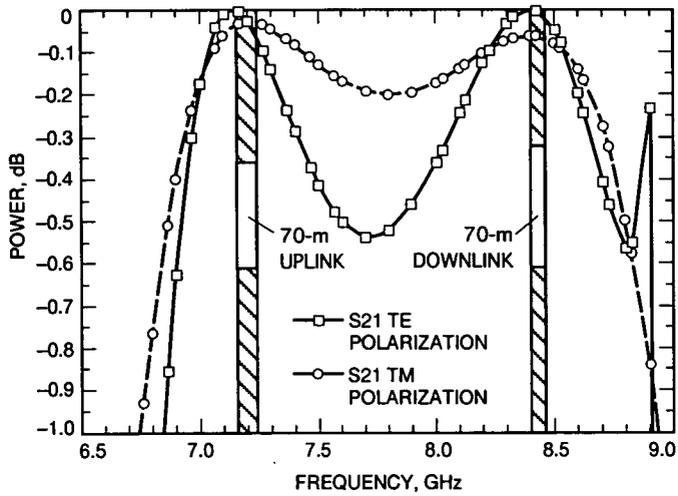


Fig. 18. Power transmission coefficient for a dichroic plate with cross-shaped holes ($a_x = 2.65$ cm, $b_x = 0.6$ cm, $a_y = 2.76$ cm, $b_y = 0.7125$ cm, depth = 3.35 cm).

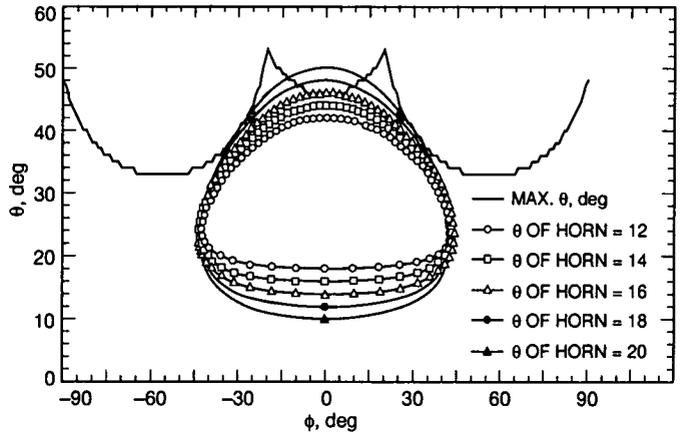
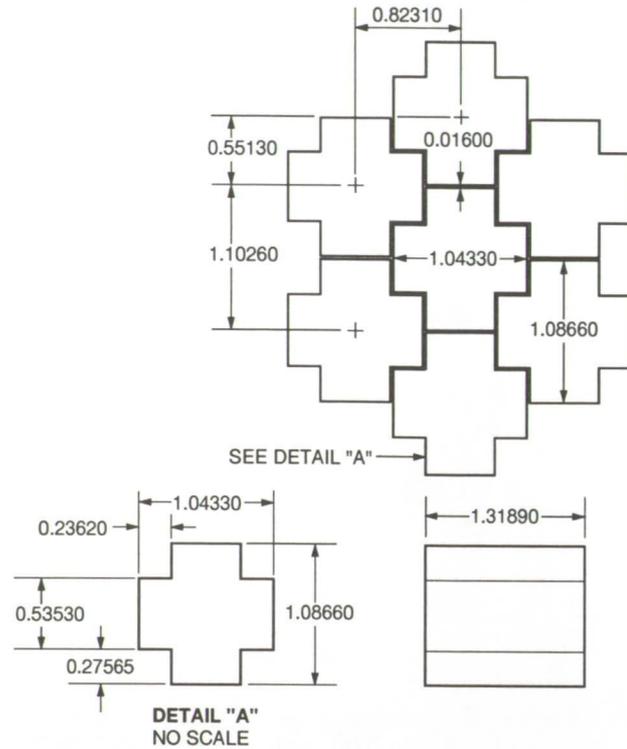


Fig. 20. Maximum angle, θ , at which a grating lobe propagates, $f = 8.625$ GHz ($a_x = 2.65$ cm, $b_x = 0.6$ cm, $a_y = 2.76$ cm, $b_y = 0.7125$ cm).



NOTE: MEASUREMENTS ARE IN INCHES

Fig. 21. Cross-shaped prototype detailed drawing.

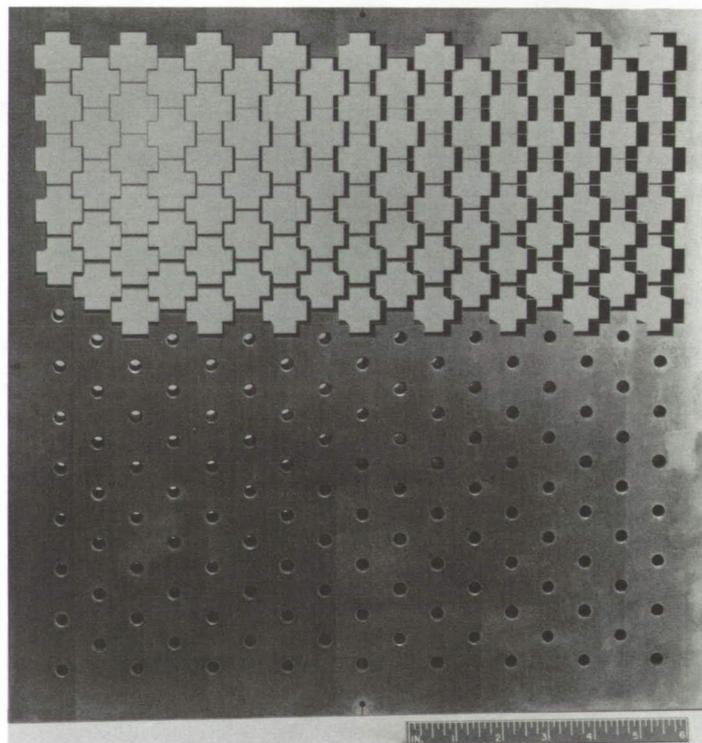


Fig. 22. Prototype of a dichroic plate with cross-shaped holes.

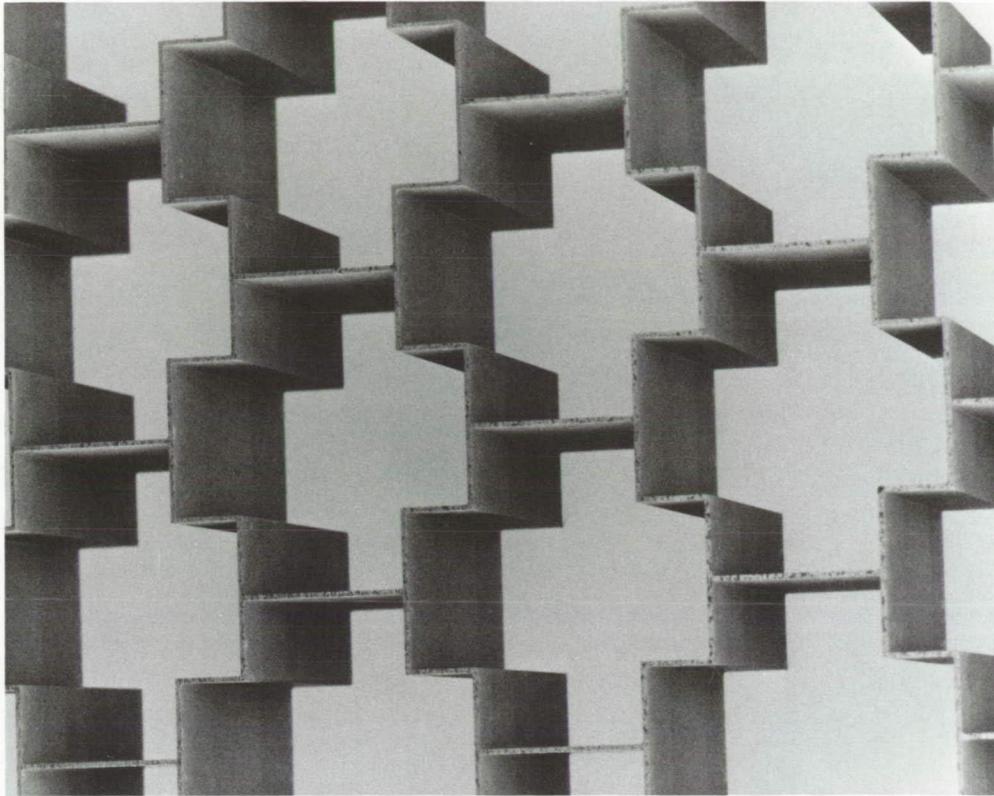


Fig. 23. Cross-shaped holes.

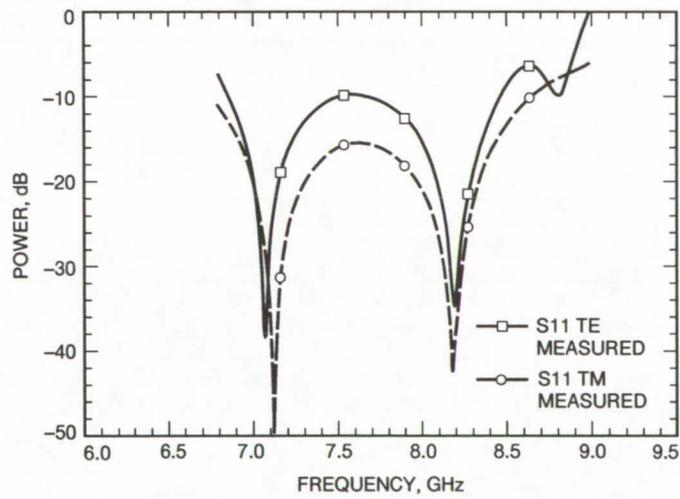


Fig. 24. Measured power reflection coefficient for a dichroic plate with cross-shaped holes.

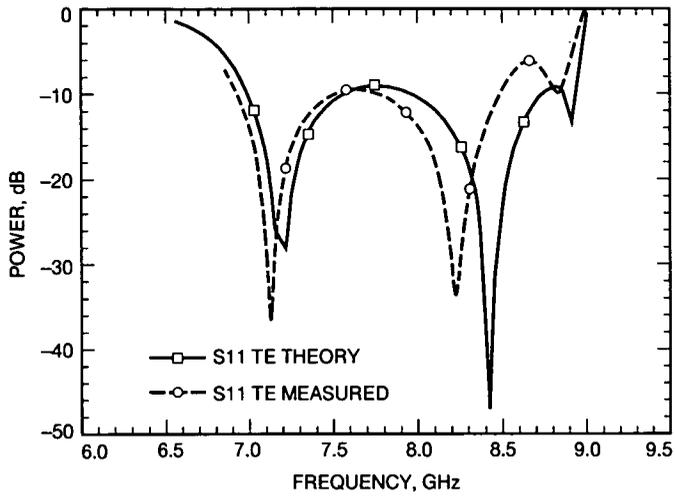


Fig. 25. Power reflection coefficient, initial comparison of measurement and theory for TE polarization.

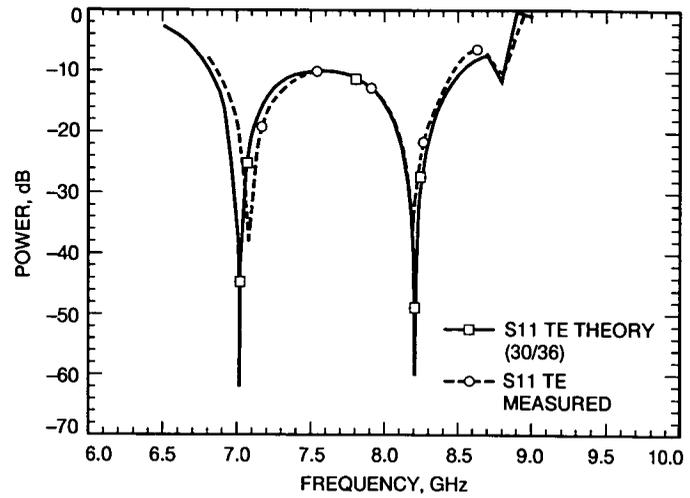


Fig. 27. Power reflection coefficient, corrected theory and measurement for TE polarization.

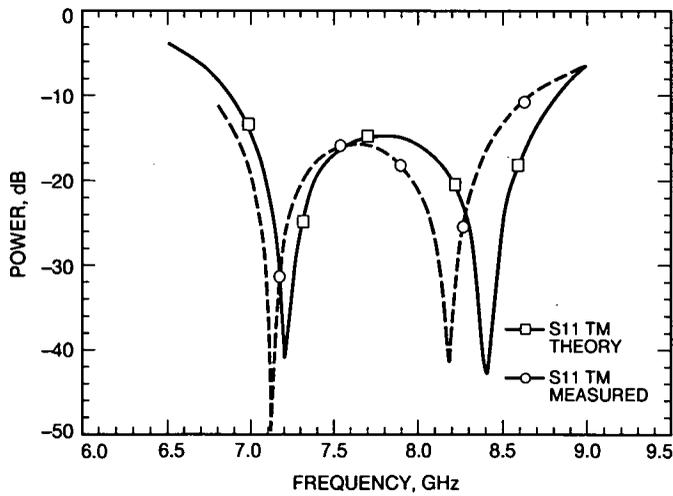


Fig. 26. Power reflection coefficient, initial comparison of measurement and theory for TM polarization.

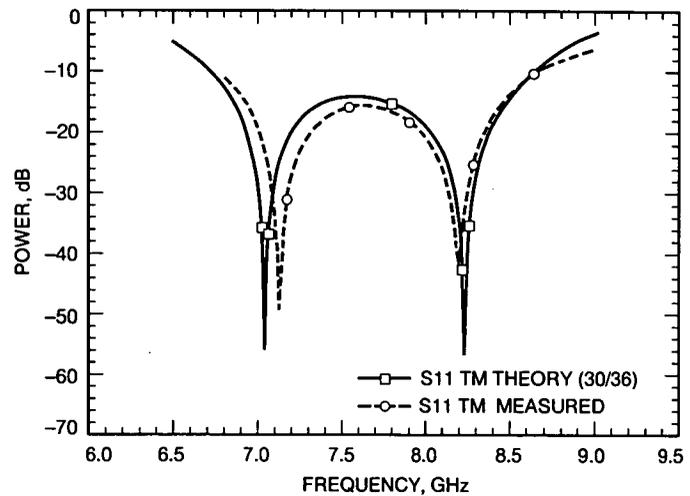


Fig. 28. Power reflection coefficient, corrected theory and measurement for TM polarization.

56-35
167926
P-11

Gravity Referenced Elevation Encoder Development

R. E. Goddard
Guidance and Control Section

Recent progress in the development of a gravity-sensor-based instrument for determining the elevation angle of DSN antennas is described. The benefits of such a system include the capability to locate the Gravity Referenced Elevation Encoder (GREE) directly on the primary reflector (thus bypassing structural flexure and deformation error sources), anticipated lower maintenance costs compared to the present gimbal encoders, direct replaceability or supplementation of the present gimbal encoders and the utilization of off-the-shelf components to construct the GREE. This article includes a description of the nominal GREE design. Test results on a laboratory breadboard model are given. Rigid-body dynamics of the GREE are derived and the simulated performance in response to measured antenna vibrations is given.

I. Introduction

Large structure ground-based antennas are subject to pointing errors which prevent achieving the full designed performance of the antennas [1-3]. Major error sources are gravity sag, thermal gradients and hysteresis of the primary reflector and its supporting structure. Therefore, the mechanical boresight of the primary reflector is not directly observable from the angle encoders on the axes' gimbals because of the large amount of structural separation. As an alternative to gimbal measurements, an instrument capable of sensing the Earth's gravity vector may be placed directly on the primary reflector to bypass the structural error sources. Knowledge of the gravity vector in antenna coordinates defines the elevation angle. In Section II, a Gravity Referenced Elevation Encoder (GREE) is described and the results of recently completed laboratory proof-of-concept tests of noise and long term drift are given. The rigid body dynamics of the GREE are derived in Section III and the nominal controller design is given. Simulated performance of the GREE in a DSN disturbance environment is detailed in Section IV.

II. Gravity Referenced Elevation Encoder

The GREE is constructed of all off-the-shelf components. Figure 1 is a pictorial of the GREE. The major components are the brushless DC motor, optical encoder, and accelerometer. The accelerometer is mounted via an adapter block onto the optical encoder rotor shaft. The opposite end of this shaft is rotated by the motor until the accelerometer output is nulled. The angular rotation required to null the accelerometer is read on the optical encoder. This angle is the elevation angle of the GREE itself. Knowledge of the calibration matrix between the GREE and the primary reflector mechanical boresight yields the antenna elevation angle. A block diagram of the GREE is shown in Fig. 2 and the laboratory unit is shown in Fig. 3. It is anticipated that thermal stabilization will be utilized in the field deployable unit to prevent temperature residual bias drift of the accelerometer. However, the laboratory model has neither heaters nor thermal insulation.

Long term stability of the sensor bias determines the time between recalibrations. Test results (Fig. 4) over 27

days indicate a 3σ of 0.8 mdegs after detrending. A trend of 0.2 mdeg is apparent in the data. The suspected cause is cyclic microtilting of the laboratory pier. The test was limited to an accelerometer on hand; however a superior version of the accelerometer, available at a modest cost increase, is expected to reduce the long term drift by 50 percent per manufacturer's specifications. Further, the drift may not be linear, which means that extrapolation of the test results is not applicable. Although the results are encouraging, longer duration tests should be completed to characterize the bias drift. A Hewlett-Packard spectrum analyzer recorded the short term noise spectrum. The rms noise is 0.32 mdeg at a 3.63-Hz sampling resolution and an accelerometer bandwidth of 250 Hz.

III. Rigid Body Dynamics

In this section, the GREE rotational dynamics are derived. The accelerometer is a torque-balanced pendulous mass sensor as shown in Fig. 5. The torque τ is developed by internal control electronics to maintain the angular position of the mass in accelerometer body coordinates. The amount of torque required is $\tau = mgl \sin \theta$ and the output is zero when the accelerometer sensitive axis is aligned with gravity.

The GREE dynamics are derived utilizing a two degree-of-freedom planar model. One degree of freedom represents the rotation axis of the optical encoder and the second degree of freedom is the accelerometer rotation axis. Figure 6 details the dynamic model where for simplicity the misalignments are assumed zero, i.e., the calibration matrix is the identity transformation. The following quantities are defined:

ϵ = elevation angle

θ_1 = rotor shaft angle with respect to gravity

x_i, y_i = location of center of mass of member $i, i = 1, 2$

x_0, y_0 = rotor axis of rotation

COM_i = center of mass of member $i, i = 1, 2$

k_1 = distance of accelerometer axis from rotor COM_1

θ_I = accelerometer pendulous mass angle with respect to accelerometer case

θ_2 = accelerometer pendulous mass angle with respect to gravity

$f x_i, f y_i$ = resultant coupling forces on rotational axis $i, i = 1, 2$

g = gravitational constant

l_i = COM_i offset from rotational axis $i, i = 1, 2$

I_i = inertia of member i about $COM_i, i = 1, 2$

The axis of rotation of the rotational member consisting of the optical encoder shaft, accelerometer, and actuator rotor (see Fig. 1) is given by the coordinates x_0, y_0 . The kinematics of the system defined in Fig. 6 are ($s_i = \sin \theta_i, c_i = \cos \theta_i$):

$$x_1 = x_0 + l_1 c_1 \quad x_2 = x_1 + k_1 c_1 + l_2 c_2$$

$$y_1 = y_0 + l_1 s_1 \quad y_2 = y_1 + k_1 s_1 + l_2 s_2$$

Twice time differentiating the above equations yields:

$$\left. \begin{aligned} \ddot{x}_1 &= \ddot{x}_0 - l_1 (\dot{\theta}_1^2 c_1 + \ddot{\theta}_1 s_1) \\ \ddot{y}_1 &= \ddot{y}_0 + l_1 (-\dot{\theta}_1^2 s_1 + \ddot{\theta}_1 c_1) \\ \ddot{x}_2 &= \ddot{x}_0 - l_1 (\dot{\theta}_1^2 c_1 + \ddot{\theta}_1 s_1) - k_1 (\dot{\theta}_1^2 c_1 + \ddot{\theta}_1 s_1) - l_2 (\dot{\theta}_2^2 c_2 + \ddot{\theta}_2 s_2) \\ \ddot{y}_2 &= \ddot{y}_0 + l_1 (-\dot{\theta}_1^2 s_1 + \ddot{\theta}_1 c_1) + k_1 (-\dot{\theta}_1^2 s_1 + \ddot{\theta}_1 c_1) + l_2 (-\dot{\theta}_2^2 s_2 + \ddot{\theta}_2 c_2) \end{aligned} \right\} \quad (1)$$

The translational dynamics are

$$\left. \begin{aligned} m_1 \ddot{x}_1 &= f_{x_1} - f_{x_2} + m_1 g & m_1 \ddot{y}_1 &= f_{y_1} - f_{y_2} \\ m_2 \ddot{x}_2 &= f_{x_2} + m_2 g & m_2 \ddot{y}_2 &= f_{y_2} \end{aligned} \right\} \quad (2)$$

Solving for the coupling forces (f_{x_i} , f_{y_i}) in terms of the translational accelerations and substituting out these accelerations via Eq. (1) yields

$$\left. \begin{aligned} f_{x_1} &= (m_1 + m_2) [\ddot{x}_0 - l_1 (\dot{\theta}_1^2 c_1 + \ddot{\theta}_1 s_1)] - m_2 k_1 [\dot{\theta}_1^2 c_1 + \ddot{\theta}_1 s_1] - m_2 l_2 [\dot{\theta}_2^2 c_2 + \ddot{\theta}_2 s_2] - (m_1 + m_2) g \\ f_{x_2} &= m_2 \ddot{x}_0 - m_2 l_1 (+\dot{\theta}_1^2 c_1 + \ddot{\theta}_1 s_1) - m_2 k_1 [+ \dot{\theta}_1^2 c_1 + \ddot{\theta}_1 s_1] - m_2 l_2 [\dot{\theta}_2^2 c_2 + \ddot{\theta}_2 s_2] - m_2 g \\ f_{y_1} &= (m_1 + m_2) \ddot{y}_0 + (m_1 + m_2) l_1 (-\dot{\theta}_1^2 s_1 + \ddot{\theta}_1 c_1) + m_2 k_1 (-\dot{\theta}_1^2 s_1 + \ddot{\theta}_1 c_1) + m_2 l_2 (-\dot{\theta}_2^2 s_2 + \ddot{\theta}_2 c_2) \\ f_{y_2} &= m_2 \ddot{y}_0 + m_2 l_1 (-\dot{\theta}_1^2 s_2 + \ddot{\theta}_1 c_1) + m_2 k_1 (-\dot{\theta}_1^2 s_1 + \ddot{\theta}_1 c_1) + m_1 l_2 (-\dot{\theta}_2^2 s_2 + \ddot{\theta}_2 c_2) \end{aligned} \right\} \quad (3)$$

The rotational dynamics may now be written and the coupling forces substituted out to arrive at the dynamics in rotational coordinates. The results are

$$\left. \begin{aligned} [I_1 + l_1^2 (m_1 + m_2) + 2l_1 k_1 m_2 + k_1^2 m_2] \ddot{\theta}_1 + [(l_1 + k_1) l_2 m_2 \cos(\theta_2 - \theta_1)] \ddot{\theta}_2 \\ - [(l_1 + k_1) l_2 m_2 \sin(\theta_2 - \theta_1)] \dot{\theta}_2^2 + l_1 s_1 (m_1 + m_2) g + k_1 s_1 m_2 g \\ = +l_1 s_1 (m_1 + m_2) \ddot{x}_0 - l_1 c_1 (m_1 + m_2) \ddot{y}_0 + k_1 s_1 m_2 \ddot{x}_0 - k_1 c_1 m_2 \ddot{y}_0 + \tau_1 - \tau_2 \end{aligned} \right\} \quad (4a)$$

$$\left. \begin{aligned} [(l_1 + k_1) l_2 m_2 \cos(\theta_2 - \theta_1)] \ddot{\theta}_1 + [(l_1 + k_1) l_2 m_2 \sin(\theta_2 - \theta_1)] \dot{\theta}_1^2 \\ + [I_2 + m_2 l_2^2] \ddot{\theta}_2 + l_2 m_2 s_2 g = +l_2 m_2 s_2 \ddot{x}_0 - l_2 m_2 c_2 \ddot{y}_0 + \tau_2 \end{aligned} \right\} \quad (4b)$$

In the dynamics, the accelerations (\ddot{x}_0, \ddot{y}_0) represent composite motion of the GREE due to antenna vibrations (\ddot{x}'_0, \ddot{y}'_0) and a jitter on the elevation angle ($\dot{\epsilon}, \ddot{\epsilon}$). This is expressed as

$$\ddot{x}_0 = \ddot{x}'_0 - l_0 (\sin \epsilon \ddot{\epsilon} + \cos \epsilon \dot{\epsilon}^2) \quad (5a)$$

$$\ddot{y}_0 = \ddot{y}'_0 - l_0 (\cos \epsilon \ddot{\epsilon} - \sin \epsilon \dot{\epsilon}^2) \quad (5b)$$

where l_0 is the off-axis mounting of the GREE. Table 1 lists the parametric values utilized for the GREE. The values are obtained from actual measurements of the GREE components, vendor specifications, or calculations. The

parameter l_0 represents the assumed off-axis mounting of the GREE on the antenna; and the parameter l_1 represents the assumed center-of-mass imbalance on the optical encoder rotor. These two quantities are installation and fabrication constraints, respectively.

The linearized GREE model is shown in Fig. 7. The actuator is modeled in the frequency domain as a simple first-order torque generator:

$$\tau = \frac{\left(\frac{k_\tau}{l}\right) (v - k_v s \theta)}{\left(s + \frac{r}{l}\right)} \quad (6)$$

where k_r , k_v , r and l are the torque constant, back electromotive force (EMF) constant, coil resistance, and coil inductance, respectively, and s is the complex frequency. The accelerometer model is a second-order critically damped system with an appropriate scale factor (SF). The linearized dynamics of the rotor are derived utilizing $I_2 \ll I_1$:

$$\frac{\theta}{\tau} = \frac{1}{s^2 + \left(\frac{m_1 l_1 g}{I_1 + m_1 l_1^2} \right)} \equiv \frac{k_r}{(s^2 + \omega_0^2)} \quad (7)$$

A generalized model of bearing friction [4], Fig. 8, describes the restoring torque as a function of rotation. This represents the resistance to rotation exhibited by the optical encoder bearing action. The model is characterized by coulomb torque (T_c), the spatial time constant (τ_c), and the material factor (i). All of the curves in Fig. 8 have equal slope (T_c/τ_c) at the origin of the torque-rotation curve and therefore an exponential approximation may be utilized. An analytical torque-rotation curve with hysteresis is shown in Fig. 9. This model is valid for $\dot{\beta} > 0$ and $\dot{\beta} < 0$. The expression for this model utilized in the subsequent simulations is

$$T(\beta) = \text{sgn}(\dot{\beta})T_c + [T(0) - \text{sgn}(\dot{\beta})T_c][e^{-(\beta/\tau_c)}] \quad (8)$$

where $T(0)$ is the starting torque and β is the angular change from the starting (last reversal) angle. Since a linearized model is required for the controller design, a spring constant is defined by

$$\Delta T = \frac{dT}{d\beta} \Delta\beta = \frac{T_c}{\tau_c} \Delta\beta \equiv k \Delta\beta \quad (9)$$

Table 2 lists values for the GREE components.

The linearized dynamics (Fig. 7) are utilized to compute a controller. The calculations are considerably simplified if higher order modes are neglected. Referring to Table 2, the motor and accelerometer dynamics are widely separated from the plant dynamics. Also combining the back EMF and friction loops into the plant dynamics yields the reduced-order block diagram in Fig. 10. A classical proportional, integral, and derivative (PID) controller is

$$H = b_1 s + b_0 + \frac{b_{-1}}{s} \quad (10)$$

The rotor angle θ_1 is the error signal. If, for example, $\theta_1(\omega) = 0$ for all ω , then $\beta = \epsilon$, and the system is error free. The controller H must keep θ_1 small. From Fig. 10, the closed-loop error transfer function is

$$\frac{\theta_1}{\epsilon} = \frac{k k_r s}{s^3 + \left[2\zeta\omega_n + \frac{k_r SF k_r b_1}{r} \right] s^2 + \left[\frac{k_r SF k_r b_0}{r} + \omega_n^2 \right] s + \left[\frac{k_r SF k_r b_{-1}}{r} \right]} \quad (11)$$

For simplicity let the closed loop response be three first order responses possessing break frequencies ω_1 , ω_2 , and ω_3 . The controller solution is

$$b_1 = \left[\frac{r}{k_t SF k_r} \right] [\omega_1 + \omega_2 + \omega_3 - 2\zeta\omega_n]$$

$$b_0 = \left[\frac{r}{k_t SF k_r} \right] [\omega_1\omega_2 + \omega_1\omega_3 + \omega_2\omega_3 - \omega_n^2]$$

$$b_{-1} = \left[\frac{r}{k_t SF k_r} \right] [\omega_1\omega_2\omega_3] \quad (12)$$

The PID controller places a zero at the origin in the error transfer function, Eq. (11). For the analytical models used, a constant input will result in zero error. This is desirable since the nominal trajectory of a celestial track has a fundamental period of $1/(86,400)$ Hz. The error frequency response can be made flat over a particular band. For example, let $\omega_1 = 2\pi(0.01)$ and $\omega_2 = \omega_3 = 2\pi(500)$ rad/sec. The resulting Bode diagram for the error $\theta_1(\omega)$, Fig. 11, is flat from 0.01 to 500 Hz with a minimum error attenuation of -80 dB.

IV. Nonlinear Simulations

The nonlinear equations derived in Section III are utilized to simulate the GREE response to DSN antenna

vibrations. The nonlinear friction model (Fig. 9) is included. Measured vibration data on DSS 14, as given in Fig. 12, are the input. Also included are the effects of angular jitter on the elevation angle $\varepsilon(t)$ at a 20-Hz rate and 0.01-rad amplitude. A fourth-order Runge-Kutta integration at a step size of 10 msec computed the GREE response given in Fig. 13. The expected error is less than 1 mdeg under these conditions.

V. Summary

The present state of development of an inertial instrument for DSN antennas has been described. A breadboard model GREE tested in the laboratory demonstrated acceptable noise and long-term drift characteristics. Rigid body analytical models and simulations have been devel-

oped for design and analysis. A nominal controller design maintained the simulated elevation angle error to less than 1 mdeg when driven by measured 70-m antenna vibrations.

The performance of a GREE under actual antenna vibrations can be determined by a field test. Because the GREE will retrofit both mechanically and electrically onto existing antennas, a performance comparison study with the existing gimbal encoders can be designed. The recently completed laboratory demonstration unit can be upgraded for the field test by the addition of heaters, thermal insulation, mechanical mounting hardware and an electrical interface. Prior to the field testing, the analytical model needs to be upgraded to incorporate the dominant flexible modes and the controller upgraded to maintain the desired performance in the specific antenna environment.

References

- [1] C. N. Guiar, F. L. Lansing, and R. Riggs, "Antenna Pointing Systematic Error Model Derivations," *The Telecommunications and Data Acquisition Progress Report 42-88*, vol. October-December 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 36-46, February 15, 1987.
- [2] K. Abichandani, "Unified Analysis for Antenna Pointing and Structural Error Part 1. Review," *The Telecommunications and Data Acquisition Progress Report 42-73*, vol. January-March 1983, Jet Propulsion Laboratory, Pasadena, California, pp. 40-59, April 15, 1983.
- [3] S. Katow, J. J. Cucchissi, K. L. Chuang, R. Levy, J. L. Lansing, F. W. Stoller, and F. Menninger, "Structural Design Options for the New 34 Meter Beam Waveguide Antenna," *The Telecommunications and Data Acquisition Progress Report 42-88*, vol. October-December 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 183-186, February 15, 1987.
- [4] P. R. Dahl, "Solid Friction Damping of Mechanical Vibrations," *AIAA Journal*, vol. 14, no. 12, pp. 1675-1688, December 1976.

Table 1. Mass properties for simulation.

Parameter	Value	
Inertia, kg-m ²	$I_1 = 1.28 \times 10^{-2}$	$I_2 = 5.65 \times 10^{-4}$
Mass, kg	$m_1 = 1.02 \times 10^0$	$m_2 = 1.20 \times 10^{-4}$
Distance to center of mass, m	$l_1 = 5.00 \times 10^{-3}$	$l_2 = 8.89 \times 10^{-3}$
Distance to accelerometer rotation axis, m	$k_1 = 1.00 \times 10^{-2}$	
Distance to elevation axis, m	$l_0 = 1.00 \times 10^0$	
Vibration scale factor, kg-m	$k_x = 5.09 \times 10^{-3}$	
Rotor transfer function, (kg-m ²) ⁻¹	$k_r = 7.78 \times 10^1$	
Rotor transfer function, rad/sec	$\omega_0 = 1.97 \times 10^0$	

Table 2. Component mechanical and electrical properties.

Component	Value
Brushless dc motor	
Torque constant, newton-m/ampere	$k_r = 1.91 \times 10^{-1}$
Inductance, henries	$l = 4.32 \times 10^{-2}$
Resistance, ohms	$r = 2.40 \times 10^1$
Voltage constant, volts/rad/sec	$k_v = 1.91 \times 10^{-1}$
Half-power frequency, rad/sec	$p_3 = 5.55 \times 10^2$
Accelerometer	
Scale factor, volts/rad	$SF = 2.87 \times 10^2$
Break frequency, Hz	$p_1, p_2 = 2.00 \times 10^3$
Optical encoder friction	
Coulomb torque, newton-m	$T_c = 5.30 \times 10^{-3}$
Spatial constant, rad	$\tau_c = 4.00 \times 10^{-4}$
Spring constant, newton-m/rad	$k = 1.32 \times 10^1$

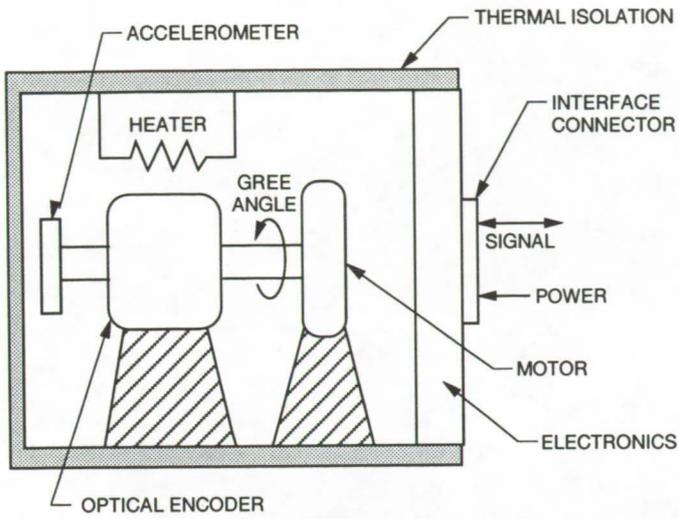


Fig. 1. GREE components.

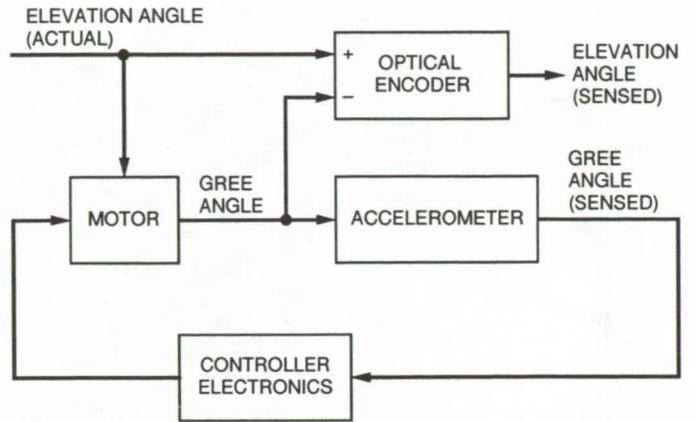


Fig. 2. GREE block diagram.

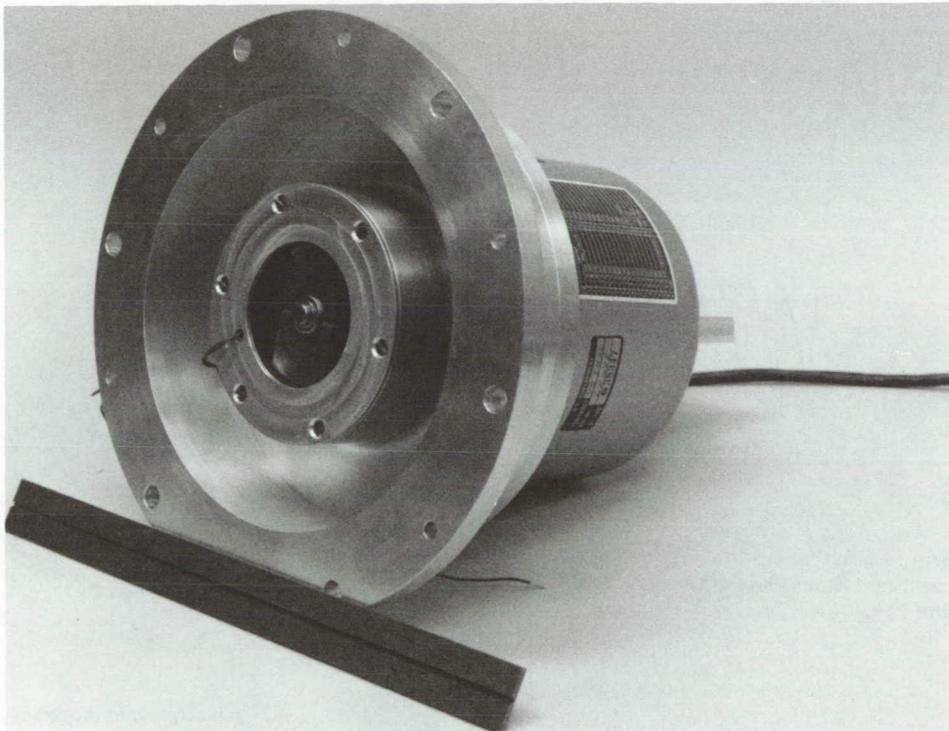


Fig. 3. Laboratory demonstration unit.

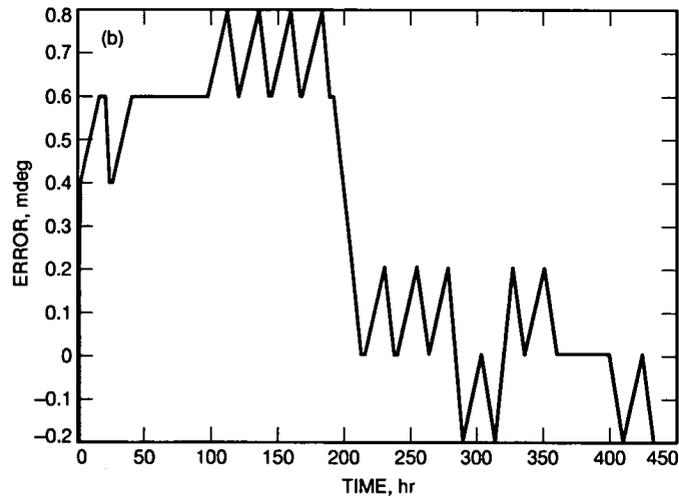
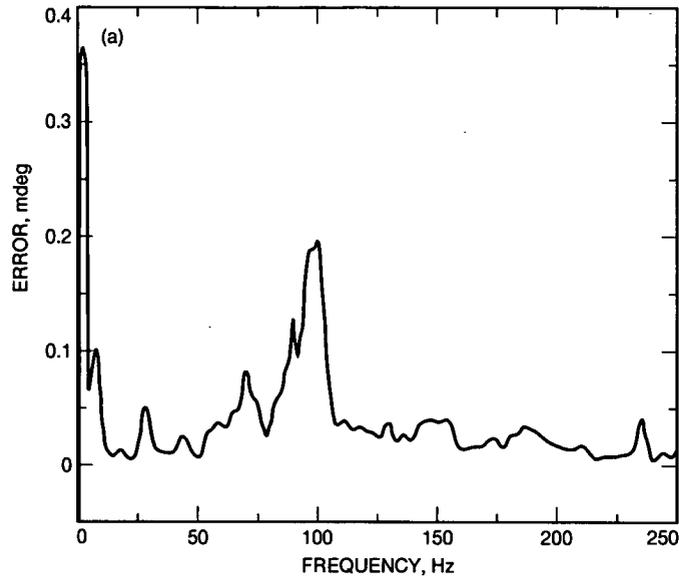


Fig. 4. Accelerometer-measured laboratory data: (a) noise spectrum and (b) 27-day drift test (Substrand QA-2000 accelerometer bias stability test, June 22 to July 17).

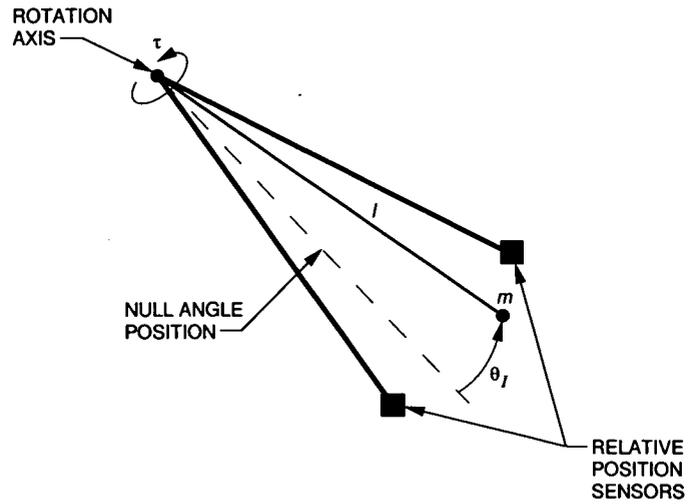


Fig. 5. Accelerometer principle of operation.

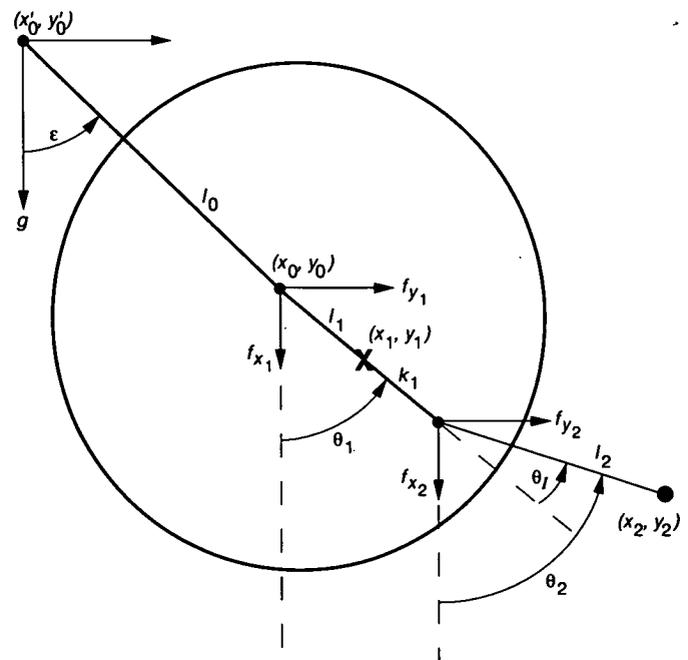


Fig. 6. Dynamic model of GREE.

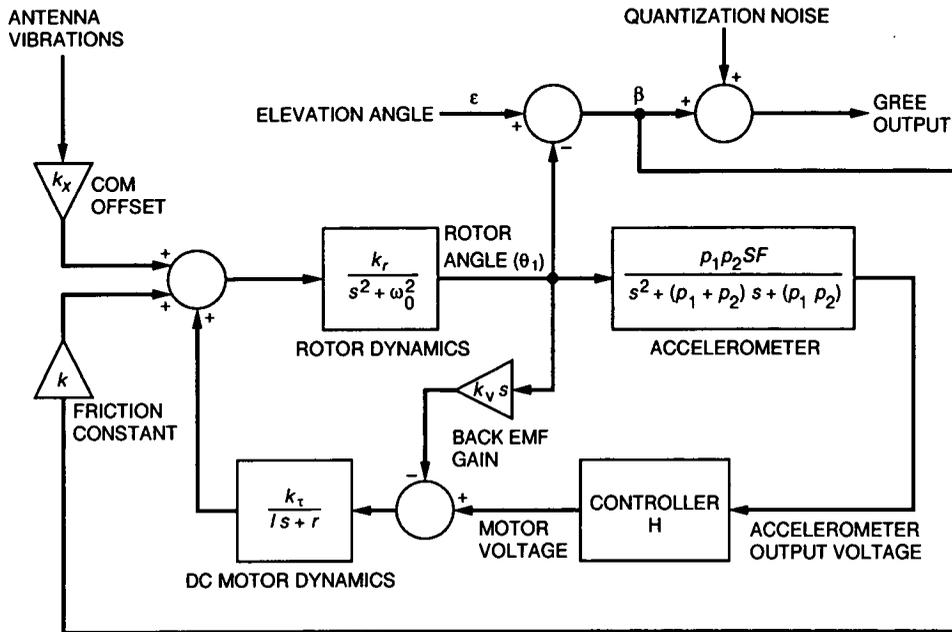


Fig. 7. Linearized dynamics for controller design.

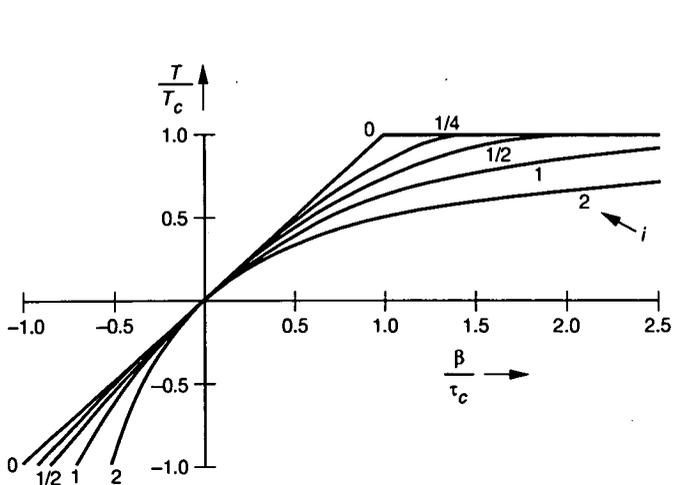


Fig. 8. Generalized force-displacement curve of solid friction damping.

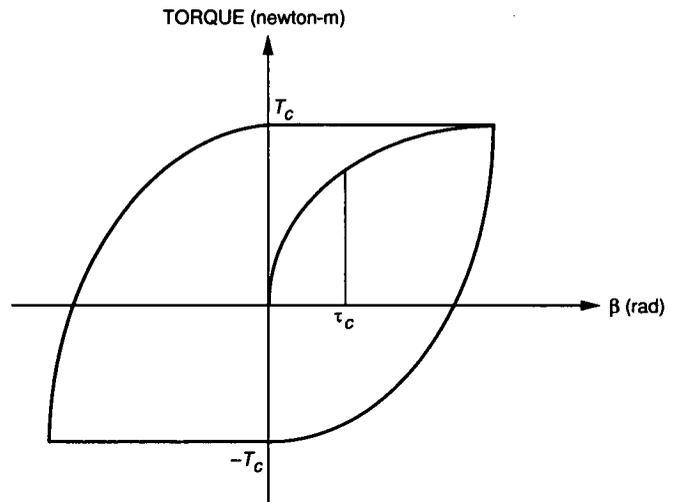


Fig. 9. Torque-rotation curve for linear controller design.

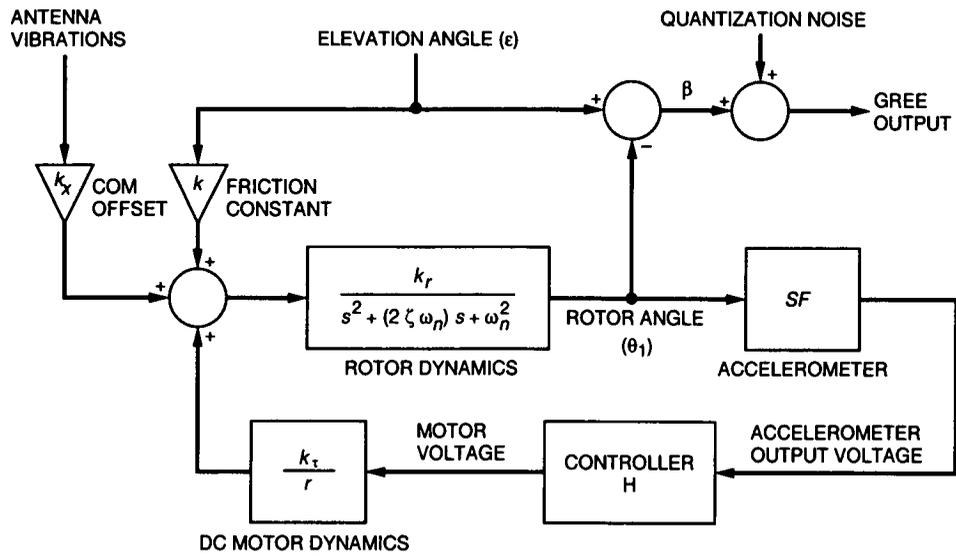


Fig. 10. PID controller and reduced-order GREE model ($\omega_N = 32.2$ rad/sec; $\zeta = 0.002$).

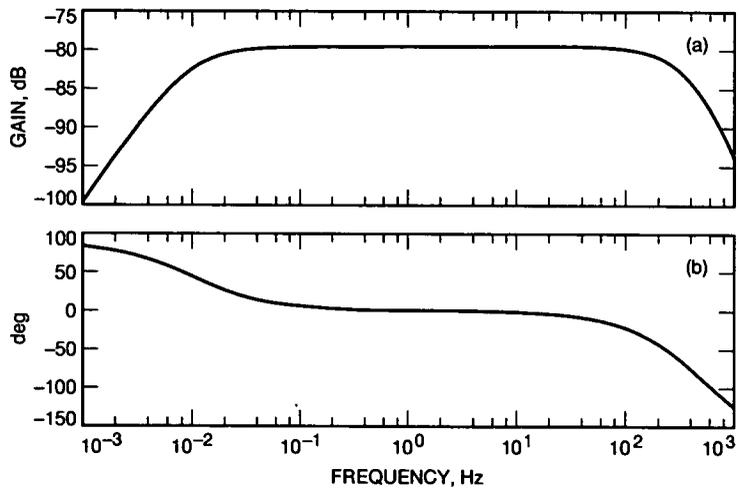


Fig. 11. Error frequency response: (a) error magnitude (small elevation changes) and (b) error phase (small elevation changes).

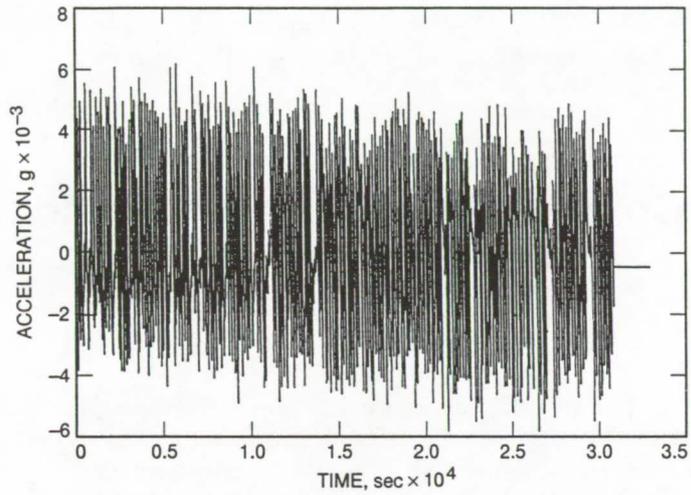


Fig. 12. Measured DSS-14 (70-m) antenna vibration data for simulation.

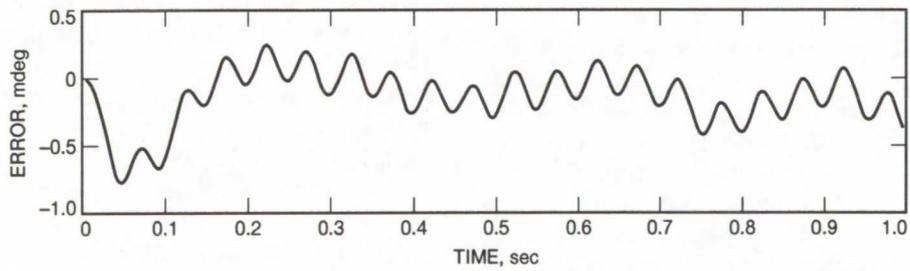


Fig. 13. Simulated GREE response to DSS-14 disturbance environment.

57-32
167927

500715

p-4

A Note on the Computation of Antenna-Blocking Shadows

R. Levy

Ground Antennas and Facilities Engineering Section

A simple and readily applied method is provided to compute the shadow on the main reflector of a Cassegrain antenna, when cast by the subreflector and the subreflector supports. The method entails some convenient minor approximations that will produce results similar to results obtained with a lengthier, mainframe computer program.

I. Introduction

The performance of conventional Cassegrain antennas is affected by the interference of the tripod or quadripod structures that support the subreflector. The interference is in the form of a shadow cast on the main reflector surface. The shadow blocks a portion of the microwave energy in either transmission or reception and is caused by the supporting tripod or quadripod legs, and also, to a less serious extent, by the subreflector. Such shadows can be avoided by offset antenna configurations, but the extra difficulties associated with offset construction are usually evaluated to be more objectional than the effects of blocking. Consequently, the symmetrical Cassegrain antenna predominates. Thus, it is important to be able to evaluate the extent of the blocking.

II. Analysis

The shadowed area consists of two types of blocking: plane wave and spherical wave. The plane-wave blocking is caused by the projections of the subreflector and an upper portion of the support legs on the aperture plane. The spherical-wave blocking is the shadow on the aperture plane of rays emanating from the focal point that

intersect the lower portion of the support leg. Figure 1, which is taken from [1], shows typical shadows from the two effects projected on the aperture plane. J. Herndon [1] developed a comprehensive numerical-integration computer program to calculate the blocked areas. However, results close to those from his computer program, which requires a mainframe computer, can be readily obtained with some simple approximations. The method presented here is suitable for execution on either a programmable calculator or a personal computer.

Figure 2(a) is a profile sketch of the reflector, taken in the plane of one of the support legs. The leg is assumed to have a trapezoidal cross section that is opaque with respect to transmission of microwave energy. Symbols of the figure are

F = focal length

R = main-reflector radius

R_S = subreflector radius

S = radial distance to centerline of leg at the intersection with the main-reflector surface

Z_0 = Z coordinate at S

Z_S = Z coordinate at back of subreflector

h = half of the leg depth

w_I = width of leg structure at inside face

w_O = width of leg structure at outside face

β = angle from the focal point to the rim of the main reflector

ϕ = slope of surface at intersection with the centerline of the leg

ψ = slope angle of the leg

Figure 2(b) is an expanded detail at the intersection of the leg with the surface. S_I and S_O are the radial distances to the points where the extensions of the inner and outer faces of the leg would intersect the surface, and Q is the distance along the tangent from the leg centerline to either of the intersection points at S_I or S_O . The relatively small curvature makes it reasonable to replace the curved surface by the tangent in the vicinity of S . Q is given by

$$Q = \frac{h}{\sin(\psi + \phi)} \quad (1)$$

Therefore

$$S_I = S - Q \cos \phi \quad (2)$$

and

$$S_O = S + Q \sin \phi \quad (3)$$

Figure 2(c) shows the spherical-wave shadow of the leg as a trapezoid of length $R - S_O$. To find the maximum width of the trapezoid at the rim of the antenna, w_M , it is necessary to find the distances X_I and X_O where a ray from the focal point to the rim crosses the inner and outer faces of the leg. To find X_I , for example, one has

$$F - Z_I = X_I / \tan \beta + (S_I - X_I) \tan \psi \quad (4)$$

in which Z_I is the Z coordinate at S_I . By introducing Z_O , the Z coordinate at S_O , a similar expression can be formed for X_O , and these expressions can be used to determine X_I and X_O .

If the width at the outer face of the leg governs the spherical-wave shadow, then the width of the trapezoid w_B at S_O is w_O , and the width at the rim is

$$w_M = w_O R / X_O \quad (5)$$

If the width at the inner face of the leg governs, it is necessary to find the width of the trapezoid at S_O . To do this, one uses the distance X_{IO} , which is where a ray from the focal point to the surface at S_O intersects the inner leg face. X_{IO} can be found from the following expression:

$$F - Z_I = X_{IO} / \tan \beta_P + (S_I - X_{IO}) \tan \psi \quad (6)$$

in which

$$\tan \beta_P = S_O / (F - Z_O) \quad (7)$$

and in this case the width of the spherical-wave-blocking trapezoid at its base is

$$w_B = w_I S_O / X_{IO} \quad (8)$$

and the width at the rim is

$$w_M = w_I R / X_I \quad (9)$$

The ideal profile for the leg cross-section is when the outer face provides the same width at the rim as the inner face does. In this case, the outer width would be

$$w_O(\text{opt}) = w_I S_I / X_{IO} \quad (10)$$

The foregoing computations imply several approximations that are expected to have only a minor effect on the results. These are

- (1) The leg is assumed to be entirely opaque. This is usually an accurate assumption for the spherical-wave shadow of the leg because the inner, narrowest face is often an opaque solid plate. The plane-wave portion of the leg shadow could be reduced because of a less-than-unity solidity ratio of the projection of the leg trusses on the aperture plane. Consequently, this openness could allow microwave energy to pass through the open spaces. It is conservative to ignore the energy that could pass through the leg regions, and, if it is thought to be significant, the plane area shadow can be reduced by the complement of the solidity ratio. This could possibly reduce the total leg shadow by 10 to 15 percent.

- (2) The spherical-wave leg shadow is modelled by the projection of a trapezoid on the aperture plane. The long sides of the trapezoid actually are curved, and the approach here slightly overestimates the shadow.
 - (3) The curve of the outer reflector rim is replaced by the straight edge of the trapezoid.
 - (4) The leg profile is taken to have a constant cross-section for the full length, and any customary tapering towards a narrow point at the leg base is ignored.
- (4) At the prompt "supply—," the user can type NLEGS= ..., PSI= ..., etc., and when all the requested data are supplied, the user should type "return"; for versions of MATLAB prior to Version 4.0, "CTRL-Z" should be supplied to satisfy the requirements of the command "keyboard." If no data are supplied before providing "return" or "CTRL-Z," the default data, which are for DSS-15 class antenna quadripods, will be executed.

III. Method

Figure 3 provides a MATLAB program¹ to calculate the blocked shadow essentially as described above and to complete some of the omitted details. The total plane- and spherical-wave shadow areas and the relative proportions of each are provided. In addition, the user-furnished dimension Z_S is used to determine the clearance between the back edge of the subreflector and the inner support leg. A moderate acquaintance with any high-level coding language, such as FORTRAN, should make the code understandable, even for one with no prior exposure to MATLAB. However, the following may help one unfamiliar with the program:

- (1) The % symbol is interpreted as the beginning of a nonexecutable comment.
- (2) MATLAB is case sensitive, and almost all instructions and built-in functions require lowercase.
- (3) The choice (in this program) was to represent all the variables (including those of Fig. 2) in uppercase. WI represents w_I , TANBETA represents $\tan \beta$, PSI is ψ , and so forth.

¹ MATLAB is a registered trademark of The Math Works, Inc., Chituate Place, 24 Prime Park Way, Natick, MA 01760.

IV. Summary and Conclusions

The algorithms contained in the program of Fig. 3 were previously executed many times by a programmable desk calculator to compute the blocking during the design stages of the new quadripods for the 70-m antennas, the 34-m high-efficiency antennas, and the 34-m beam-waveguide tripod. The UNIVAC mainframe computer [1] program was used a number of times, and it was established that results agreed within practical requirements. In particular, in [1] an example computation for a 64-m antenna showed a blocking shadow of 7.35 percent and the procedure here showed 7.38 percent. The total leg-shadow area agreement was within 0.5 percent, although the allocations for the plane and spherical shadow contributions had larger differences. These differences could have been caused by the treatment of the pointed bases of the legs in the reference, which may have influenced whether a particular area was assigned to plane- or spherical-wave blocking. Both sets of computations, of course, agreed exactly on the subreflector shadow, which was only about 15 percent of the total.

The sample data built into the Fig. 3 program (data which the user is given the opportunity to replace) will result in a total shadow of 5.478 percent. The effect on the microwave antenna is more severe than the geometric aperture area reduction, perhaps by a factor of about two.

Reference

- [1] J. Herndon, "Efficient Antenna Systems: A Program to Calculate the Optical Blockage by the Quadripod on Large Microwave Antennas," *Space Programs Summary 37-48*, vol. II, Jet Propulsion Laboratory, Pasadena, California, pp. 58-63, November 30, 1967.

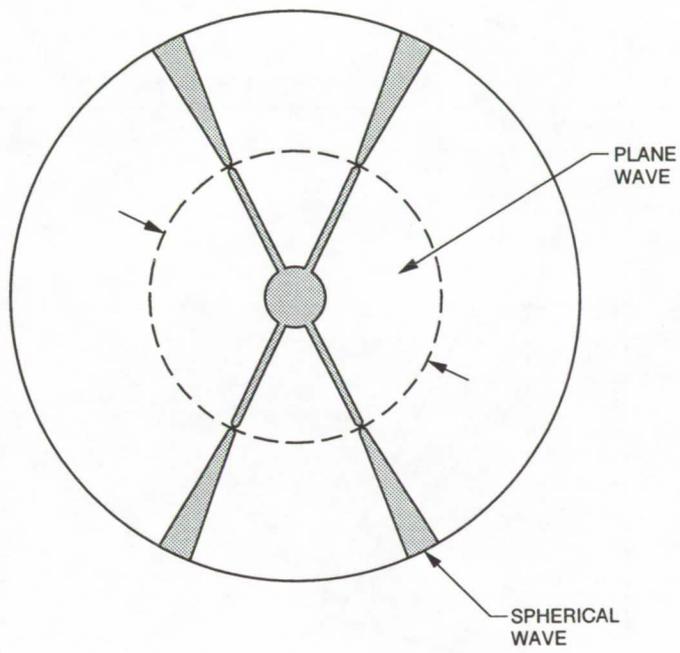


Fig. 1. Plane- and spherical-wave blocking.

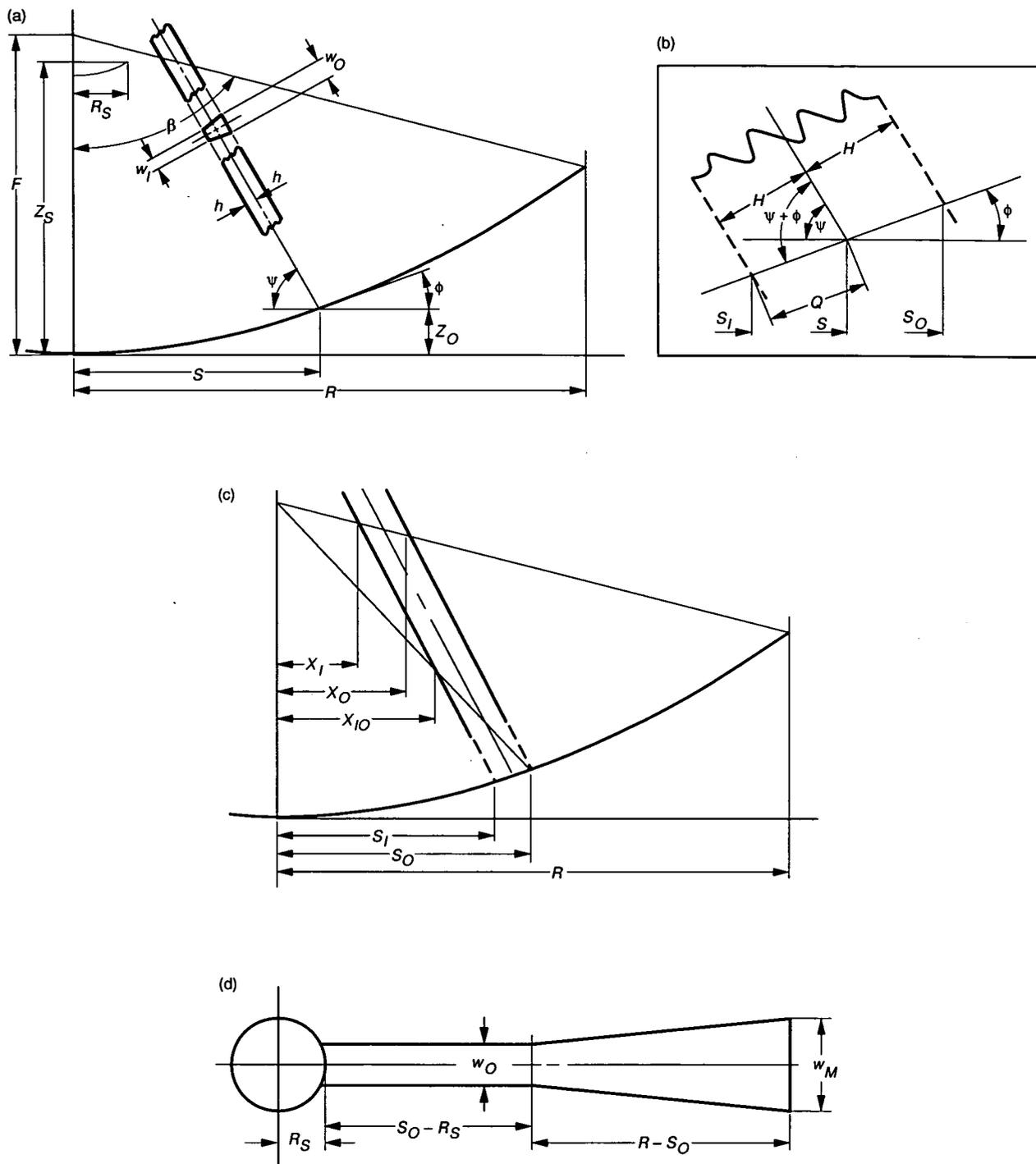


Fig. 2. Blocking geometry: (a) overall geometry, (b) detail at base of leg, (c) intersection with leg, and (d) leg-shadow projection.

```

%This is MATLAB\MISCPROBS\BLOCKING.M, Feb.10,1993
% compute blocking of subreflector and tripod or quadripod
% The following two functions are expected to be available to MATLAB:
%   function y=sine(x)           function y=cosine(x)
%   y=sin(x*pi/180);           y=cos(x*pi/180);

format compact
% Set some default values for 34M HEF antenna
NLEGS=4; PSI=61.3967; F=434; ZS=406.7;
R=669.3; RS=75.; S=328; H=19.45; WI=9.5; WO=14.

disp('supply- NLEGS,PSI,F,ZS,R,RS,S,H,WI,WO, AND "return" ')
keyboard
Tanphi=S/2/F;
PHI=atan(Tanphi)*180./pi;
Q=H/sine(PSI+PHI)
SI=S-Q*cosine(PHI)
SO=S+Q*cosine(PHI)
ZI=SI*SI/4/F
ZO=SO*SO/4/F
ZMAX=R*R/4/F;
TANBETA=R/(F-ZMAX)
TANP=sine(PSI)/cosine(PSI)
DEN=1/TANBETA-TANP
XI=(F-ZI-SI*TANP)/DEN
XO=(F-ZO-SO*TANP)/DEN
TANBETAP=SO/(F-ZO)
DENP=1/TANBETAP-TANP
XIO=(F-ZI-SI*TANP)/DENP
MAGI=R/XI
MAGO=R/XO
MAGIO=SO/XIO

WOPT=MAGIO*WI
AFACT=pi/144.
ASUB=RS*RS*AFACT %SQUARE FEET
AMAIN=R*R*AFACT
if WO>=WOPT
  ASPH=(R-SO)*WO*.5*(1+MAGO);
else
  ASPH=(R-SO)*WI*.5*(MAGIO+MAGI);
end
ASPH=ASPH*NLEGS/144
APLANE=WO*(SO-RS)*NLEGS/144
LEGSHAD=ASPH+APLANE
TOTSHAD=LEGSHAD+ASUB
TOTPCT=TOTSHAD/AMAIN*100
LEGPCT=LEGSHAD/AMAIN*100
% Blocking calculations completed above
% Now get leg-to-subreflector clearances
CLH=SI-(ZS-ZI)/TANP-RS % horizontal clearance
CLP=CLH*sine(PSI) % perpendicular to leg face clearance

```

Fig. 3. Program to calculate blocking.

58-32

167928

P-9

Maximum and Minimum Return Losses From a Passive Two-Port Network Terminated With a Mismatched Load

T. Y. Otoshi

Ground Antennas and Facilities Engineering Section

This article presents an analytical method for determining the exact distance a load is required to be offset from a passive two-port network to obtain maximum or minimum return losses from the terminated two-port network. Equations are derived in terms of two-port network S-parameters and load reflection coefficient. The equations are useful for predicting worst-case performances of some types of networks that are terminated with offset short-circuit loads.

I. Introduction

With the advent of automatic network analyzers, it has become common practice to characterize microwave two-port networks in terms of measured S-parameters [1]. As a result, many measurement techniques and network configurations are now being analyzed in terms of S-parameters rather than in terms of the impedance and admittance characterization parameters that have been used in the past. Microwave measurement techniques often employ the use of a sliding load, sliding short, or mismatched loads of known reflection-coefficient values. A frequently encountered problem of interest is determining the distance a particular load has to be offset from the two-port network output port to obtain a maximum or minimum return loss from the terminated two-port network. When the load is fixed (rather than sliding), line spacers can be inserted between the output of the two-port network and the load for producing the desired return losses. It is desirable to know what spacer-line length is required without having to resort to a trial-and-error process. A literature search indicates that analytical solutions, in terms of S-parameters, have not been published for these purposes.

This article will present these analytical expressions of interest.

Section II of this article sets forth the theory and analytical derivations of the conditions for maximum and minimum return losses. Also presented are design equations for dielectric spacers that can be used to obtain desired return losses. Section III presents examples showing how these equations can be used for predicting and avoiding the occurrence of a type of resonance phenomenon (leading to large dissipative losses) for some types of reflector antennas.

II. Theory

A. Derivations

The circuit under consideration is shown in Fig. 1 and the following conditions exist. It is assumed that the S-parameters of the two-port network and load reflection-coefficient values are known. In practice, if the load is fixed, the phase of its reflection coefficient as defined at the network output port can be varied through the use of

dielectric spacers. The following presents the derivation of equations required to calculate the specific conditions under which maximum or minimum return losses will be obtained from the terminated network.

Given the two-port network terminated in Γ_L as shown in Fig. 1, the input reflection coefficient as seen looking into port 1 is

$$\Gamma_1 = S_{11} + \frac{S_{12}S_{21}\Gamma_L}{1 - S_{22}\Gamma_L} \quad (1)$$

where S_{11} , S_{12} , S_{21} , and S_{22} are scattering parameters [1]. From Eq. (1),

$$|\Gamma_1| = \sqrt{\frac{|S_{11}|^2 + 2|S_{11}U\Gamma_L|\cos\theta_1 + |U\Gamma_L|^2}{1 - 2|S_{22}\Gamma_L|\cos\theta_2 + |S_{22}\Gamma_L|^2}} \quad (2)$$

where

$$U = S_{12}S_{21} - S_{11}S_{22} \quad (3)$$

$$\theta_1 = \psi_L + \psi_U - \psi_{11} \quad (4)$$

$$\theta_2 = \psi_L + \psi_{22} \quad (5)$$

and ψ_L , ψ_U , ψ_{11} , and ψ_{22} are the phase angles of Γ_L , U , S_{11} , and S_{22} , respectively.

In the following analysis, $|\Gamma_L|$ is assumed to be a constant and only ψ_L is allowed to vary. To find the maximum or minimum values of $|\Gamma_1|$ under these conditions, set

$$\frac{\partial |\Gamma_1|^2}{\partial \psi_L} = 0 \quad (6)$$

This leads to an equation of the form

$$A \sin \theta_1 + B \sin \theta_2 + C = 0 \quad (7)$$

where

$$A = |S_{11}U| (1 + |S_{22}\Gamma_L|^2) \quad (8)$$

$$B = |S_{22}| (|S_{11}|^2 + |U\Gamma_L|^2) \quad (9)$$

$$C = 2 |S_{11}S_{22}U\Gamma_L| \sin(\psi_{22} + \psi_{11} - \psi_U) \quad (10)$$

Substituting Eqs. (4) and (5) into Eq. (7), using trigonometric sum angle identities, and collecting terms leads to the expression

$$D \cos \psi_L + E \sin \psi_L + C = 0 \quad (11)$$

where

$$D = A \sin(\psi_U - \psi_{11}) + B \sin \psi_{22} \quad (12)$$

$$E = A \cos(\psi_U - \psi_{11}) + B \cos \psi_{22} \quad (13)$$

Upon expressing Eq. (11) in the equivalent form

$$\sqrt{D^2 + E^2} \cos(\psi_L - \varphi) = -C \quad (14)$$

where

$$\varphi = \tan^{-1} \left(\frac{E}{D} \right) \quad (15)$$

then, from Eq. (14), the values of ψ_L that correspond to maximum or minimum values of $|\Gamma_1|$ are

$$\psi_{Lm} = \varphi \pm \cos^{-1} \left(\frac{-C}{\sqrt{D^2 + E^2}} \right) \quad (16)$$

Substitution of $\psi_L = \psi_{Lm}$ into Eqs. (4), (5), and (2) gives the desired values of $|\Gamma_1|_{\max}$ and $|\Gamma_1|_{\min}$. Return losses are computed from

$$RL_{dB} = -20 \log_{10} |\Gamma_1| \quad (17)$$

B. Design of Spacers

As shown in Fig. 1, if the actual load reflection coefficient is Γ_T , but a particular Γ_L load reflection coefficient is desired to obtain a maximum or minimum return loss, a dielectric spacer can be utilized. The relationship for the dielectric spacer length of a nonferrous material is given as follows:

$$\ell = \frac{1}{2\beta} \arg R \quad (18)$$

where β is the phase constant whose expression will follow and $\arg R$ is the phase angle of a complex value R

$$R = \left(\frac{\Gamma_D + \Gamma_T}{1 + \Gamma_D \Gamma_T} \right) \left(\frac{1 + \Gamma_D \Gamma_L}{\Gamma_D + \Gamma_L} \right) \quad (19)$$

Eq. (19) was derived from input impedance and reflection coefficient relationships, starting at termination and transforming through the dielectric to the Z_{01} input interface.

Defining ϵ' to be the relative permittivity, $\tan \delta$ the loss tangent, λ_0 the free-space wavelength, and θ_i the incident angle, then the dielectric-air interface reflection coefficient for perpendicular polarization is

$$\Gamma_D = (\Gamma_D)_\perp = \frac{\sqrt{(\epsilon' - \sin^2 \theta_i) - j \epsilon''} - \cos \theta_i}{\sqrt{(\epsilon' - \sin^2 \theta_i) - j \epsilon''} + \cos \theta_i} \quad (20)$$

and for parallel polarization,

$$\begin{aligned} \Gamma_D &= (\Gamma_D)_\parallel \\ &= \frac{(\epsilon' - j \epsilon'') \cos \theta_i - \sqrt{(\epsilon' - \sin^2 \theta_i) - j \epsilon''}}{(\epsilon' - j \epsilon'') \cos \theta_i + \sqrt{(\epsilon' - \sin^2 \theta_i) - j \epsilon''}} \end{aligned} \quad (21)$$

For either polarization, the following relationships hold:

$$\epsilon'' = \epsilon' \tan \delta \quad (22)$$

$$\gamma = \alpha + j \beta \quad (23)$$

$$\alpha = \frac{2\pi}{\lambda_0} \sqrt{\epsilon' - \sin^2 \theta_i} \sqrt{\frac{\sqrt{1 + \tan^2 \delta'} - 1}{2}} \quad (24)$$

$$\beta = \frac{2\pi}{\lambda_0} \sqrt{\epsilon' - \sin^2 \theta_i} \sqrt{\frac{\sqrt{1 + \tan^2 \delta'} + 1}{2}} \quad (25)$$

$$\tan \delta' = \frac{\epsilon''}{\epsilon' - \sin^2 \theta_i} \quad (26)$$

When measurements are made in a TE_{10} mode waveguide, in the above equations for perpendicular polarization, one may substitute

$$\sin \theta_i = \frac{\lambda_0}{\lambda_c} \quad (27)$$

$$\cos \theta_i = \sqrt{1 - \left(\frac{\lambda_0}{\lambda_c} \right)^2} \quad (28)$$

where λ_c is the cutoff wavelength for the TE_{10} mode. For a dielectric sheet of thickness ℓ , the general-case S-parameter relationships that apply to perpendicular or parallel polarization in free space or to the TE_{10} mode waveguide cases are

$$(S_{11})_D = (S_{22})_D = \frac{-\Gamma_D (1 - e^{-2\gamma \ell})}{1 - \Gamma_D^2 e^{-2\gamma \ell}} \quad (29)$$

$$(S_{12})_D = (S_{21})_D = \frac{(1 - \Gamma_D^2) e^{-\gamma \ell}}{1 - \Gamma_D^2 e^{-2\gamma \ell}} \quad (30)$$

The above formulas were derived by considering the air-to-dielectric interfaces as separate individual two-port networks, and the dielectric media of length ℓ as the third two-port network. Then, the overall S-parameters for three cascaded networks were derived using cascading formulas given in [1].

III. Applications

A. Salisbury Screen

An application of the previously derived equations is the determination of maximum return loss that will occur when a perforated plate or mesh is terminated with an offset short circuit. For this type of configuration, sometimes referred to as a Salisbury Screen [2], it is of interest to know the exact offset distance, between the short and mesh, that will lead to maximum return loss. Maximum return loss for this type of configuration will occur when there is maximum dissipation due to the resistive losses of the perforated screen or mesh.

To demonstrate the application of the derived equations, Table 1 presents measured S-parameters of a perforated plate that has been used for the antenna reflector surface material on some DSN antennas. These measured S-parameters also agree well with those calculated from theory [3]. Maximum and minimum return losses for an offset short-circuit load, as calculated through Eqs. (16) and (17), are shown in Table 2.

It is of interest to study the characteristics of this perforated plate material when terminated in an offset short-circuit load and the offset distance is allowed to vary. Return losses calculated from Eq. (1) are plotted in Fig. 2. A sharp resonance is caused by the resistive losses of the perforated plate, and the maximum return loss is 33.2 dB. If the perforated plate were lossless, then the return loss would have been 0 dB. The exact short-circuit load position corresponding to the 33.2-dB maximum return loss was calculated from equations in this article and is shown in Table 2. The resonance phenomenon shown in Fig. 2 was observed experimentally when a sample of the perforated plate test sample was inserted in the waveguide and terminated with a sliding short-circuit load. Knowledge of maximum return loss can be used to obtain information on the resistive losses of the perforated plate. Similar techniques can be applied to the determination of resistive losses of other types of meshes such as a tricot-mesh material used for spacecraft unfurlable antennas [4,5].

The equations presented in this article are not limited to short-circuit loads, but are applicable for any load whose reflection-coefficient magnitude is greater than zero. For interest, Table 2 tabulates the conditions under which maximum and minimum return losses occur when the described perforated plate is terminated with loads having various load reflection-coefficient magnitudes.

B. Water Film on the Reflector Surface

The manufacturing process of some commercial Ku-band reflectors involves pouring a fiberglass epoxy mix onto a thin metallic screen laid on a parabolic-shaped mold. After drying, the reflector consists of a fiberglass layer bonded to the metal screen. It will be shown that a fiberglass layer thickness of 1.59 mm (0.0625 in.), which sometimes results in the manufacturing process, leads to poor reflector performance at 12 GHz when the surface becomes wet due to condensation or rain.

For theoretical calculation purposes, it will be assumed that the fiberglass dielectric material has a relative permittivity of 4.5 and loss tangent of 0.002 at 12 GHz. First, the S-parameters of the fiberglass sheet of 1.59-mm thickness are calculated from Eqs. (29) and (30) for a normally incident, linearly polarized wave at 12 GHz. Then, the input reflection coefficient is calculated from Eq. (1) for the fiberglass sheet terminated in a short-circuit load. The result is an input reflection coefficient of 0.9994 with a phase angle of 123.9 deg when this fiberglass sheet is terminated with a short-circuit load. The corresponding return loss value of 0.005 dB calculated from Eq. (17) shows that this reflector with the described fiberglass layer is an excellent reflector under dry conditions.

For purposes of analyzing the wet surface condition, a thin water film will be treated as a two-port network whose S-parameters are calculated from Eqs. (29) and (30) for water-film thicknesses varying between 0.05 mm (0.002 in.) and 0.10 mm (0.004 in.). These thicknesses are typical for a water film lying on a flat, smooth surface. A relative permittivity of 55.4 and loss tangent of 0.637 are used for water at 20 deg C and 12 GHz. These values for water¹ agree closely with those published in [6,7]. The resulting calculated S-parameters of water films are shown in Table 3 for normal incidence at 12 GHz.

Next, the water film terminated by the fiberglass-screen reflector is considered. Through the use of S-parameters shown in Table 3, and a load with a reflection-coefficient magnitude of 0.9994 and phase angle of 123.9 deg, the input reflection coefficient at the surface of the water film is calculated from Eq. (1). The results show that the return losses are unexpectedly high. For example, for the water-layer thickness of 0.10 mm (0.004 in.), the return loss is 12.8 dB. This result is consistent with experimental observation of about 10-dB signal losses when the fiberglass surfaces became wet from rain.²

To study the loss behavior in more detail, the return losses are shown plotted in Fig. 3 for water film terminated by a load with a reflection-coefficient magnitude of 0.9994 and a variable phase. Note that for the water-film thickness of 0.1 mm, the return loss is about 13 dB when the load phase angle is about 124 deg, but the actual peak of the resonance curve is at a lower phase angle and is about 25 dB. The exact values for the peaks of the curves for the various water-film thicknesses were calculated from Eq. (18) and are shown in Table 4. The high return losses result from the position of the maximum E-field having shifted to the inside of the film of water, which is very dissipative.

The values for minimum return losses are also tabulated in Table 4, as well as the fiberglass-layer thicknesses that correspond to maximum and minimum return losses. It is of interest to note that to obtain minimum return loss, the thickness of the fiberglass layer should be 5.84 mm (0.230 in.). The corresponding load phase angle (represented by the fiberglass-screen reflector) would then be close to 180 deg. Except for small losses due to the dielectric, this condition would be equivalent to the reflector

¹ S. J. Keihm, private communication, Microwave Observational Systems Section, Jet Propulsion Laboratory, Pasadena, California, September 1991.

² D. A. Bathker, private communication, Ground Antennas and Facilities Engineering Section, Jet Propulsion Laboratory, Pasadena, California, July 1989.

having no dielectric layer, where the rain film would lie directly on the metal surface. This result of a minimum return loss is consistent with experimental observations that very little degradation or signal loss occurs when rainwater wets a solid metallic reflector surface.

If the reflector were manufactured with a fiberglass-layer thickness of 5.84 mm (0.230 in.), rather than the current 1.59 mm (0.0625 in.), the return loss in dry conditions for the new thickness would be degraded only to 0.026 dB as compared to 0.005 dB for the original 1.59-mm thickness. However, if the fiberglass surface were to become wet with water thicknesses varying from 0.05 to 0.10 mm, the return loss would still be the 0.026-dB worst case as compared to 12.8 dB for a fiberglass-layer thickness of 1.59 mm.

The equations used for this study of a fiberglass-layered reflector can also be applied for performance studies of graphite-epoxy-layered reflector surfaces being proposed for future ground-space communication antennas. It is important that the thicknesses of the dielectric layer be

designed properly so as to obtain good performance in wet conditions due to rain or condensation.

IV. Conclusions

The analytical equations derived in this article are useful for determining worst-case return losses that can occur when a passive two-port network is terminated with a mismatched load. In one of the examples presented, the equations were useful for predicting worst-case return losses that could occur when water film was on a reflector surface that had a dielectric layer. The same type of analysis might be applied to a similar type of reflector painted with thermal diffusive paint suspected of being very lossy.

Although it was not shown through a specific example, for ($|\Gamma_T| < 1$), the equations presented in this article can also be used to design dielectric- or air-line spacers that can be inserted between a fixed load ($|\Gamma_T| < 1$) and the network if the goal is to obtain maximum or minimum return loss from the terminated two-port network.

Acknowledgments

The Ku-band reflector antenna resonance phenomenon in rainy conditions was brought to the author's attention by D. Bathker of the Ground Antennas and Facilities Engineering Section. Discussions with him motivated the studies and derivation of the equations presented in this article. A "Dielectric Sheet" computer program written in 1972 by Dr. C. Yeh, a consultant to JPL, was helpful in confirming the general case S-parameter equations presented in this article. His derivations were based on free-space considerations, rather than the network theory approach used in this article.

References

- [1] D. M. Kerns and R. W. Beatty, *Basic Theory of Waveguide Junctions and Introductory Microwave Network Analysis*, New York: Pergamon, 1967.
- [2] R. L. Fante and M. T. McCormack, "Reflection Properties of the Salisbury Screen," *IEEE Trans. on Antennas and Propagation*, vol. 36, no. 10, pp. 1443-1454, October 1988.
- [3] T. Y. Otoshi, "Precision Reflectivity Loss Measurements of Perforated-Plate Mesh Materials by a Waveguide Technique," *IEEE Trans. on Instrumentation and Measurement*, vol. IM-21, pp. 451-457, November 1972.

- [4] K. Woo and T. Y. Otoshi, "Further RF Study of Reflector Surface Materials for Spacecraft Antennas," *Space Programs Summary No. 36-65*, vol. III, Jet Propulsion Laboratory, Pasadena, California, pp. 47-52, October 31, 1970.
- [5] W. A. Imbriale, V. Galindo-Israel, and Y. Rahmat-Samii, "On the Reflectivity of Complex Mesh Surfaces," *IEEE Trans. on Antennas and Propagation*, vol. 39, no. 9, pp. 1352-1365, September 1991.
- [6] D. S. Engelder and C. R. Buffler, "Measuring Dielectric Properties of Food Products at Microwave Frequencies," *Microwave World*, vol. 12, no. 2, Summer 1991.
- [7] H. J. Liebe, G. A. Hufford, and T. Manabe, "A Model for the Complex Permittivity of Water at Frequencies Below 1 THz," *International Journal of Infrared and Millimeter Waves*, vol. 12, pp. 664, Table 1, November 7, 1991. (Note: 1 THz = 1000 GHz.)

Table 1. Measured S-parameters of the 64-m diameter antenna perforated panel test sample at 8.448 GHz, with perpendicular polarization, 38.5-deg incidence angle, and $\phi = 0$ deg.

S-parameter	Magnitude	Phase, deg
S_{11}	0.9990	172.4
S_{21}	0.0305	82.2

Notes: 1. $S_{12} = S_{21}$ and $S_{22} = S_{11}$.

2. S-parameters are defined for free-space characteristic impedance Z_{01} at the input and output ports.

3. Perforated-plate hole diameter, hole-to-hole spacing, with thicknesses of 4.763 mm (0.1875 in.), 6.350 mm (0.250 in.), and 2.286 mm (0.090 in.).

Table 2. Load reflection coefficient phase angle for minimum and maximum return losses for the 64-m diameter antenna perforated plate (see Table 1).

Γ_L magnitude	Γ_L phase, deg	Return loss, dB	Comments
1.0	-172.40	33.21	This maximum return loss occurs when the distance from a short = $0.4894 \lambda_g^a$
1.0	-4.00	0.0040	This minimum return loss occurs when the distance from a short = $0.2556 \lambda_g$
0.5	-172.27	0.1641	Maximum return loss
0.5	8.80	0.0056	Minimum return loss
0.1	-172.07	0.0092	Maximum return loss
0.1	8.09	0.0076	Minimum return loss
0.0	All values	0.0087	$20 \log_{10} S_{11} $

$$^a \lambda_g = \frac{\lambda_0}{\sqrt{1 - \sin^2 \theta_i}} = 1.785''$$

Table 3. Calculated S-parameters of water film for relative permittivity = 55.4, loss tangent = 0.637, at 12 GHz, and with normal incidence angle.

S-parameter	0.05-mm thick		0.076-mm thick		0.10-mm thick	
	Magnitude	Phase, deg	Magnitude	Phase, deg	Magnitude	Phase, deg
S_{11}	0.3249	-139.48	0.4319	-145.23	0.5137	-149.82
S_{12}	0.7855	-16.56	0.6975	-22.38	0.6239	-27.07

Notes: 1. $S_{12} = S_{21}$ and $S_{22} = S_{11}$.

2. S-parameters are defined for free-space characteristic impedance Z_{01} at the input and output ports (see Fig. 1).

Table 4. Fiberglass plus screen load reflection coefficients and corresponding return losses for various water-layer thicknesses whose S-parameters are shown in Table 3.

Fiberglass thickness, mm	Fiberglass thickness, in.	Γ_L magnitude	Γ_L phase, deg	Return loss (no water layer), dB	Water layer thickness, mm	Return loss (with water layer), dB
1.59	0.0625	0.99940	123.86	0.0052	0.050	3.24
1.59	0.0625	0.99940	123.86	0.0052	0.076	6.69
1.59	0.0625	0.99940	123.86	0.0052	0.100	12.77
2.344	0.0923	0.99700	70.22	0.0261	0.050	8.46
2.075	0.0817	0.99824	93.37	0.0153	0.076	14.20
1.836	0.0723	0.99895	109.83	0.0091	0.100	24.92
5.865	0.2309	0.99704	-179.32	0.0257	0.050	0.026
5.847	0.2302	0.99704	-178.81	0.0257	0.076	0.026
5.834	0.2297	0.99704	-178.44	0.0257	0.100	0.026

Note: Γ_L is defined for free-space characteristic impedance Z_{01} (see Fig. 1).

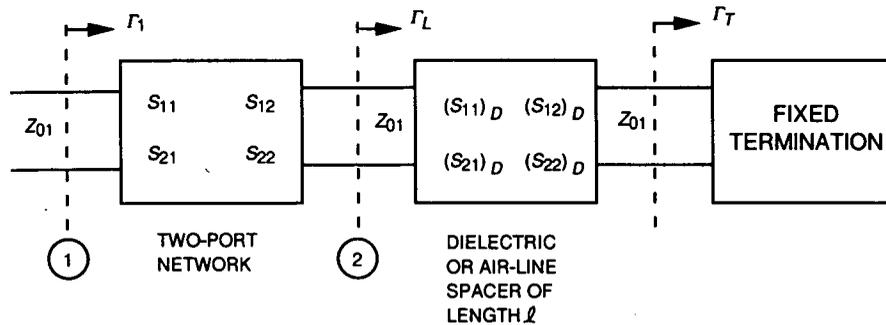


Fig. 1. Two-port network terminated in a Γ_L load.

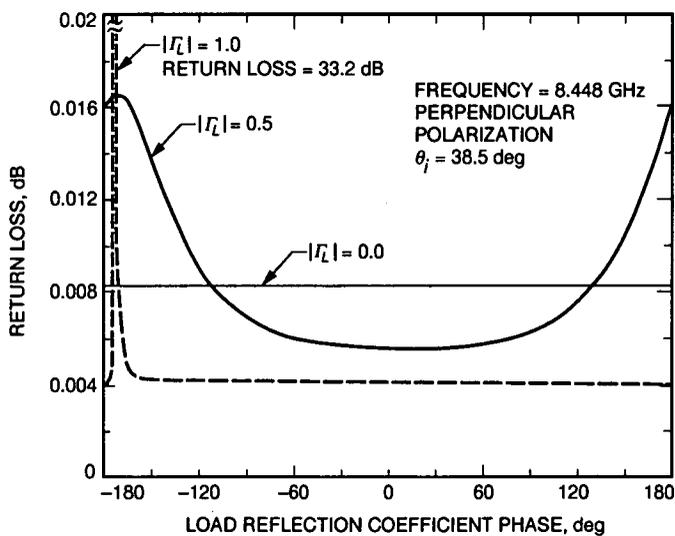


Fig. 2. Return loss of a perforated plate as a function of load-reflection coefficient phase angle.

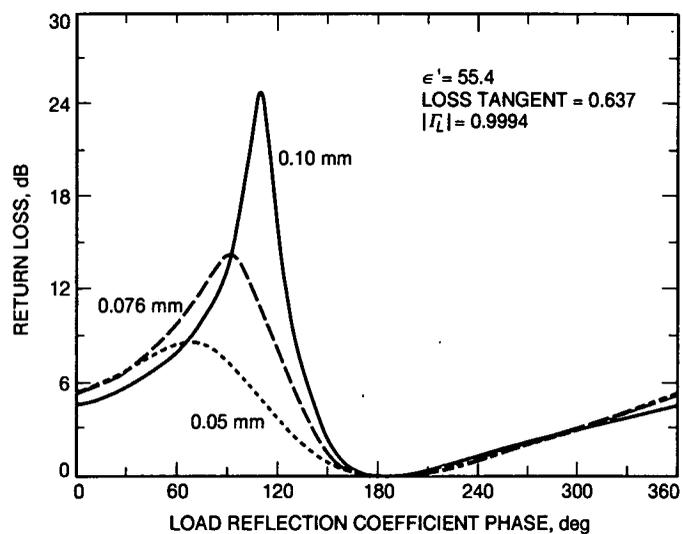


Fig. 3. Return loss versus load phase angle for various thicknesses of water films terminated in a load having a reflection coefficient magnitude of 0.9994 at 12 GHz.

500728

59-61

167929

p. 10

The General Theory of Convolutional Codes

R. J. McEliece¹

Communications Systems Research Section

R. P. Stanley²

This article presents a self-contained introduction to the algebraic theory of convolutional codes, which is partly tutorial, but at the same time contains a number of new results which will prove useful for designers of advanced telecommunication systems. Among the new concepts introduced here are the Hilbert series for a convolutional code and the class of compact codes.

I. Introduction

Convolutional codes have played a central part in NASA's deep-space telecommunications systems for many years. In all such applications to date, the convolutional codes have been codes of *dimension 1*, which are commonly, but not strictly correctly, referred to as "rate $1/n$ " codes. However, as systems become more sophisticated, the coding subsystems must keep pace, and this article is an outline of algebraic theory for the most general class of convolutional codes known, the so-called " (n, k, m) " codes, of which the usual "rate $1/n$ " codes form the special case $k = 1$. Much of this theory was originally developed by Forney [1-4], but this article adds to what is already known, as well as placing many of the older results into a modern, "system-theoretic" context. In particular, introduced here for the first time is the "Hilbert series" for a convolutional code, which is a generating function from which the dimensions of certain polynomial subcodes can be easily computed in terms of the "Forney indices" of the code. The Forney indices provide a derivation of an upper

bound on the free distance of a convolutional code which in some cases improves the bounds previously known, and whose derivation makes no use of the structure of any particular encoder structure. Finally, the notion of "compact" and "noncompact" convolutional codes is introduced, and it is argued that only compact codes are likely to be interesting for applications.

II. Convolutional Codes: Polynomial Generator Matrices

Let F be a field, usually $GF(2)$, and let $F(D)$ be the field of rational functions over F . An (n, k) convolutional code over F is a k -dimensional subspace of $F(D)^n$. The elements of the code are called its *codewords*. A codeword is thus an n -tuple of rational functions over $F(D)$. The *weight* of a codeword is defined to be the sum of the weights of its components, where the weight of a component (i.e., rational function) is the number of nonzero coefficients in its expansion as a Laurent series in increasing powers of D . The *free distance* of a convolutional code is defined to be the minimum nonzero weight of any codeword.

¹ Consultant, California Institute of Technology, Engineering Department.

² Consultant, Massachusetts Institute of Technology, Department of Mathematics.

If C is an (n, k) convolutional code over F , a *generator matrix* $G(D)$ for C is a $k \times n$ matrix over $F(D)$ whose rows form a basis for C . If the entries of $G(D)$ are polynomials, then $G(D)$ is called a *polynomial generator matrix* (PGM) for C . Any convolutional code has a polynomial generator matrix, since if G is an arbitrary generator matrix for C , the matrix obtained from G by multiplying each row by the least common multiple of the denominators of the entries in that row is a PGM for C .

Let $G(D) = (g_{ij}(D))$ be a $k \times n$ PGM for C . The i th row of G , i.e., the n -vector (g_{i1}, \dots, g_{in}) , is denoted by g_i , and the *degree* of g_i is defined as the maximum degree of its components. In a similar way, the degree of any n -tuple of polynomials is defined as the maximum degree of any component. The *internal degree* and *external degree* of $G(D)$ are defined as follows:

$$\text{int.deg. } G(D) = \text{maximum degree of } G(D)\text{'s } k \times k \text{ minors}$$

$$\text{ext.deg. } G(D) = \text{sum of the row degrees of } G(D)$$

The following two definitions will be essential in the discussion of convolutional codes.

A. Definition 1

A $k \times n$ polynomial matrix $G(D)$ is called *basic* if, among all polynomial matrices of the form $T(D)G(D)$, where $T(D)$ is a nonsingular $k \times k$ matrix over $F(D)$, it has the minimum possible internal degree.

B. Definition 2

A $k \times n$ polynomial matrix $G(D)$ is called *reduced* if, among all matrices of the form $T(D)G(D)$, where $T(D)$ is unimodular,³ $G(D)$ has the minimum possible external degree. Since any unimodular matrix is a product of elementary matrices, an equivalent definition is that a matrix

is reduced if its external degree cannot be reduced by a sequence of elementary row operations.

Before continuing along the main line, it is helpful to present a simple theorem that provides several useful facts about the internal and external degrees of a polynomial matrix.

Theorem 1. Let $G(D)$ be a $k \times n$ polynomial matrix.

- (1) If $T(D)$ is any nonsingular $k \times k$ polynomial matrix, then $\text{int.deg. } T(D)G(D) = \text{int.deg. } G(D) + \text{deg.det. } T(D)$. In particular $\text{int.deg. } T(D)G(D) \geq \text{int.deg. } G(D)$, with equality if and only if $T(D)$ is unimodular.
- (2) $\text{int.deg. } G(D) \leq \text{ext.deg. } G(D)$.

Proof:

- (1) The $k \times k$ submatrices of $T(D)G(D)$ are just the $k \times k$ submatrices of $G(D)$, each multiplied by $T(D)$. Thus the $k \times k$ minors of $T(D)G(D)$ are just the $k \times k$ minors of $G(D)$, each multiplied by $\det T(D)$. The result now follows.
- (2) Denote the degree of the i th row of $G(D)$ by e_i . In the expansion of any $k \times k$ minor of $G(D)$, each term is the product of k entries of $G(D)$, one from each row (and column). Since each entry from the i th row has degree $\leq e_i$, it follows that the degree of any $k \times k$ minor is at most $e_1 + \dots + e_k = \text{ext.deg. } G(D)$. \square

Basic and reduced polynomial matrices enjoy many useful and surprising properties. The Appendix gives two theorems, Theorem A-1 and Theorem A-2, delineating these properties. These Theorems will be referenced constantly in the rest of this article.

Example 1. Here are eight generator matrices for a $(4, 2)$ convolutional code over $GF(2)$. Of these eight, six, G_2 through G_7 , are PGMs.

$$G_1 = \begin{pmatrix} \frac{1}{1+D+D^2} & 1 & \frac{1+D^2}{1+D+D^2} & \frac{1+D}{1+D+D^2} \\ 1 & \frac{1+D+D^2}{D} & D & \frac{1}{D} \end{pmatrix}$$

$$G_2 = \begin{pmatrix} 1 & 1+D+D^2 & 1+D^2 & 1+D \\ D & 1+D+D^2 & D^2 & 1 \end{pmatrix}$$

³ A unimodular matrix is a square polynomial matrix whose determinant is a nonzero scalar.

$$G_3 = \begin{pmatrix} 1 & 1+D+D^2 & 1+D^2 & 1+D \\ 0 & 1+D & D & 1 \end{pmatrix}$$

$$G_4 = \begin{pmatrix} 1 & D & 1+D & 0 \\ 0 & 1+D & D & 1 \end{pmatrix}$$

$$G_5 = \begin{pmatrix} 1+D & 0 & 1 & D \\ D & 1+D+D^2 & D^2 & 1 \end{pmatrix}$$

$$G_6 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1+D & D & 1 \end{pmatrix}$$

$$G_7 = \begin{pmatrix} 1+D & 0 & 1 & D \\ 1 & D & 1+D & 0 \end{pmatrix}$$

$$G_8 = \begin{pmatrix} 1 & 0 & \frac{1}{1+D} & \frac{D}{1+D} \\ 0 & 1 & \frac{D}{1+D} & \frac{1}{1+D} \end{pmatrix}$$

Table 1 lists the properties of the generator matrices G_1, \dots, G_8 . (Note that only the polynomial generator matrices, viz., G_2 – G_7 , have external or internal degrees, or can be basic or reduced.) These properties are easily verified by referring to Theorems A-1 and A-2. For example, G_3 is *basic* because the gcd of its 2×2 minors is 1 (Theorem A-1, condition (2)), and it is *not reduced* because its internal and external degrees are unequal (Theorem A-2, condition (2)). \square

It follows from Definition 1 that among all PGMs for a given convolutional code, those for which the *internal* degree is as small as possible are exactly the *basic* PGMs. It turns out, however, that the set of PGMs for which the *external* degree is as small as possible form a much more interesting class, the class of *minimal* PGMs.

C. Definition 3

Among all PGMs for a given convolutional code C , those for which the external degree is as small as possible are called *minimal* PGMs. This minimal external degree is called the *degree* of the code C , and is denoted $\deg C$.

It will be shown that minimal generator matrices have many remarkable properties. The key to these properties is the following theorem:

Theorem 2. A PGM $G(D)$ for the convolutional code C is minimal if and only if it is both basic and reduced.

Proof: First it will be shown that a minimal PGM must be both basic and reduced. Then it will be shown that a PGM that is both basic and reduced must be minimal.

To prove the first assertion, denote by m_0 the common internal degree of all the basic PGMs for C , and among all the basic PGMs choose one, say $G_0(D)$, for which the *external* degree is as small as possible. Then G_0 must be reduced, since if $T(D)$ is unimodular, $\text{int.deg. } T(D)G_0(D) = \text{int.deg. } G_0(D) = m_0$ by Theorem 1, and so by the definition of G_0 , $\text{ext.deg. } T(D)G_0(D) \geq \text{ext.deg. } G_0(D)$. Now let $G(D)$ be any minimal PGM for C . Then

$$\text{int.deg. } G_0 \leq \text{int.deg. } G \leq \text{ext.deg. } G \leq \text{ext.deg. } G_0 \quad (1)$$

(The first inequality in Definition 1 follows from the fact that G_0 is chosen to have minimum possible internal degree. The second inequality follows from Theorem 1(2). The third inequality is because G , as a minimal PGM, has minimum possible external degree.) But since $G_0(D)$ is reduced, by Theorem A-2, condition (2), $\text{int.deg. } G_0 = \text{ext.deg. } G_0$, so that equality holds throughout Definition 1. Thus $\text{int.deg. } G = \text{int.deg. } G_0 = m_0$, so G is basic; and $\text{int.deg. } G = \text{ext.deg. } G$, so that G is reduced, by Theorem A-2, condition (2).

Conversely, suppose that $G(D)$ is basic and reduced, and $G_0(D)$ is any other PGM for C . Then by Theorem 1(2), $\text{ext.deg. } G_0(D) \geq \text{int.deg. } G_0(D)$. Since $G(D)$ is basic, $\text{int.deg. } G_0(D) \geq \text{int.deg. } G(D)$; since $G(D)$ is reduced, by Theorem 1 $\text{int.deg. } G(D) = \text{ext.deg. } G(D)$. Combining these inequalities, $\text{ext.deg. } G_0(D) \geq \text{ext.deg. } G(D)$, which proves that $G(D)$ is minimal. \square

In the proof of Theorem 2, m_0 , the common internal degree for all PGMs for C , is equal to $\deg C$, i.e., the minimum possible external degree. Thus there are two corollaries to Theorem 2.

Corollary 1. The minimal internal degree of any PGM for a given convolutional code C is equal to the degree of C .

Corollary 2. If G is any basic generator matrix for C , then $\text{int.deg. } G = \deg C$.

The following theorem shows that minimal generator matrices are "minimal" in a very strong sense.

Theorem 3. If $e_1 \leq e_2 \leq \dots \leq e_k$ are the row degrees of a minimal generator matrix for a convolutional code C , and if $f_1 \leq f_2 \leq \dots \leq f_k$ are the row degrees of any other polynomial generator matrix, say G' , for C , then $e_i \leq f_i$, for $i = 1, \dots, k$.

Proof: If the statement is false, there exists an index j such that $e_1 \leq f_1, \dots, e_j \leq f_j$, but $e_{j+1} > f_{j+1}$. It then follows from the minimality of G (use the properties in Theorem A-1, condition (5), and Theorem A-2, condition (3)) that the first $j+1$ rows of G' must be polynomial linear combinations of the first j rows of G , which contradicts the fact that the rows of G' are linearly independent. \square

Theorem 4. The set of row degrees is the same for all minimal PGMs for a given code.

Proof: This result follows immediately from Theorem 3. \square

The row degrees referred to in Theorems 3 and 4, say (e_1, e_2, \dots, e_k) , are called the *Forney indices* of the code. The sum $e_1 + \dots + e_k$ of the Forney indices, i.e., the minimum possible external degree of any PGM for C , is the degree of the code. The maximum of the Forney indices is called the *memory* of the code. From now on, reserve the letter m to denote the degree of a given convolutional code, and refer to an (n, k) code with degree m as an (n, k, m) code. An (n, k, m) code is called *optimal* if it has the maximum possible free distance among all codes with the same value of n, k , and m .

Example 2: Continuing the study of the $(4, 2)$ code from Example 1, of the eight given generator matrices, only G_6 is minimal (it satisfies condition (2) of Theorem A-1 and condition (2) of Theorem A-2, so it is both basic and reduced), so that the *degree* of the code C is 1, and the *Forney indices* are $(0, 1)$. The code is thus a $(4, 2, 1)$ code. From Example 3, below, it is in fact an *optimal* $(4, 2, 1)$ code. \square

III. The Hilbert Series and Free Distance Bounds

If C is a fixed (n, k) convolutional code, a *polynomial codeword* of C is a codeword all of whose components are polynomials. Recalling that the degree of a polynomial vector is defined to be the maximum degree of any component, for any integer $L \geq 0$, C_L is defined as the set of *polynomial codewords of degree $\leq L$* . C_L is a vector space over F . Indeed, it is a subspace of the set of all possible

n -dimensional polynomial vectors of degree $\leq L$ over F . The F -dimension of C_L is denoted by δ_L . The following theorem shows that the δ_L 's can be computed from the Forney indices.

Theorem 5. If C is an (n, k) convolutional code with Forney indices (e_1, \dots, e_k) and polynomial subcode dimensions δ_L , then

Note: The power series appearing in Theorem 5 is called the Hilbert series for the code.

Proof: Let $G(D)$ be a minimal PGM for C , whose row degrees are the ordered Forney indices, say $e_1 \leq \dots \leq e_k$, and let g_1, \dots, g_k be the rows of G . Let $y(D)$ be any polynomial codeword of degree $\leq L$. Then it follows from Theorem A-1, condition (5), that $y(D) = x(D)G(D)$, where $x(D) = (x_1(D), \dots, x_k(D))$ is a k -vector of polynomials, and it follows from the predictable degree property (Theorem A-2, condition (3)) that $\deg x_i + e_i \leq L$. Thus a basis for the F -space C_L is the set $\{D^j g_i(D) : j + e_i \leq L\}$. Hence

$$\begin{aligned} \sum_{L \geq 0} \delta_L t^L &= \sum_{i=1}^k \sum_{j \geq 0} (t^{e_i+j} + t^{e_i+j+1} + \dots) \\ &= \sum_{i=1}^k \sum_{j \geq 0} \frac{t^{e_i+j}}{1-t} \\ &= \sum_{i=1}^k \frac{t^{e_i}}{(1-t)^2} \quad \square \end{aligned}$$

Corollary 3. The following explicit formula for δ_L holds:

$$\delta_L = \sum_{i=1}^k \max(L + 1 - e_i, 0) \quad (3)$$

Proof: By elementary calculus, $(1-t)^{-2} = \sum_{j \geq 0} (j+1)t^j$. Applying this fact to Eq. (2),

$$\begin{aligned} \sum_{L \geq 0} \delta_L t^L &= \sum_{i=1}^k \frac{t^{e_i}}{(1-t)^2} \\ &= \sum_{i=1}^k \sum_{j \geq 0} (j+1)t^{e_i+j} \\ &= \sum_{i=1}^k \sum_{j \geq e_i} (j+1 - e_i)t^j \end{aligned}$$

Thus the coefficient of t^L in the Hilbert series $\sum_{L \geq 0} \delta_L t^L$ is $\sum_{i=1}^k \max(L+1-e_i, 0)$, which is the desired proof. \square

Corollary 4. Let C be an (n, k, m) convolutional code. Then for all $L \geq 0$

$$\delta_L \geq \max((L+1)k - m, 0) \quad (4)$$

Furthermore, there is equality for all $L \geq 0$ in Eq. (4) if and only if the Forney indices assume only the two values $\lfloor m/k \rfloor$ and $\lceil m/k \rceil$. A code for which this is true will be called a *compact* code.

Proof: Since $\max(x, 0) \geq x$, Eq. (3) implies that $\delta_L \geq \sum_{i=1}^k (L+1-e_i) = (L+1)k - \sum_{i=1}^k e_i = (L+1)k - m$. Since $\delta_L \geq 0$, too, it follows that $\delta_L \geq \max((L+1)k - m, 0)$ for all $L \geq 0$. Assuming that the Forney indices are ordered so that $e_1 \leq \dots \leq e_k$, it follows from Eq. (3) that $\delta_L = (L+1)k - m = \max((L+1)k - m, 0)$ if $L+1 \geq e_k$, and $\delta_L = 0 = \max((L+1)k - m, 0)$ if $L+1 \leq e_1$. If $e_k - e_1 \leq 1$, one of these alternatives must hold for all L . On the other hand, if $e_k - e_1 \geq 2$, then there is at least one value of L for which $e_1 < L+1 < e_k$, in which case $0 < \delta_L < (L+1)k - m$, so that $\delta_L \neq \max((L+1)k - m, 0)$. \square

Incidentally, it is easy to show that in a compact code, there are exactly $(m \bmod k)$ Forney indices equal to $\lfloor m/k \rfloor$, and $k - (m \bmod k)$ Forney indices equal to $\lceil m/k \rceil$. Thus for example if $k = 4$ and $m = 13$, a compact code will have Forney indices $(3, 3, 3, 4)$.

Since the subcode C_L forms an $(n(L+1), \delta_L)$ linear block code over F , the *free distance* of C cannot exceed the *minimum distance* of C_L , for $L = 0, 1, \dots$, which leads to Theorem 6.

Theorem 6. If C is an (n, k) convolutional code with Forney indices (e_1, \dots, e_k) , then

$$d_{\text{free}}(C) \leq \min_{L \geq 0} \Delta_F(n(L+1), \delta_L)$$

where $\Delta_F(n, k)$ denotes the maximum possible minimum distance of an (n, k) linear block code over F .

Note from Corollary 4 that δ_L is minimized for all L , and so $\Delta_F(n(L+1), \delta_L)$ is maximized for all L , for a compact code. This suggests, but does not prove, that among all (n, k, m) codes, the compact codes will have the largest free distances. In any case, for applications a bound on

d_{free} is needed that applies to all (n, k, m) codes, regardless of their Forney indices. Thus is offered the following corollary to Theorem 6, which gives an upper bound on d_{free} for all (n, k, m) codes, regardless of the Forney indices. It is good to bear in mind, however, that it may be possible to improve the bound if the code is noncompact.

Corollary 5. If C is an (n, k, m) code, then

$$d_{\text{free}}(C) \leq \min_{L \geq 0} \Delta_F(n(L+1), k(L+1) - m)$$

Proof: Combine Theorem 6 with Corollary 4. \square

The upper bound on d_{free} of Theorem 6 is attained for many, but not all, values of n, k , and m . The following examples will illustrate this. (All examples in this article are over the field $GF(2)$.)

Example 3. Continuing the study of the $(4, 2, 1)$ code in Examples 1 and 2, since the Forney indices are $(0, 1)$, by Theorem 5 the ‘‘Hilbert Series’’ for the code is

$$\frac{1+t}{(1-t)^2} = 1 + 3t + 5t^2 + \dots + (2L+1)t^L + \dots$$

Thus the dimension of the L th subcode C_L is $2L+1$. It then follows from Theorem 6 that the free distance of the code satisfies

$$d_{\text{free}}(C) \leq \min_{L \geq 0} \Delta_{GF(2)}(4(L+1), 2L+1)$$

In particular, for $L = 0$ the above bound gives $d_{\text{free}}(C) \leq \Delta_2(4, 1) = 4$. But in fact $d_{\text{free}} = 4$ for this code (use the generator matrix G_4 to check this fact), so this particular code has the largest possible free distance for a $(4, 2, 1)$ convolutional code, i.e., it is an *optimal* $(4, 2, 1)$ code. \square

Example 4. Consider $(n, k, m) = (2, 1, 2)$ codes. Since $k = 1$, there is only one Forney index, and so any $(2, 1, 2)$ code is compact. By Corollary 3, $\delta_0 = 0$, and, for $L \geq 1$, $\delta_L = L - 1$, so that by Theorem 6 (here and hereafter Verhoeff’s tables [7] of the values of $\Delta_F(n, k)$ are used when $F = GF(2)$),

$$\begin{aligned} d_{\text{free}} &\leq \min(\Delta(2, 0), \Delta(4, 0), \Delta(6, 1), \Delta(8, 2), \Delta(10, 3) \dots) \\ &= \min(\infty, \infty, 6, 5, 5, \dots) \\ &= 5 \end{aligned}$$

In fact, there is a well-known $(2, 1, 2)$ code with $d_{\text{free}} = 5$, whose (unique) minimal generator matrix is

$$G(D) = (1 + D^2 \quad 1 + D + D^2) \quad (5)$$

(See [6], Chapter 9.) It follows then that the $(2, 1, 2)$ code defined by Theorem 6 is optimal. \square

Example 5: Consider binary $(4, 3, 2)$ codes. The Forney indices of such codes must be either $(0, 1, 1)$ (compact) or $(0, 0, 2)$ (noncompact). In the first case, by Corollary 3, $\delta_0 = 1$, $\delta_1 = 4$, $\delta_2 = 7$, etc., and so by Theorem 6,

$$\begin{aligned} d_{\text{free}} &\leq \min(\Delta(4, 1), \Delta(8, 4), \Delta(12, 7), \Delta(16, 10), \dots) \\ &= \min(4, 4, 4, \dots) \\ &= 4 \end{aligned}$$

However, it turns out that there is no $(4, 3, 2)$ code with $d_{\text{free}} = 4$ [9]. The largest possible d_{free} turns out to be $d_{\text{free}} = 3$, which is achieved by the second-order Wyner-Ash code [8]. A minimal PGM for such a code is

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & D \\ D & 0 & 1 & 1 \end{pmatrix}$$

On the other hand, if the Forney indices are $(0, 0, 2)$, then Corollary 3 tells one that $\delta_0 = 2$, $\delta_1 = 4$, $\delta_2 = 7$, etc., and so by Theorem 6,

$$\begin{aligned} d_{\text{free}} &\leq \min(\Delta(4, 2), \Delta(8, 4), \Delta(12, 7), \Delta(16, 10), \dots) \\ &= \min(2, 4, 4, \dots) \\ &= 2 \end{aligned}$$

And indeed there is a $(4, 3, 2)$ code with Forney indices $(0, 0, 2)$ and $d_{\text{free}} = 2$. A minimal PGM for one such code is

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 + D^2 \end{pmatrix}$$

Therefore, among $(4, 3, 2)$ codes, only the compact ones can be optimal. \square

References

- [1] G. D. Forney, Jr., "Convolutional Codes I: Algebraic Structure," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 720-738, November 1970.
- [2] G. D. Forney, Jr., "Structural Analysis of Convolutional Codes via Dual Codes," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 512-518, July 1973.
- [3] G. D. Forney, Jr., "Minimal Bases of Rational Vector Spaces with Applications to Multivariable Linear Systems," *SIAM J. Control*, vol. 13, pp. 493-502, May 1975.
- [4] G. D. Forney, Jr., "Algebraic Structure of Convolutional Codes, and Algebraic System Theory," in *Mathematical System Theory*, edited by A. C. Antoulas, Berlin, Germany: Springer-Verlag, pp. 527-557, 1991.
- [5] T. Kailath, *Linear Systems*. Englewood Cliffs, New Jersey: Prentice Hall, 1980.
- [6] R. McEliece, *The Theory of Information and Coding*, Reading, Massachusetts: Addison-Wesley, 1977.
- [7] T. Verhoeff, "An Updated Table of Minimum-Distance Bounds for Binary Linear Codes," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 665-680, September 1987.
- [8] A. D. Wyner and R. B. Ash, "Analysis of Recurrent Codes," *IEEE Trans. Inform. Theory*, vol. IT-9, pp. 143-156, July, 1963.
- [9] Ø. Ytrehus, "A Note on High Rate Binary Convolutional Codes," in *Report in Informatics No. 68*, Bergen, Norway: University of Bergen, August 1992.

Appendix

Basic and Reduced Matrices

This appendix—a reference collection of many useful properties of basic and reduced matrices—begins with the basic matrices.

Theorem A-1. A $k \times n$ polynomial matrix $G(D)$ is basic (see Definition 1) if and only if any one of the following six conditions is satisfied:

- (1) The invariant factors of $G(D)$ are all 1.
- (2) The gcd of the $k \times k$ minors of $G(D)$ is 1.
- (3) $G(\alpha)$ has rank k for any α in the algebraic closure of F .
- (4) $G(D)$ has a right $F[D]$ inverse, i.e., there exists an $n \times k$ polynomial matrix $H(D)$ such that $G(D)H(D) = I_k$.
- (5) If $y(D) = x(D)G(D)$, and if $y(D) \in F[D]^n$, then $x(D) \in F[D]^k$. (“Polynomial output implies polynomial input.”)
- (6) $G(D)$ is a submatrix of a unimodular matrix, i.e., there exists an $(n - k) \times n$ matrix $L(D)$ such that the $n \times n$ matrix $\begin{pmatrix} G(D) \\ L(D) \end{pmatrix}$ has determinant 1.

Proof: The proof is logically rather involved. The following implications will be proved: (Basic) \rightarrow (1) \rightarrow (2) \rightarrow (4) \rightarrow (5) \rightarrow (Basic); (2) \leftrightarrow (3); (1) \leftrightarrow (6).

• (Basic) \rightarrow (1): Suppose that Γ is the $k \times n$ invariant-factor form for G , i.e., $\Gamma = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_k)$ where $\gamma_i = \Delta_i/\Delta_{i-1}$, Δ_i being the gcd of the $i \times i$ minors of G . (Take $\Delta_0 = 1$ by convention.) Then there is a $k \times k$ unimodular matrix X and an $n \times n$ unimodular matrix Y such that

$$XGY = \Gamma \tag{A-1}$$

(For a proof of this “invariant factor decomposition,” see [5], Theorem 6.3.16.) Thus if Γ_k denotes the $k \times k$ matrix formed by the leftmost k columns of Γ , the matrix $G' = \Gamma_k^{-1}XG$ is a polynomial matrix equivalent to G . Furthermore, since $\det(\Gamma_k^{-1}X) = \det \Gamma_k^{-1} = (\det \Gamma_k)^{-1} = (\gamma_1 \cdots \gamma_k)^{-1}$, unless the γ_i 's are all 1, the internal degree of G' is strictly less than that of Γ . Thus if the invariant factors of G are not all 1, then G is not basic, which proves (Basic) \rightarrow (1).

• (1) \leftrightarrow (2): According to the definitions given in the previous paragraph, the product of the invariant factors of G is

$$\begin{aligned} \gamma_1 \gamma_2 \cdots \gamma_k &= \frac{\Delta_1}{\Delta_0} \cdot \frac{\Delta_2}{\Delta_1} \cdots \frac{\Delta_k}{\Delta_{k-1}} \\ &= \frac{\Delta_k}{\Delta_0} \\ &= \Delta_k = (\text{the gcd of the } k \times k \text{ minors of } G) \end{aligned}$$

Hence the g_i 's are all 1 if and only if $\Delta_k = 1$.

• (2) \rightarrow (4): Suppose that the gcd of the $k \times k$ minors of $G(D)$ is 1, and denote the individual minors by $\Delta_\nu(D)$, for $\nu = 1, 2, \dots, \binom{n}{k}$. Then by Cramer's rule for each ν , there will exist a “pseudo-inverse” for $G(D)$, with factor $\Delta_\nu(D)$, i.e., an $n \times k$ matrix $H_\nu(D)$ such that $G(D)H_\nu(D) = \Delta_\nu(D)I_k$. Since the gcd of the $\Delta_\nu(D)$'s is 1, there exists a polynomial linear combination of the $\Delta_\nu(D)$'s equal to 1, say $\sum_\nu \lambda_\nu(D)\Delta_\nu(D) = 1$. It follows that $H(D) = \sum_\nu \lambda_\nu(D)H_\nu(D)$ is an $n \times k$ polynomial inverse for $G(D)$.

• (4) \rightarrow (5): Suppose that $G(D)$ has an $n \times k$ polynomial inverse $H(D)$, and that $x(D) = (x_1(D), \dots, x_k(D))$ is a k -vector of rational functions such that $y(D) = x(D)G(D)$ is an n -vector of polynomials. Multiplying this equation on the right by $H(D)$, $y(D)H(D) = x(D)$, which implies that $x(D)$ is in fact a polynomial vector.

• (5) \rightarrow (Basic): Suppose that property (5) holds, and let $T(D)$ be an arbitrary nonsingular $k \times k$ matrix of rational functions such that $G' = TG$ is a polynomial matrix. Then by property (5), T must in fact be a polynomial matrix, so that by Theorem 1(1), $\text{int.deg. } G' \geq \text{int.deg. } G$, which means that G is basic.

• (2) \leftrightarrow (3): Suppose that the gcd of the $k \times k$ minors of $G(D)$ is 1, let α be an arbitrary element of the algebraic closure of F , and let $p(D)$ be the minimal polynomial of α . Then there must be at least one $k \times k$ subdeterminant of $G(D)$ which is not divisible by $p(D)$, which means that the corresponding $k \times k$ submatrix of $G(\alpha)$ is nonsingular. Thus $G(\alpha)$ must have rank k . Conversely, suppose that the gcd of the $k \times k$ minors of G is not 1, which means

that it is divisible by some irreducible polynomial $p(D)$. If α is a root of $p(D)$ in some extension field of F , it follows that every $k \times k$ minor of $G(\alpha)$ is zero, which in turn means that $G(\alpha)$ has rank less than k .

• (1) \leftrightarrow (6): Suppose the invariant factors of G are all 1. Then the invariant-factor decomposition in Eq. (A-1) can be written as

$$G = A \begin{pmatrix} I_k & 0_{k,n-k} \end{pmatrix} B$$

where $A = X^{-1}$ and $B = Y^{-1}$. Thus if $B = \begin{pmatrix} B_U \\ B_L \end{pmatrix}$ where B_U is $k \times n$ and B_L is $(n-k) \times n$, it follows that $G = AB_U$. But the matrix $\begin{pmatrix} AB_U \\ B_L \end{pmatrix}$ is unimodular, since it is obtained from the unimodular matrix B via a sequence of elementary row operations on the first k rows. Conversely, if $B = \begin{pmatrix} G(D) \\ H(D) \end{pmatrix}$ is unimodular, then the equation $G(D) = I_k \begin{pmatrix} \Gamma_k & 0_{k,n-k} \end{pmatrix} B$ shows that the invariant factors of $G(D)$ are all 1. \square

Theorem A-2. A $k \times n$ polynomial is reduced (see Definition 2) if and only if one of the following three conditions is satisfied:

(1) If the “matrix of high-order coefficients” \overline{G} is defined by

$$\overline{G}_{ij} = \text{coeff}_{D^{e_i}} g_{ij}(D)$$

where e_i is the degree of $G(D)$'s i th row, then \overline{G} has rank k .

(2) $\text{ext.deg. } G(D) = \text{int.deg. } G(D)$.

(3) The “predictable degree property”: For any k -dimensional polynomial vector, i.e., any $x(D) \in F[D]^k$

$$\text{deg}(x(D)G(D)) = \max_{1 \leq i \leq k} (\text{deg } x_i(D) + \text{deg } g_i(D))$$

Proof: The logical organization of this proof is as follows: It shall be proved that (Reduced) \rightarrow (1) \rightarrow (2) \rightarrow (Reduced), and (1) \leftrightarrow (3).

• (Reduced) \rightarrow (1): Suppose property (1) is false. Then there is a nonzero k -dimensional vector from F , say $\alpha =$

$(\alpha_1, \dots, \alpha_k)$, such that $\alpha \overline{G} = 0$. Now suppose that the rows of G are (g_1, \dots, g_k) with $\text{deg } g_i = e_i$, and $e_1 \leq e_2 \leq \dots \leq e_k$. Then from $\alpha \overline{G} = 0$, it follows that the coefficient of D^{e_k} in the linear combination

$$g'_k = \alpha_1 D^{e_k - e_1} g_1 + \alpha_2 D^{e_k - e_2} g_2 + \dots + \alpha_k D^{e_k - e_k} g_k$$

is zero, so that the unimodular transformation of $G(D)$ that replaces g_k with g'_k —and leaves the remaining rows of G unchanged—reduces the external degree of G . In other words, if property (1) is false, G is not reduced, which is (the contrapositive of) what the authors have set out to prove.

• (1) \rightarrow (2): Suppose that \overline{G} has rank k , and denote the $k \times k$ submatrices of \overline{G} by \overline{G}_ν , for $\nu = 1, 2, \dots, \binom{n}{k}$. Then since $\text{rank } \overline{G} = k$, there is at least one index ν_0 such that $\det \overline{G}_{\nu_0} \neq 0$. If now the row degrees of G are e_1, \dots, e_k , then (cf. the proof of Theorem 1(2)) the coefficient of $D^{e_1 + \dots + e_k}$ in $\det G_{\nu_0}$ is $\det \overline{G}_{\nu_0} \neq 0$. Thus $\text{int.deg. } G \geq \text{ext.deg. } G$. The opposite inequality is true for any matrix, as was shown in Theorem 1(2).

• (2) \rightarrow (Reduced): Suppose $\text{int.deg. } G(D) = \text{ext.deg. } G(D)$, and that $T(D)$ is an arbitrary $k \times k$ unimodular matrix. Then $\text{ext.deg. } TG \geq \text{int.deg. } TG$ by Theorem 1(2); $\text{int.deg. } TG = \text{int.deg. } G$, by Theorem 1(2); and $\text{int.deg. } G = \text{ext.deg. } T$, by assumption. Combining this string of inequalities and equalities, $\text{ext.deg. } TG \geq \text{ext.deg. } G$, which proves that G is reduced.

• (1) \leftrightarrow (3): Let $x(D) = (x_1(D), \dots, x_k(D))$ be a k -vector of polynomials, and let $y(D) = (y_1(D), \dots, y_n(D))$ be defined by the equation $y(D) = x(D)G(D)$. If the k rows of $G(D)$ are denoted by $g_1(D), \dots, g_k(D)$, then

$$\begin{aligned} y(D) &= x(D)G(D) \\ &= x_1(D)g_1(D) + \dots + x_k(D)g_k(D) \end{aligned} \tag{A-2}$$

If the degree of $x_i(D)$ is d_i for $i = 1, \dots, k$, and the degree of $g_i(D)$ is e_i for $i = 1, \dots, k$, it follows from Eq. (A-2) that the degree of $y(D)$ is at most $d = \max_i (d_i + e_i)$. Call d the “prediction” of the degree of $y(D)$. To test the prediction, note that the vector of coefficients of D^d in $y(D)$ is $\alpha = (\alpha_1, \dots, \alpha_k) \overline{G}$, where α_i is the coefficient of D^{d-e_i} in $x_i(D)$. (At least one of the α_i is nonzero, since $d_i + e_i = d$ must hold for at least one index i .) But $\alpha \overline{G} \neq 0$ for all nonzero α 's if and only if \overline{G} has rank k , and so the prediction is true for all $x(D)$'s if and only if $\text{rank } \overline{G} = k$. \square

Table 1. Generator-matrix properties.

Property	Basic?	Reduced?	Int.deg.	Ext.deg.
G_1	-	-	-	-
G_2	No	No	3	4
G_3	Yes	No	1	3
G_4	Yes	No	1	2
G_5	No	Yes	3	3
G_6	Yes	Yes	1	1
G_7	No	Yes	2	2
G_8	-	-	-	-

500732

510-32

167930

P. 7

Uncorrectable Sequences and Telecommand

L. Ekroot, R. McEliece,¹ S. Dolinar, and L. Swanson
Communications Systems Research Section

The purpose of a tail sequence for command link transmission units is to fail to decode, so that the command decoder will begin searching for the start of the next unit. A tail sequence used by several missions and recommended for this purpose by the Consultative Committee on Space Data Standards is analyzed. A single channel error can cause the sequence to decode. An alternative sequence requiring at least two channel errors before it can possibly decode is presented. (No sequence requiring more than two channel errors before it can possibly decode exists for this code.)

I. Introduction

When a *command link transmission unit* (CLTU) consisting of many *codeblocks* is received by a spacecraft, the command decoder verifies that each codeblock is a valid codeword and accepts it, or that it is a slightly corrupted codeword and corrects it, or that it is too far from a valid codeword and rejects it. Rejecting a codeblock causes the receiver to give up on the unit and begin searching for the start of the next unit. At the end of the CLTU, there is a *tail sequence* designed to be rejected as a codeword, sending the decoder into a "search mode." This article analyzes the performance of the tail sequence recommended by the Consultative Committee on Space Data Standards (CCSDS) and used by several missions. So instead of the

usual question about a code, i.e., how many errors can the code correct or detect, the question here is how many errors can occur before an uncorrectable sequence becomes correctable.

II. Analysis of Uncorrectable Sequences

In order for a sequence to be uncorrectable, it must be far enough from a codeword to cause the decoder to not decode. At the very least, it must differ from the nearest codeword in more positions than the decoder is able to correct. However, channel errors can make such a sequence decodable. The more errors that must occur before the sequence becomes correctable, the less likely it is that the sequence will accidentally decode. In order to maximize the number of channel errors before the sequence will decode, it is necessary to characterize and find sequences that are as far away from codewords as possible.

¹ Consultant, California Institute of Technology, Engineering Department.

sequence has odd weight and differs from the nearest codeword in three positions; if the syndrome is $\begin{pmatrix} \alpha^j \\ 1 \end{pmatrix}$, the sequence has odd weight and differs from the nearest codeword in the j th position; if the syndrome is $\begin{pmatrix} \alpha^j \\ 0 \end{pmatrix}$, the weight is even and the sequence differs from the nearest codewords in two positions.

For this example sequence, the top part of the syndrome (modulo $\alpha^6 + \alpha + 1$) is

$$s_1(\mathbf{r}) \equiv \sum_{i=0}^{N-1} r_i \alpha^i \pmod{\alpha^6 + \alpha + 1} \quad (4)$$

$$\equiv 1 + \alpha^2 + \alpha^4 + \alpha^6 + \alpha^7 + \alpha^9 + \alpha^{11} + \dots + \alpha^{61} \quad (5)$$

$$\equiv 1 + \alpha^2 + \alpha^4 + (\alpha + 1) + \frac{\alpha^7 + \alpha^{63}}{1 + \alpha^2} \quad (6)$$

$$\equiv \alpha + \alpha^2 + \alpha^4 + \frac{\alpha\alpha^6 + 1}{1 + \alpha^2} \quad (7)$$

$$\equiv \frac{(\alpha + \alpha^2 + \alpha^4)(1 + \alpha^2)}{1 + \alpha^2} + \frac{\alpha(\alpha + 1) + 1}{1 + \alpha^2} \quad (8)$$

$$\equiv \frac{\alpha + \alpha^2 + \alpha^4 + \alpha^3 + \alpha^4 + \alpha^6 + \alpha^2 + \alpha + 1}{1 + \alpha^2} \quad (9)$$

$$\equiv \frac{\alpha^3 + \alpha^6 + 1}{1 + \alpha^2} \quad (10)$$

$$\equiv \frac{\alpha^3 + (1 + \alpha) + 1}{1 + \alpha^2} \quad (11)$$

$$\equiv \frac{\alpha^3 + \alpha}{1 + \alpha^2} \quad (12)$$

$$\equiv \alpha \quad (13)$$

where Eqs. (6), (8) and (11) follow from equivalence modulo $\alpha^6 + \alpha + 1$. The bottom part of the syndrome (modulo $\alpha + 1$) is

$$s_2(\mathbf{r}) \equiv \sum_{i=0}^{N-1} r_i \alpha^i \pmod{\alpha + 1} \quad (14)$$

$$\equiv \text{weight}(\mathbf{r}) \pmod{2} \quad (15)$$

$$\equiv 32 \pmod{2} \quad (16)$$

$$\equiv 0 \quad (17)$$

The syndrome $\binom{\alpha}{0}$ indicates that the sequence is not a codeword in the (63,56) code, and is two away from the nearest codewords. A single error in any bit except r_1 will make the sequence differ from a codeword by one, and thus decodable.¹

If instead a sequence is considered that is an odd-weight word in the perfect (63,57) code, it is three away from the codewords of the (63,56) code. This means that two errors must occur before it becomes decodable by a single-error-correcting (63,56) decoder. Such a sequence would be a better choice for a tail sequence because it is more resistant to accidental decoding in the presence of errors.

In selecting a particular sequence for the command coding application, the effects on the distance and uncorrectability properties as the code is shortened must be taken into account.

C. Shortening the (63,56) Code

Select a subset of a code, where all the codewords in the subset have zeros in some specified positions. Since all of the codewords in the subset have zeros in the specified positions, those positions carry no information and can be ignored. The resulting set of codewords forms a *shortened code*. Shortening cannot decrease the minimum distance, and will only increase the minimum distance if the code is shortened severely.²

For this application, shortening will be done by taking only the codewords which have zeros in the leftmost or first m positions. Note that for each shortened word there is a corresponding full-length word that has zeros in the first m positions.

Example 2: Consider the sequence of length 55

01 1010101

It corresponds to the full-length sequence

0000000001 1010101

with weight 28.

The top part of the syndrome for the full-length sequence is

$$s_1(\mathbf{r}) \equiv \sum_{i=0}^{N-1} r_i \alpha^i \pmod{\alpha^6 + \alpha + 1} \quad (18)$$

$$= 1 + \alpha^2 + \alpha^4 + \alpha^6 + \alpha^7 + \alpha^9 + \alpha^{11} + \dots + \alpha^{53} \quad (19)$$

$$\equiv 1 + \alpha^2 + \alpha^4 + (1 + \alpha) + (\alpha^7 + \alpha^9 + \alpha^{11} + \dots + \alpha^{53}) \quad (20)$$

$$= \alpha + \alpha^2 + \alpha^4 + \frac{\alpha^7 + \alpha^{55}}{1 + \alpha^2} \quad (21)$$

$$\equiv \alpha + \alpha^2 + \alpha^4 + \frac{\alpha(\alpha + 1) + \alpha(\alpha + 1)^9}{(1 + \alpha)^2} \quad (22)$$

¹ If the error is in r_1 , the syndrome becomes $\binom{0}{1}$, indicating that it is three away from a codeword.

² For the (63,56) code, it can be shown that as long as the shortened length is greater than 32, the minimum distance will remain 4, and the maximum distance of any sequence to the nearest codewords will remain 3.

$$= \alpha + \alpha^2 + \alpha^4 + \frac{\alpha(1 + (\alpha + 1)^8)}{1 + \alpha} \quad (23)$$

$$= \alpha + \alpha^2 + \alpha^4 + \frac{\alpha(1 + \alpha^8 + 1)}{1 + \alpha} \quad (24)$$

$$\equiv \alpha + \alpha^2 + \alpha^4 + \frac{\alpha\alpha^2(\alpha + 1)}{1 + \alpha} \quad (25)$$

$$= \alpha + \alpha^2 + \alpha^4 + \alpha^3 \quad (26)$$

$$= \alpha(\alpha + 1)^3 \quad (27)$$

$$\equiv \alpha^{19} \quad (28)$$

The bottom part of the syndrome is zero since the weight is even. The two-part syndrome $\begin{pmatrix} \alpha^{19} \\ 0 \end{pmatrix}$ indicates that the sequence is not a codeword, and that it is two away from the nearest codewords. A single error can make the sequence differ from a codeword by one, and thus decodable.

D. Finding a Good Uncorrectable Sequence

The concepts in Section II.B and Section II.C lead to the definition of a *good uncorrectable sequence* as one for which it and all the desired truncations of it are maximally distant from the codewords in the corresponding code, i.e., an odd-weight codeword in the perfect (63,57) code. The syndrome for good uncorrectable sequences is $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The analysis below shows that an uncorrectable sequence can be chosen so that, when it is truncated by octets, its syndrome does not change, and therefore it does not become correctable.

If a given sequence is truncated by m bits, and if it is desired that the syndrome not be changed by the truncation, then the smallest possible m is eight. This is because in order to not change either part of the syndrome, the truncated bits must correspond to a polynomial that is zero modulo both $\alpha + 1$ and $\alpha^6 + \alpha + 1$; the lowest order nonzero polynomial satisfying that requirement is $g(\alpha) = \alpha^7 + \alpha^6 + \alpha^2 + 1$. Therefore the shortest such nonzero sequence is 11000101, which has length eight.

There are engineering reasons for truncating by octets in the application considered in Section III. Also, bit synchronization requirements often make it preferable to have many transitions in the sequence, and thus a mostly zeros sequence is undesirable. For the remainder of the article, it is assumed that shortening will be done only by multiples of eight bits, and that octets of zeros are not of interest.

Since the bottom part of the syndrome must be 1 for the sequence and all of its truncations, and the truncated bits 11000101 have even weight, the nontruncated part of the sequence must have odd weight. Since the top part of the syndrome must be 0 for the sequence and all of its truncations, the nontruncated part of the sequence must correspond to a polynomial which is zero modulo $\alpha^6 + \alpha + 1$. The shortest such sequence is 1000011.

A simple construction of a good uncorrectable sequence is a concatenation of octets of the form 11000101 with the seven bits 1000011 in the rightmost positions. This is not the only good uncorrectable sequence, but it does have good distance and bit synchronization properties. The syndrome for this sequence is confirmed in the next example.

Example 3: Consider

11000101 11000101 11000101 11000101 11000101 11000101 11000101 1000011

This sequence has weight 31, so the bottom part of the syndrome $s_2(\mathbf{r})$ is 1. The top part of the syndrome (modulo $\alpha^6 + \alpha + 1$) for the full-length sequence is

$$s_1(\mathbf{r}) \equiv \sum_{i=0}^{N-1} r_i \alpha^i \pmod{\alpha^6 + \alpha + 1} \quad (29)$$

$$= 1 + \alpha + \alpha^6 + \sum_{k=0}^6 \alpha^{7+8k} (1 + \alpha^2 + \alpha^6 + \alpha^7) \quad (30)$$

$$= 1 + \alpha + \alpha^6 + \sum_{k=0}^6 \alpha^{7+8k} (1 + \alpha + \alpha^6)(1 + \alpha) \quad (31)$$

$$\equiv 0 \quad (32)$$

Thus, the sequence is three away from the nearest codewords. A single error cannot make this sequence decodable.

III. Command Link Coding

The CCSDS recommendation uses a tail sequence that is specially constructed to not decode. By not decoding, it causes the receiver to begin searching for the next CLTU. It will be shown that the sequence that has been recommended is not the best sequence in terms of distance, and a better one will be given.

In order to apply the results of Section II to the command coding problem, the operations of the CLTU must be detailed. Refer to [1] for more detailed explanations of what is summarized here.

A. Telecommand Codeblock

The telecommand codeblock has K information bits, 7 *inverted* parity check bits, and a fill bit for a total length of L . Because of the code selected and the desire to shorten in units of 8 bits, the number of information bits and block lengths considered are $K = 32, 40, 48, \text{ and } 56$, and $L = 40, 48, 56, \text{ and } 64$, respectively.

B. Command Link Transmission Unit

The CLTU consists of

- (1) a 16-bit start sequence, namely 1110101110010000,
- (2) a number of telecommand codeblocks (the information may be padded with fill to make the information into a multiple of K bits), and
- (3) a tail sequence which is a sequence of bits the same length as a codeblock and designed to be uncorrectable. The idea is to cause the receiver to stop decoding and begin looking for the start sequence of the next CLTU.

C. Tail Sequences

The CCSDS recommendations specify that the tail sequence \mathbf{t} be a sequence of alternating zeros and ones beginning with a zero. Ignoring the fill bit, and noting that the parity bits are inverted, it can be seen that for a block length L of 64, this corresponds to the sequence in Example 1. As illustrated in the example, this sequence has distance properties such that a single channel error can, and almost certainly will, make the sequence decodable.

For the shorter block lengths $L = 40, 48,$ and $56,$ the sequences of alternating zeros and ones beginning with a zero have corresponding full-length sequences with zeros filled in the first positions. All three of these corresponding sequences have even weights, and have nonzero syndromes. The calculation for $L = 56$ is done in Example 2.

If instead the sequence

11000101 11000101 11000101 11000101 11000101 11000101 11000101 0111100 0

is used as the tail sequence, it corresponds (when the fill bit is removed and the parity bits are inverted) to the sequence in Example 3. If one channel error occurs, then this sequence will still be uncorrectable.

D. Augmentation Using the Fill Bit

The CCSDS recommendations propose to use the fill bit as a flag to tell the decoder to operate in error-detect mode only. This augmentation is only suggested for use with the tail sequence. This improves the probability of spotting the tail sequence by not allowing the decoder to correct any errors when the fill bit is 1. In this mode, two errors are sufficient to make the CCSDS tail sequence (of alternating zeros and ones) decodable, while a minimum of three errors is required to make the sequence presented here decodable.

IV. Conclusions

The analysis in this article shows that there are uncorrectable sequences that can tolerate one more channel error than the CCSDS tail sequence before becoming decodable. It is also shown that this property may be preserved for the shortened as well as the full-length codes recommended by the CCSDS. A sequence satisfying these requirements should be considered for the role of the CCSDS tail sequence since it is more resistant to channel errors than the proposed tail sequence, and still has many transitions to aid bit synchronization.

Tables of probabilities of missing the tail sequence, the operation of the decoder on the shortened codes, and the proof that no sequence requiring at least three channel errors before it can possibly decode exists for this code may be the subject of future work.

Reference

- [1] NASA, Consultative Committee on Space Data Standards, *Telecommand Part 1 Channel Service*, Blue Book, 201.0-B-1, Washington, D.C., January 1987.

511-91
167931

p-12

500744

Galileo Post-Gaspra Cruise and Earth-2 Encounter

P. E. Beyer and M. M. Andrews
TDA Mission Support and DSN Operations

This article documents DSN support for the Galileo cruise after the October 1991 encounter with the asteroid Gaspra. This article also details the Earth-2 encounter and the special non-DSN support provided during the Earth-2 closest approach.

I. Introduction

The launch, initial acquisition, Venus encounter, Earth-1 encounter, asteroid Gaspra encounter, and associated cruise periods have been documented [1]. This article documents the cruise period following the Gaspra encounter and leading up to the Earth-2 encounter, as well as the Earth-2 activities.

Also documented are the first solar conjunction experienced by the JPL Galileo Project in January 1992 and the many High-Gain Antenna (HGA) anomaly recovery windows that were identified. These windows were used to perform minisequences designed to attempt to free the stuck ribs [1]. These activities included cooling turns, warming turns, and Dual-Drive Actuator (DDA) motor calibrations. DSN support requirements are noted.

Finally, due to an unexplained increase in velocity noticed during closest approach to Earth in the first flyby, Galileo requested special non-DSN coverage during the Earth-2 DSN tracking coverage gap at the closest approach. The DSN negotiated support with the National Space Development Agency of Japan (NASDA), the Euro-

pean Space Agency (ESA), the NASA/Tracking and Data Relay Satellite System (TDRSS), and the University of Chile. This special support is also documented.

II. Solar Conjunction 1992

On January 22, 1992, the spacecraft passed behind the Sun, with a minimum Sun-Earth-craft (SEC) angle of approximately 2 deg. As part of the HGA anomaly recovery effort, a minisequence for the Cooling Turn No. 4 needed to be uplinked prior to the conjunction. The sequence was scheduled to go active on January 28, 1992, and last through February 8, 1992.

The upload was originally planned for January 13, 1992, from Deep Space Station (DSS) 14, as the Project did not plan to command inside of a 5-deg SEC angle. The SEC angle was greater than 5 deg on that day. Unfortunately, problems were encountered, which caused several commands to not be accepted by the spacecraft. The DSN verified configurations to ensure that this anomaly was not due to the ground equipment. DSS 43 was scheduled in real time to resend the commands, and again not all com-

mands were correctly received and processed by the spacecraft. Solar effects had begun disrupting the data downlink from approximately an 8-deg SEC angle. The incidence of errors detected in the telemetry by the Project increased coincident with the smaller SEC angles.

Also, radio science data collected during this time confirmed highly variable solar activity. Several plasma ejection events were detected by the radio scintillation experiment. That experiment is a conjunction support requirement for the DSN for all conjunctions. Due to the HGA anomaly, the first conjunction was supported on the LGA. The effect was that the Project had to reschedule some DSN support, to relinquish other support, and also on occasion to reconfigure the spacecraft for 34-m standard antenna passes. The experiment was successfully supported by the DSN.

After the uplink problems on January 13, 1992, more commanding was attempted on January 14, 1992, from DSS 14 again. Both the Project and the DSN were surprised by the inability to command at greater than a 5-deg SEC angle, since other projects had not experienced such problems. Investigation of the ground equipment revealed no ground problems that could have been contributors. The low SEC angle and associated solar interference appeared to be the root cause of the problem. Due to the spacecraft range and the HGA anomaly, the uplink power of 100 kW was the norm for commanding to the LGA. All attempts through the DSS 14 pass on January 14, 1992, utilized 100 kW. Though all commands were sent twice, no commands were processed by the spacecraft.

At this time the transmitter uplink power of 400 kW was requested and authorized. Commands were again radiated, with two attempts planned for each. However, the high-power transmitter began to periodically trip off with crowbar fire alarms. Finally, the transmitter was declared red, and the pass was completed downlink only. Thus, all planned commands could not be sent.

One last attempt to transmit the cooling turn sequence before conjunction was planned over DSS 63 on January 15, 1992. If it proved to be unsuccessful, the sequence would not be executed, and this attempt in the series of activities to free the stuck ribs on the HGA would be lost.

Again, the 400-kW transmitter uplink power was requested and utilized. Early in the pass, the transmitter tripped off once for a short duration. Then it was solidly operational for several hours, allowing multiple sets of commands to be transmitted. Surprisingly, the commands sent later in the pass were the ones processed by the spacecraft. Then the automatic gain controls (AGC's), both on

the ground and on the spacecraft, appeared to stabilize. Incidentally, solar activity had begun to subside. At last the commanding was successful and the cooling turn was executed as planned.

III. DSN Support for High-Gain Antenna Activities

The HGA was not successfully deployed as planned in April 1991 [1]. An anomaly recovery team was quickly formed to assess the problem and potential solutions. Since the Probe mission was still intact, and the spacecraft was otherwise perfectly healthy, only minimum-risk ideas were considered. The most plausible failure scenario was that the mid-rib guide pins were stuck in their receptacles. The best engineering judgment led to a walking pin theory as a possible solution to this HGA anomaly.

This theory was based on a model which showed that by alternately warming and then cooling the HGA tower, the tower would expand and contract and would allow small movement of the guide pins until they were eventually walked out. Cooling turns (CTs) and warming turns (WTs) were carried out throughout 1991 and 1992 (Fig. 1). These turn events involved turning the spacecraft either about 45 deg off-Sun to expose and thus heat the HGA tower, or about 165 deg to shield the tower and thus contract it.

The DSN support for the various cooling and warming turns was unique in many respects. For instance, during Cooling Turn No. 2, the Low-Gain Antenna 2 (LGA-2) was switched on because of the large aspect angle of the LGA-1. Even with the LGA-2, the lowest Galileo data rate of 10 bps was not always supportable on the telecommunications link.

Furthermore, nonstandard temperature control was first attempted to maximize cooling. This involved turning some heaters off, etc.; but of more importance was that the Ultra Stable Oscillator (USO) was turned off. This meant that the DSN had to generate separate predictions for the Auxiliary Oscillator. The DSN predictions unit also had an unknown in the USO predictions after the USO was powered back on. This unknown consisted of the effects of temperature changes on USO frequency, which had been minimally characterized in flight.

Temperature changes during the turns resulted in significant drift of the Voltage Controlled Oscillator (VCO). For example, the VCO drifted about 4.5 kHz in the week of precooling before the actual Cooling Turn No. 2. Two-way

operation with the VCO is the prime support configuration for the DSN.

The DSN tracking predictions unit was able to adapt well to these frequency changes. They maintained close interaction with the Galileo spacecraft team to ensure good temperature monitoring. The USO was again turned off during Cooling Turns Nos. 3-6.

One other way that the DSN support was affected during cooling turns was that the Travelling Wave Tube (TWT) was powered to its low power state, again to minimize temperature. This decreased the telecommunications link by 4.8 dB and further complicated the data acquisition process. As the spacecraft moved toward aphelion, the telemetry link margin was negative for the cooling turns. The DSN contributed to the Project's strategy for monitoring the spacecraft state. The strategy developed was that the Project would schedule DSN tracking support for about 2 hr every 8-10 hr to monitor for presence of carrier. If the spacecraft were to safe itself, it would orient itself towards the Sun, default to 10-bps telemetry, and change subcarriers. The received AGCs would indicate the respective changes.

The DSN supported these alternate cooling and warming turns through July 1992. As time went on, the walking pin theory was examined repeatedly, and the Project determined that the expectation of the stuck pins walking out was unlikely. The model showed that there was little to be gained after about six turns, as the curves became quite steep; i.e., if the model was correct, the pins should have been freed by completion of a half-dozen maneuvers.

One other HGA activity was supported by the DSN in July 1992. In conjunction with the DDA motors' activity of July 21, 1992, the LGA-2 was retracted. In spacecraft testing prelaunch, some recalled that an LGA-2 retraction test shook the spacecraft when the LGA hit the stops on the Radioisotope Thermoelectric Generator (RTG) boom on which this antenna is suspended. While the retraction was successful, no change in the HGA configuration was observed.

Alternate theories and possibilities had been put forth and examined. The one that seemed to hold the most promise was Dual-Drive Actuator (DDA) hammering. The DDA includes the motors that are used to drive the HGA open. This theory, extensively modelled with tests on the spare flight HGA in the laboratory at JPL, showed that short on/off pulses of the HGA motors generated force, and could drive the ball-screw up its shaft with a total of about 1.5 revolutions maximum before stalling out again.

The force generated on the push rods would be increased about threefold, thereby generating the hope that the antenna would be opened.

Due to the highly critical Earth-gravity assist (EGA), the Project did not want to perform hammering prior to that flyby. However several motors on calibrations were conducted in the April-October 1992 time frame. The first few were just 2-sec turn-ons at various temperatures to measure the DDA performance and to calibrate it against the laboratory measurements. Then in October 1992, a hammer test was performed. This was a shortened demonstration of the extensive hammering planned for late-December 1992 and early 1993.

The DSN supported extensive DDA hammering activities beginning on December 29, 1992. These hammering exercises continued for about 3 weeks. Though some ball-screw movement was detected, the HGA ribs were not freed.

Galileo had an agreement with NASA to declare that the Galileo mission would be supported on the LGA on March 1, 1993, should the HGA still be only partially deployed. Since the aforementioned activities apparently were not successful, the DSN will in fact support the mission at Jupiter only on the LGA. Significant new TDA developments, as well as spacecraft enhancements, are planned to increase the flexibility of, and maximize the supportable, data rates at the Jupiter ranges. These developments include, among other things, DSN antenna arraying, new coding techniques, data compression, and full carrier suppression operations with the Block-V receiver. These developments are planned to be discussed in detail in future TDA Progress Reports.

IV. DSN Support of Earth-2 Flyby

The Earth-2 flyby was essential to the Jupiter mission. If it were not executed properly, the spacecraft would not reach Jupiter. Navigation was near perfect, however, and Galileo is now finally on its way to completion of its prime mission.

DSN support for the flyby was substantial. There were a multitude of requirements to fulfill. These included, but were not limited to, Trajectory Correction Maneuvers (TCMs 14, 15, 16 and 17), the complete Gaspra tape-recorder playback [1], two very critical Probe tests, science instrument calibrations, and the closest-approach science.

TCM-14 was performed on August 4, 1992, and imparted a delta velocity of 21.27 m/sec. This was a deterministic maneuver which would put the spacecraft on

course for its second asteroid encounter. That encounter will be with Ida on August 28, 1993.

TCM-15 was performed on October 9, 1992, and imparted a delta velocity of 0.72 m/sec. TCM-16 was performed on November 13, 1993, and imparted a delta velocity of 0.89 m/sec. TCM-17 was the encounter-minus-10-day maneuver. The delta velocity was only 0.03 m/sec.

The sum effect of the targeting was that closest approach was nearly exactly on target and on time. The altitude was only 304 km. The time of closest approach was 15:09:25 UTC (Fig. 2). The targeting was so effective that the TCM-18 planned for December 21, 1992, as a cleanup maneuver was declared unnecessary and therefore canceled.

Complete Gaspra data were played back from the tape recorder on November 23-24, 1992 (Fig. 3). The data were played back twice, and over two Deep Space Complexes to ensure ground capture. They were completely recovered by the DSN on the first attempt.

The highest resolution color image of Gaspra had previously been played back shortly after the asteroid encounter [1]. And when the data rate again supported 40 bps, the highest resolution image (black and white) was relayed to the ground via a Data Management Subsystem Memory Readout (DMSMRO). This occurred over the period between May 18 and June 6, 1992.

Two highly critical Probe tests were supported flawlessly by the DSN in the pre-Earth-encounter period. The first in-flight Mission Sequence Test (MST) was conducted on November 20, 1992. The MST was critical because it will be the only complete MST in the mission if the HGA is not deployed. All Probe data are currently designed for 28.8 kbps data rate. Furthermore, the test is critical in that the Probe ground data processors are fairly intolerant to data gaps.

Some subtle communication problems had been noticed from the Madrid Deep Space Communications Complex (Spain) (MDSCC) in November 1992, which caused the Project some anxiety. Knowing this, DSN and MDSCC personnel spent great energies troubleshooting the communications system, resolving anomalies, and thus ensuring the integrity of the Probe test. A letter of commendation was received by the JPL Assistant Laboratory Director for TDA from the Galileo Project Manager as a result of the DSN support for this test.

Then on December 2, 1992, the Probe Abbreviated System Functional Test (ASFT) was conducted. Again, it was

supported flawlessly by the DSN. All test objectives were met for both Probe tests.

The Gaspra playback, Probe tests mentioned above, and the Earth's and Moon's closest-approach activities were conducted with heightened DSN personnel awareness, special support, and NASA Communications Network (NASCOM) special coverage. This support was requested by the Project and was provided successfully by the DSN.

The short range to the spacecraft during the Earth-2 flyby period allowed sufficient telemetry signal-to-noise ratio for all of the Galileo data rates to be exercised. It also allowed the DSN to verify its performance at the higher rates. Some changes since launch and Earth-1 were noted. For example, the type-B Telemetry Processor Assembly (TPA) buffer capacity was reduced during the DSN telemetry system upgrade. The effect of this reduction was that simultaneous recording and transmission of data rates above 80.64 kbps were precluded.

Of more importance than just exercising the high data rate was the ability of the Project to perform science instrument calibrations. If the HGA is not freed, then detailed, high-rate calibrations will not be possible for the duration of the mission. Many science instrument calibrations occurred in the days before closest approach. Then the close Moon and Earth observations were completed to verify the calibrations in preparation for the mission at Jupiter. The DSN supported all these calibrations without incident.

More detail on the closest-approach activities is now noted (see Fig. 4). For this flyby, the Moon's closest approach was at 110,000 km, contrasted with 350,000 km for the Earth-1 flyby. Also, the latitude for the Earth-2 encounter was 61 deg north as opposed to equatorial for Earth-1. This increased the interest in the lunar observations due to this trajectory providing the North Polar Moon coverage. The DSN supported these observations flawlessly beginning about 5 hr before the Moon's closest approach. The closest approach was at 03:47:45 UTC. The lunar observations continued until only a few hours before the Earth's closest approach.

At the Earth's closest approach, Galileo began to concentrate on Solid State Imaging (SSI) of the Andes, Antarctica, Hawaii, Indonesia, and other places. The Andes data were placed on the tape recorder for later playback, as the DSN was not in view during this time. The next section discusses special, non-DSN coverage during the DSN tracking gap. This coverage was limited to carrier tracking only.

After the closest approach, the LGA-1 aspect angle to the DSN antennas was more than 100 deg. Poor telemetry performance was noted. It was most noticeable as black streaks in the real-time images that the Project was displaying on the JPL monitors. As noted earlier, the LGA-2 had been stowed in July 1992; and with the HGA anomaly, the only available antenna was the LGA-1. Also, the sequence on board the spacecraft assumed a margin to support the highest possible telemetry data rate, i.e., the data rate was lowered only when the higher data rate was not deemed to be supportable.

This condition lasted for several days. It was determined quite early that the high aspect angle was likely reflecting the signal through the magnetometer boom, thus causing the unpredicted poor performance. The Project had little choice but to accept this anomalous condition until the aspect angle became much better, as a modified sequence was unattainable. The DSN had to provide support through these periodic data gaps.

The Galileo Optical Experiment (GOPEX) was very successfully supported from December 9–16, 1992. The DSN provided special pointing predictions to the Table Mountain Observatory and the Starfire Optical Range. The DSN also provided special communications equipment and arranged for communication lines for these activities. More detail on the experiment is provided in [2].

V. Special Non-DSN Support During Earth-2 Closest Approach

A trajectory anomaly was observed during the Earth-1 encounter. It coincided with the closest approach, which was out of view of the DSN. After intense examination by the Galileo navigation and science teams, no plausible argument could explain the trajectory anomaly, a 3- to 4-mm/sec velocity increase, at the Earth-1 closest approach. The trajectory for Earth-1 flew almost directly over Madrid, Spain, with a closest approach at 25 deg north latitude and 63 deg west longitude. The minimum altitude was 960 km. Goldstone rise was then 15 min after Madrid set, thus leaving this flyby with a short DSN tracking coverage gap. The Earth-2 closest approach was at 34 deg south latitude, 6 deg west longitude, and only 304 km altitude. This led to a DSN tracking gap of about 2 hr.

A Galileo celestial mechanics investigator, John Anderson, convinced the Project that there was reasonable

probability that the answer may have been of fundamental science significance, i.e., gravitation or general relativity related. In April 1992, the Project requested that the TDA team pursue options available for closing the DSN tracking gap during the Earth-2 flyby.

Many ground sites were considered, along with NASA's own TDRSS. After some deliberation, it was decided to enter into negotiations with NASA and Goddard Space Flight Center for use of a TDRSS satellite; with the University of Chile for a Santiago tracking station; with the European Space Agency (ESA) for its Perth, Australia, tracking station; and with the National Space Development Agency of Japan (NASDA) for an Okinawa tracking station.

Ultimately, TDRS-3 at 62 deg west longitude was used for support. Also utilized were the 9-m antenna at Santiago, Chile, the 15-m antenna at Perth, Australia, and one of two 18-m antennas at Okinawa, Japan (Fig. 5).

The TDRSS operations were made easier by transporting a special receiver called the Experimental Tone Tracker (ETT) to White Sands, New Mexico, to communicate with the TDRS-3 (Fig. 6). This receiver was developed by the JPL Tracking Systems and Applications Section 335, for use with the Orbiting Very Long Baseline Interferometry (OVLBI) experiment, and so had previously been installed and operated at White Sands. The Galileo Doppler data were recorded from the ETT and transported to JPL.

The Santiago interface was the only real-time interface for the non-DSN supporters and was conveniently connected into NASCOM. The Perth data were sent electronically to Darmstadt, Germany, from where the ESA personnel forwarded them electronically to the JPL navigation team. The Okinawa data were formatted onto a floppy disk and mailed to JPL.

The time line shows that 3-way Doppler was recorded by the non-DSN stations until DSS 42 had the transmitter off at 1340 UTC. Then they reacquired 1-way until their respective set times. Thus there was a 22-min ground gap until Santiago had an acquisition of signal (AOS) at 1513 UTC. The link from Galileo to TDRS-3 to the ETT at White Sands maintained lock during this time. Goldstone acquired the signal before Santiago lost the signal. In fact, DSS 12 established an uplink at 1555 UTC and Santiago was in 3-way lock at 1556 UTC. The DSN tracking gap was thus completely covered by the non-DSN resources, and the operation was a complete success. Analysis of the data did not reveal an orbit anomaly.

References

- [1] P. E. Beyer, R. C. O'Connor, and D. J. Mudgway, "Galileo Early Cruise, Including Venus, First Earth, and Gaspra Encounters," *The Telecommunications and Data Acquisition Progress Report 42-109*, vol. January-March 1992, Jet Propulsion Laboratory, Pasadena, California, pp. 265-281, May 15, 1992.
- [2] K. E. Wilson, J. R. Lesh, T.-Y. Yan, J. Schwartz, M. D. Rayman, and S. Wee, "GOPEX: A Deep-Space Optical Communications Demonstration With the Galileo Spacecraft," *The Telecommunications and Data Acquisition Progress Report 42-109*, vol. July-September 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 262-277, November 15, 1990.

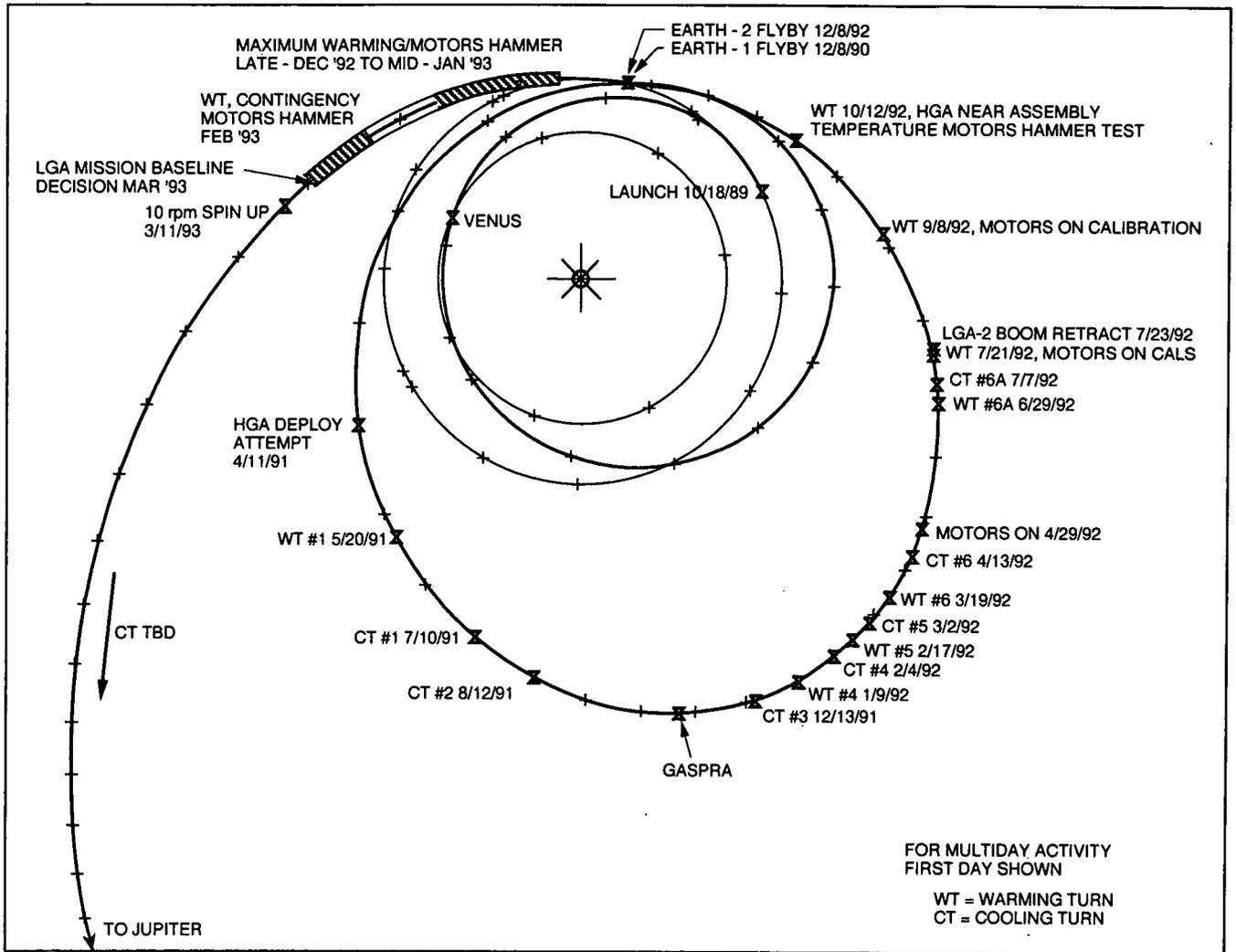


Fig. 1. Galileo HGA events.

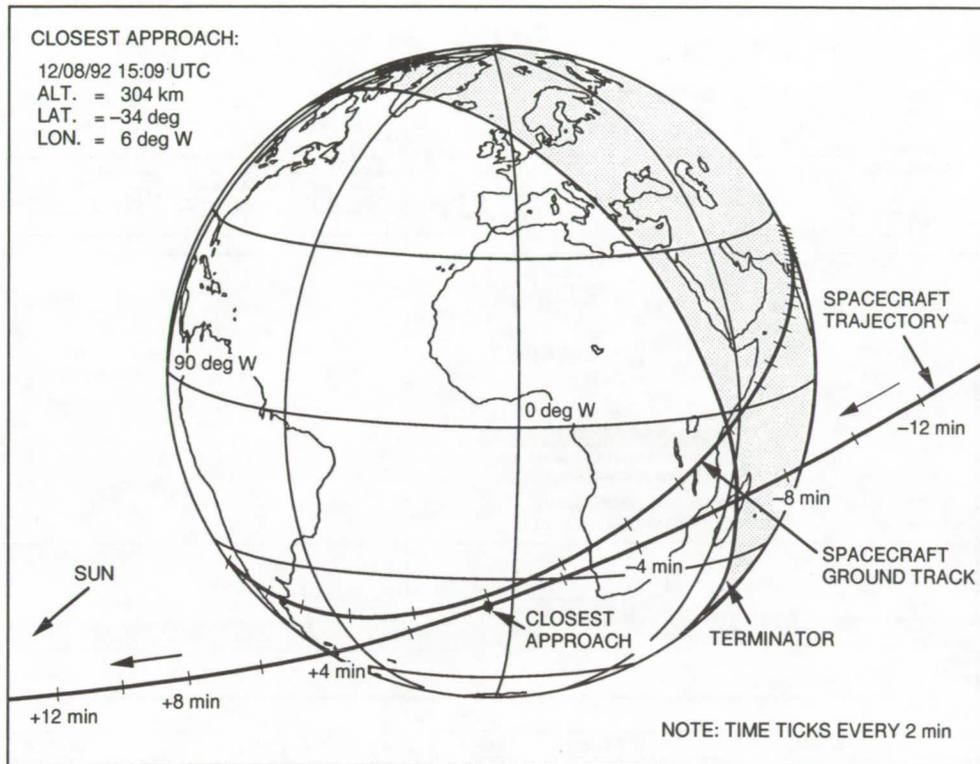


Fig. 2. Galileo ground track of Earth-2 flyby.

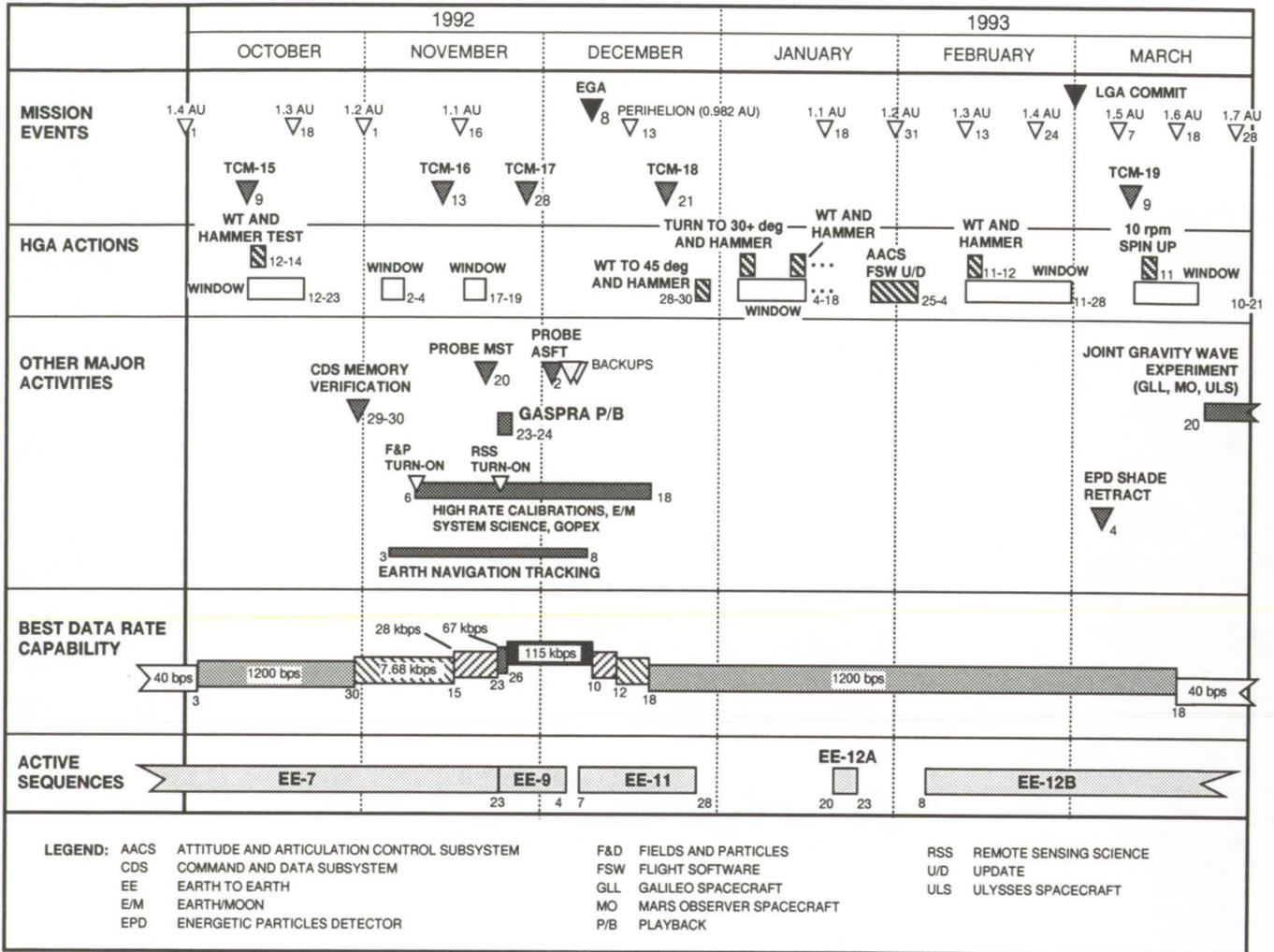


Fig. 3. Galileo time line of events.

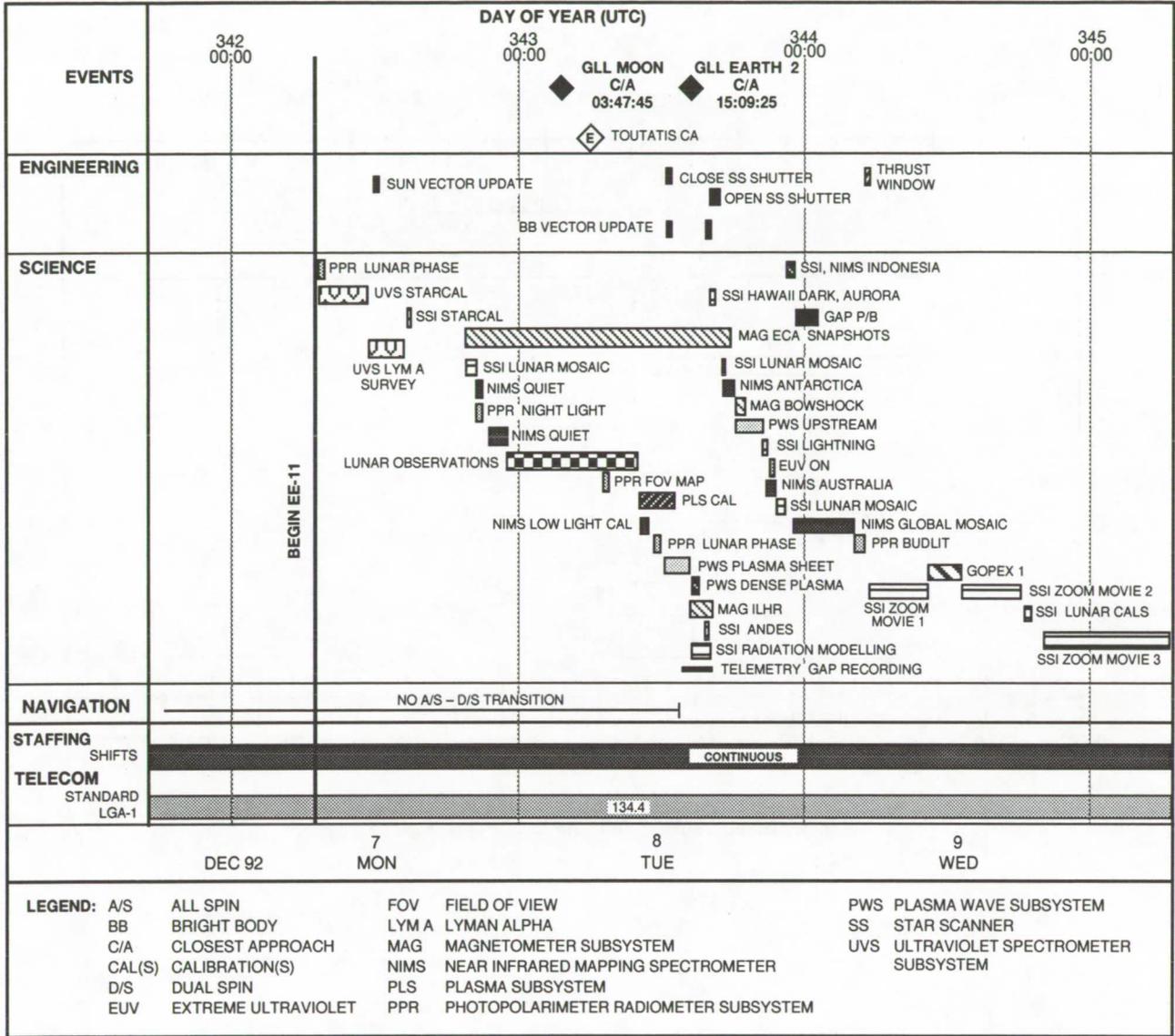


Fig. 4. Earth-2 encounter overview Earth closest-approach activities.

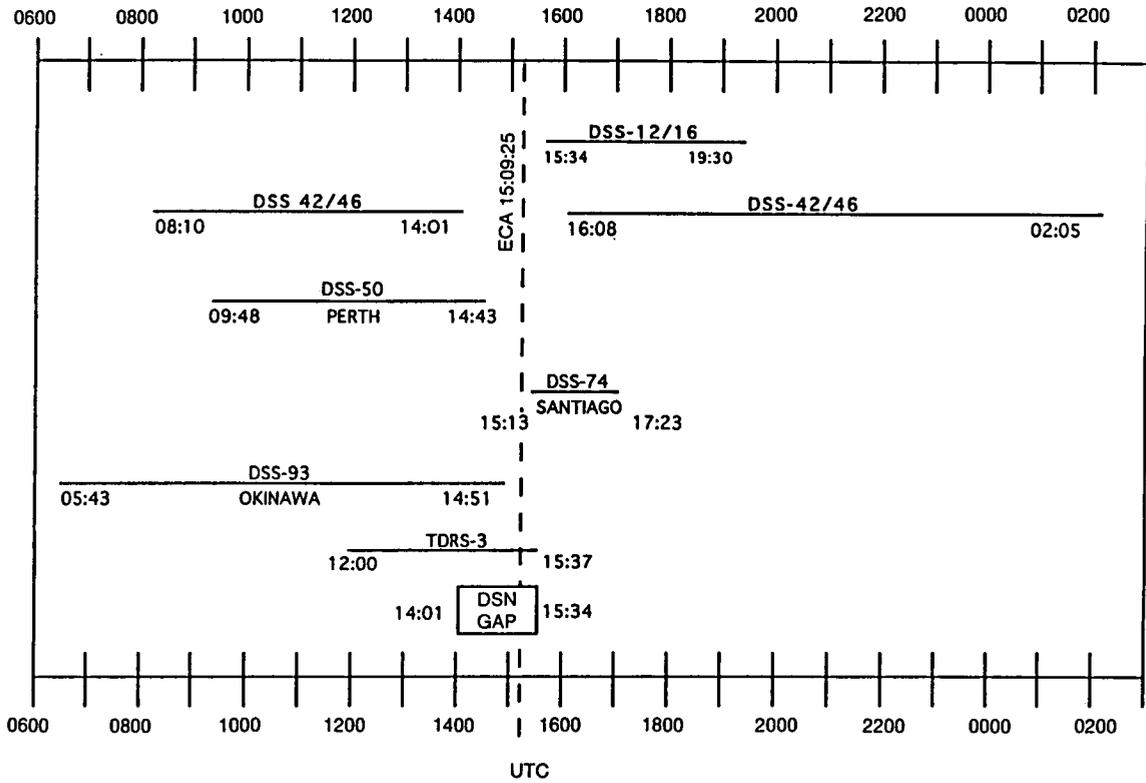


Fig. 5: Galileo tracking time line for Earth closest approach.

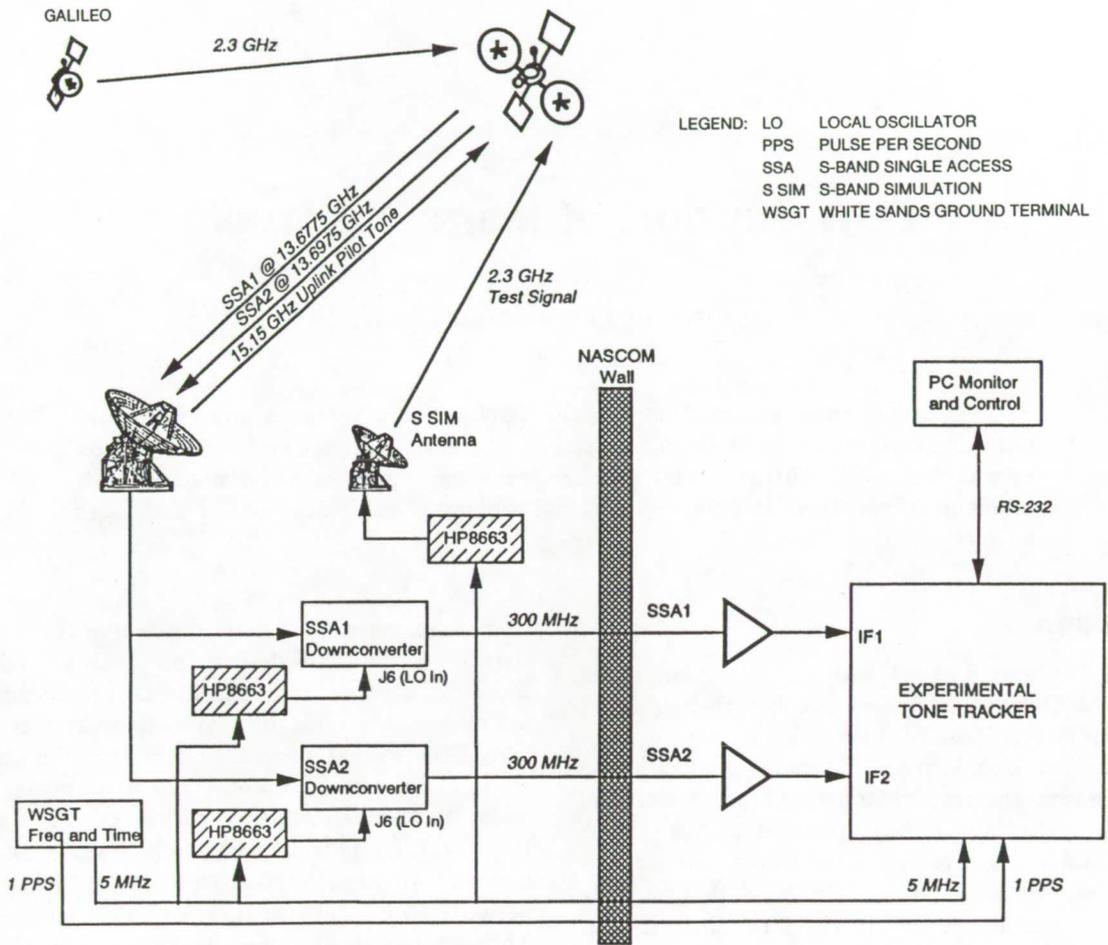


Fig. 6. Galileo Earth-2 TDRSS tracking.

S/2-81

167932

500751

P-5

DSN Support of Mars Observer

M. R. Traxler

TDA Mission Support and DSN Operations

This article provides a summary of the DSN prelaunch, launch and cruise support of Mars Observer through Trajectory Correction Maneuver 2 (TCM-2) on February 8, 1993. This summary includes planning, implementation, testing, DSN special configurations and DSN operational problems, and successes and challenges to date.

I. Introduction

The Mars Observer Project was originally planned as the first of the Observer series with a launch in August 1990. The primary purpose of the mission is to gather scientific information in orbit around Mars for one Mars year (687 days). Several factors dictated by budget constraints mandated that the launch be replanned to September 1992. This mission was directed to use Consultative Committee on Space Data Systems (CCSDS) data standards which played a significant role in the final Multimission Operations Systems Office (MOSO) and DSN configurations selected for this mission.

The spacecraft was built by GE Astro, East Windsor, New Jersey and required an overview function which had DSN participation. This spacecraft has an X-band 7164.624229-MHz uplink and an X-band downlink at two frequencies: 8417.716050 (coherent or auxiliary oscillator) or 8423.148147 (ultrastable oscillator). The spacecraft has a high-gain antenna, one low-gain transmit antenna and two low-gain receive antennas. All antennas are right-hand circular polarized.

II. Prelaunch Planning and Implementation

For several years the DSN worked in partnership with the Project to determine the specific DSN stations and

configurations which would best serve the mission objectives. The 34-m High Efficiency subnet, the newest of the DSN subnets, was selected to provide daily X-band uplink and downlink communications with the spacecraft. This DSN subnet provides telemetry at data rates from 10 bps through 85.3 kbps, commanding at data rates from 7.8 through 500 bps, radio metric data (two-way Doppler, ranging and VLBI), radio science, and DSN monitor data to the Project. The 70-m subnet also provides periodic VLBI and real-time high-rate telemetry support. These requirements were documented in the Mars Observer Support Instrumentation Requirements Document (SIRD) and a NASA Support Plan developed to meet these requirements.

The joint DSN/MOSO/Project plan for Mars Observer data developed into a real-time data flow of all DSN data types via the Advanced Multimission Operations System (AMMOS) to the Project Data Base. The exception is the radio metric data, which is routed via the DSN Multimission Navigation team for preprocessing before routing to the Project in near-real time.

The DSN implementation improvements for Mars Observer at launch were part of significant electronics upgrades (new computers, software and hardware) at the DSN complexes, Network Operations Control Center, and Ground Communications Facility. The DSN improve-

ments included several specific capabilities for Mars Observer. These improvements included (1) Standard Format Data Unit (SFDU) headers for all data types except command, (2) Reed-Solomon decoders at each complex, and (3) X-band acquisition aid at Canberra. The X-band acquisition aid in particular was required to provide a ± 2.4 deg beamwidth capability to ensure the first DSN initial acquisition at X-band. The 34-m high-efficiency (HEF) antenna has a pencil beam (± 0.0375 deg) at X-band. Therefore, the X-band acquisition aid was mounted on the 26-m antenna where angle information was transferred to the 34-m antenna to ensure timely spacecraft acquisition. All these upgrades were completed and thoroughly tested before launch.

III. DSN Compatibility Testing

The DSN conducted a comprehensive and intensive series of tests with the spacecraft telecommunications panel and associated equipment in April 1991. These two weeks of testing included telemetry, command, radio metric and radio science configurations required for the Mars Observer.

There were five anomalies found which are listed in Table 1. Anomalies 1-3 presented configurations which, while not desirable, were acceptable to the DSN. At the end of this testing, the spacecraft telecommunications system to the level it was tested was declared compatible with the DSN.

The DSN performed another 40-hour intensive set of compatibility tests with the final spacecraft at the Kennedy Space Center (KSC) launch site in July 1992. Four additional anomalies were identified and are listed in Table 2. The KSC testing was the first end-to-end command testing with the spacecraft.

IV. Prelaunch Testing

The DSN Network Operations Project Engineer (NOPE) conducted approximately 60 Mission Readiness Tests with the complexes. This large number of tests was partially required because of late changes to DSN and MOSO launch software.

A significant effort was also spent on planning, testing and documenting the DSN acquisition plan for Mars Observer. This was necessitated by the requirement that the DSN initially acquire the spacecraft at the X-band frequency. The launch vehicle and related interfaces were also new.

The launch vehicle was a Titan III/Transfer Orbit Stage (TOS). Interfaces were established with the launch vehicle to obtain prelaunch trajectory information, launch vehicle Mark Events and Titan and TOS state vectors. The DSN initial acquisition plan documented the nominal acquisition strategy and the planned use of the Mark Events, trajectory information and state vectors for nominal and non-nominal trajectories as well as spacecraft anomalies.

V. Launch and DSN Initial Acquisition Phase

The Mars Observer spacecraft was successfully launched on September 25, 1992, thirty-five minutes into the window at 17:05 UTC. The Titan Mark Events were passed over the net and indicated nominal performance. The Titan Operations Control Center at Cape Canaveral Air Force Station (CCAFS) provided the parking orbit state vector as planned, about 10 minutes after launch. The only significant problem in the launch phase was the loss of the TOS S-band (2272.5 MHz) transmitter, which was identified by the Advanced Range Instrumented Aircraft (ARIA) and confirmed by Dakar and Johannesburg.

The loss of the TOS transmitter caused all the TOS Mark Events not to be provided. The DSN was well informed of the probable reason for no TOS data and stayed with the DSN nominal acquisition plan. The TOS Project Operations Control Center (POCC) at KSC provided the DSN with a state vector based on the actual Titan parking orbit and planned TOS burn at launch plus 33 min. This vector was used by the DSN to generate updated predicts for the Canberra stations.

The DSN Canberra 26-m station had planned to track the TOS S-band (2272.5-MHz) downlink starting at station horizon break (approximately launch plus 49 min). Because of the TOS transmitter being off, there was no signal for Canberra to track.

The DSN 34-m HEF and 26-m stations acquired the X-band (8417.716-MHz) spacecraft downlink within 40 sec of the planned turn-on time of launch plus 84.7 min.

The Canberra tracking of the Mars Observer launch was excellent with all data provided to the Project. The quick acquisition of the spacecraft X-band signal by the Canberra 26- and 34-m antennas during initial contact with the spacecraft was much appreciated by the Project. This acquisition helped quickly relieve the tension from not having any TOS performance data and allowed the Project to quickly evaluate the spacecraft.

The DSN support of the launch countdown, launch, initial acquisition and first pass by Canberra was nearly flawless with no problems identified. The DSN met or exceeded all Project requirements, including the spacecraft 2-kbps data flow from a Payload Data Formatter (PDF) at Canberra to Building AO at Cape Canaveral Air Force Station.

VI. Early Cruise Through TCM-2

The 34-m HEF subnet provided continuous tracking from launch through 30 days on October 24, 1992. Trajectory Correction Maneuver 1 (TCM-1) and spacecraft checkout were accomplished during this period.

The DSN was unable to symbol synchronize and convolutionally decode the TOS telemetry data played back from the spacecraft twice over Madrid and Canberra shortly after entering cruise. The TOS telemetry playback was tried a third time, on October 7, over Canberra and the DSN was able to convolutionally decode and frame sync the data. Both digital and analog tapes of the TOS data were expedited to the TOS Project. The DSN Compatibility Test station within the Merritt Island Launch Area (MILA) complex at KSC played back the TOS analog tape to the TOS POCC at KSC with the same configuration used prelaunch to expedite TOS analysis.

Later DSN analysis indicates that the TOS testing and Network Operations Plan were flawed, leading to the stations' using the wrong Maximum Likelihood Convolutional Decoder (MCD) connection vector configuration during the first two tape-recorder playbacks of TOS data. The MCD connection vector required was Ground Spaceflight Tracking and Data Network ("GSTDN") versus "DSN."

The DSN real-time data delivered over the first 30 days to Mars Observer well exceeded the 97.5 percent required by the Project.

Since November 1992, the DSN tracking has been reduced to one or two 34-m HEF passes per day except for the continuous coverage period around a Flight Software Load centered on January 7-8 and TCM-2 on February 8, 1993. Again the DSN has generally provided good tracking, telemetry and command data exceeding the 97.5-percent requirement.

Several configuration problems have arisen at Madrid during the Mars Observer cruise period. Most of one pass was lost when two Telemetry Processor Assemblies

(TPA's) were configured and concurrently transmitting; another time they were left circular polarization (LCP) instead of right circular polarization (RCP); another time a whole pass was lost for unexplained reasons. Also, scheduling and spacecraft sequencing problems were encountered due to the Madrid strike.

One significant problem with the Mars Observer Camera (MOC) telemetry data has been identified however. The end-to-end spacecraft and ground system is losing a few telemetry data blocks each day which are periodically scattered through the tracking pass. A gap in telemetry data of even one block will cause up to 128 lines of an MOC image to be lost. Depending on the data compression scheme, even more lines could be lost.

A Mars Observer telemetry data loss team was formed and determined that the data abnormalities were one of two types: (1) Unexplained data corruption errors of downlink signal and (2) dropped blocks from the station telemetry system to the MOSO Telemetry Input Subsystem (TIS). Much analysis indicated that the Goldstone uplink was being periodically shifted in phase by a defective Frequency and Timing Subsystem feed. This problem was corrected and the unexplained errors were reduced dramatically but were not completely eliminated. The tiger team is continuing its investigation of this problem.

The dropped blocks between the station telemetry system and the MOSO TIS were identified as three problems associated with the Station Communications Processor (SCP), NASCOM circuits and the SFOC Gateway (DSN) to GCF Interface (MOSO). The Space Flight Operations Center (SFOC) Gateway to Ground Communications Facility (GCF) Interface accounted for the largest amount of numerical blocks lost. Action is currently under way to improve this real-time data flow throughput. TCM-2 was supported without incident by the DSN.

VII. Future Tracking

The plan that has been adopted is for the DSN to continue to track the spacecraft at a one- to two-pass per day rate during the outer cruise mode until July 24, 1993, when continuous 34-m HEF coverage will begin. DSN continuous 34-m HEF coverage will be required through Mars Orbit Insertion (MOI) on August 24, 1993, and will continue until start of the mapping phase on November 22, 1993. The rest of the prime mission is planned to be generally supported by the DSN at a one-to-two 34-m HEF pass per day level until February 2, 1996.

Table 1. Anomalies discovered in testing the spacecraft telecommunications panel.

No.	Anomaly	Characterization/status
1.	Subcarrier harmonics feed-through at -35 dB down from subcarrier in four configurations of the Mars Observer Transponder (MOT).	Verified to occur in Cross-Strapping Unit (XSU) configurations: MOT 1-XSU side 1, MOT 1-XSU side 2 cross-strapped, MOT 2-XSU side 2, MOT 2-XSU side 1 cross-strapped. Demonstrated to be caused by signal leakage by the MOT telemetry on-off relay.
2.	Subcarrier harmonics at all the same amplitude.	Testing indicates 320-kHz subcarrier squarewave distorted at MOT input. Distortion is being caused by overshoot on the squarewave in an XSU amplifier.
3.	Ranging signal feed through on downlink.	Ranging signal is on downlink whenever ranging modulation is on uplink. Signal is being coupled through the MOT ranging modulation on-off relay at ≈ 30 dBc. DSN can lock up ranging with spacecraft ranging switch either in the on or off position.
4.	1.024-MHz signal on downlink all the time about 30 dB down from carrier.	Signal present in all cross-strap configurations and all downlink configurations. The 1.024-MHz signal appeared to be coming from the XSU, but the exact route transferred as downlink modulation was not determined. This anomaly was later fixed by GE Astro and verified by the DSN.
5.	The spacecraft convolutional encoder had a Goddard Space Flight Center (GSFC) connection vector convention instead of the JPL (DSN) standard convention.	The DSN could not lock up the MCD in the planned configuration. The DSN reconfigured the MCD to the GSFC connection vector mode and successfully decoded the data. The DSN agreed to support Mars Observer in this mode.

Table 2. Anomalies discovered in compatibility tests with final spacecraft.

No.	Anomaly	Characterization/status
1.	Spurious signals (≈ -35 dBc) appeared approximately 20 kHz and 12.5 kHz from downlink carrier.	These spurs were thought to be intermodulation products between frequency sources. While not desirable, these were acceptable to the DSN and Radio Science personnel because of their distance from the carrier.
2.	At high signal levels the end-to-end command system could not reliably send and validate long strings of commands to the spacecraft.	The spacecraft configuration is for both MOT/Command Demodulation unit (CDU) strings to be active all the time. The problem was identified as the lower signal level CDU occasionally locking first even when deliberately set for the wrong command data rate. This problem was solved by setting the low signal CDU to 7.8125 bits/sec and offsetting the subcarrier by 2 Hz to 16,002 Hz, therefore locking out this CDU.
3.	The safe mode operation of the spacecraft (7.8125 b/sec in both CDU's) and the 16,000-Hz subcarrier would not reliably send and validate commands.	Limited testing of commanding with ten-sec instead of six-sec spacing indicated 100 percent of commands received.
4.	Long strings of commands radiated by AMMOS at 125 Hz could not be reliably radiated by the DSN.	It was determined that the DSN Command Processor Assembly (CPA) has a limitation which will not reliably radiate long strings of commands in the Mars Observer configuration if closer than 4 to 5 sec. The Project is to transmit commands with 10-sec spacing until new DSN command system software is delivered in June 1993.

500764 May 15, 1993

513-32

167933

P-7

Signal-to-Noise Ratio Losses in Full Spectrum Combining of Signals With a Downconverted Subcarrier

Y. FERIA and J. STATMAN
Communications Systems Research Section

This article presents the results of the signal-to-noise ratio loss in the process of full spectrum combining of signals with a downconverted subcarrier under imperfect conditions. These imperfect conditions not only include the misalignment of the carrier, the subcarrier, and the symbols, but they also include the nonideal filtering in the subcarrier downconversion process, the cutoff of the data bandwidth, and the distortion in signal waveform.

I. Introduction

Arraying techniques have been used to improve the signal-to-noise ratio (SNR) by combining the signals from two or more antennas [3,4]. An overview of arraying schemes was given by Mileant and Hinedi [1], where symbol stream combining, baseband combining, carrier arraying, and full spectrum combining were described. In this work, the full spectrum scheme is employed; however, it differs slightly from the one described in [1]: The carrier frequency here is at a residual carrier frequency rather than at an intermediate frequency, and the subcarrier is downconverted to a lower frequency.

Analytical symbol SNR degradations in the arraying process are given in [1]; the degradations are generally lower than the SNR losses studied here. This study computes symbol SNR losses, the additional symbol SNR needed under the imperfect conditions to achieve the same symbol error rate as under the perfect conditions, through simulations.

This article presents the simulation results of symbol SNR losses in full spectrum combining of signals with a downconverted subcarrier under imperfect conditions. These conditions not only include the misalignment of the carrier, the subcarrier, and the symbols, but they also include the nonideal filtering, the data bandwidth cutoff, and the signal distortion. The simulated results of the SNR losses are compared with the theoretical symbol error probability for binary phase-shift keying (BPSK) for the considered symbol SNR.

The following cases are simulated:

- (1) For reference, a single antenna with known carrier phase, subcarrier phase, and symbol synchronization before the downconversion, and the group delay due to the downconversion as an integer multiple of the original sample period.
- (2) Two identical antennas under the same conditions as the reference.

- (3) Two identical antennas, one with a carrier phase jitter that has a normal distribution with a standard deviation of 0.1, 0.2, and 0.3 rad.
- (4) A 70-m antenna with a 34-m antenna under the same conditions as the reference.
- (5) A 70-m antenna with a 34-m antenna with a delay of one original sample period.

These cases are presented individually later, following a description of the general procedure.

II. General Procedure

A block diagram of the general procedure is depicted in Fig. 1 and is described as follows:

A square-wave subcarrier with a fundamental frequency of 22.5 KHz is modulated by a pseudo-random sequence with a clock time of 1/1000 sec. This signal is then sampled at a rate of 288 KHz and entered as the input to the down-mixing and arraying simulator. The sampled signal is then multiplied by a residual carrier of 100 Hz and its quadrature component to form the in-phase and the quadrature components of the simulated received signal from one antenna. For the received signal from another antenna, a phase jitter of the residual carrier is simulated by adding a random noise with a normal distribution and a standard deviation of σ_ϕ . This phase-jittered residual carrier and its quadrature component are multiplied by the input signal to form the simulated received signal from the other antenna.

The two pairs of the in-phase and quadrature components of the signals received from the two different antennas are then weighted according to the gain-noise temperature ratio of each antenna. A delay is added to one of the in-phase and quadrature pairs to simulate the asynchronization between the two antennas. To each of the signal components, an additive white Gaussian noise is added. The noises are assumed to be independent and identically distributed. The four noise-contaminated signal components are downconverted individually and decimated to a rate of 36 KHz. Note that the decimator outputs represent received signals that are recordable at a low rate. The downconverted in-phase and quadrature components from the two different antennas can now be weighted and added respectively, and this completes the combining part.

To remove the residual carrier, the in-phase and the quadrature components are multiplied by the residual carrier and its quadrature component, respectively. Taking

the delay due to the downconversion into account, the results are then added. Note that the delay, in this case, is considered as an integer multiple of the original sample period, thereby introducing a round-off error. The resulting signal is the input of the symbol detector, which consists of a multiplier of a square-wave subcarrier (at the lower frequency), and an integrate-and-dump filter (IDF). The symbols are finally obtained at the output of the IDF. Comparing the obtained symbols with the original pseudo-random noise (PN) sequence, the ratio of the number of wrong detections over the total number of symbols gives the symbol-error rate.

The five cases of simulation mentioned earlier are all special cases of this general procedure. The simulation conditions and results are presented in the following section.

III. Simulation Conditions and Results

A. A Single Antenna

For reference, a single antenna is simulated first. As a special case of the general procedure, the phase jitter of the carrier is set to zero, and one of the weights is set to one and the other, zero.

The average loss due to the downconversion is found to be 0.28 dB. Within this 0.28-dB symbol SNR loss, about a 0.15-dB loss is due to the data bandwidth cutoff, the other 0.13-dB SNR loss is due to the nonideal filtering, imperfect carrier and subcarrier phase compensation, and signal distortion due to the nonlinear phase of the down-converting system. Note that this loss agrees with the result of 0.28 dB obtained in a previous study [2]. The difference between this simulation and the previous one is that in this case, the residual carrier has a frequency of 100 Hz, whereas in the previous study, this frequency was considered zero.

B. Two Identical Antennas

When combining two identical antennas, all the weights are set to one, and the carrier phase jitter is set to zero.

The average gain over one antenna is 2.6885 dB, which is about 0.3 dB lower than the ideal of a 3-dB gain. This loss is about the same as in the reference since the accuracy of the results is about ± 0.02 dB as discussed in [2].

The results of the above two cases are shown in Figs. 2 and 3.

C. Two Identical Antennas With Carrier Phase Jitter

All the weights are set to one, and the phase jitter is set active with a standard deviation of 0.1, 0.2, and 0.3 rad.

The results are shown in Table 1 and Figs. 4 and 5. The losses include the losses due to the downconversion and the carrier phase jitter.

From Table 1, it can be observed that when the carrier phase jitter is small with a standard deviation of 0.1 rad, the loss due to the phase jitter is practically zero, and it increases gradually as the phase jitter increases.

D. A 70-m Antenna With a 34-m STD Antenna

Taking the larger antenna, the 70-m, as the reference, the weight for the 34-m standard (STD) antenna at S-band (2.2 to 2.3 GHz) is set to $\sqrt{0.17}$ [1]. The phase jitter is set to zero.

The average gain in the arraying is 0.3939 dB, which is 0.2861 dB lower than the 0.68-dB ideal gain. This loss is due to the downconversion and symbol detection as in the reference case.

E. A 70-m Antenna With a 34-m With a Delay of One Sample Period

The weights are set to one and $\sqrt{0.17}$ for the 70-m and the 34-m antennas, respectively. In addition, the delay

of the 34-m antenna is set to the value of one original sample period. This simulates the worst-case scenario of the imperfect delay compensation. Since the compensation can only be made as an integer multiple of a sample period, a misalignment can be a fraction of a sample period.

The result shows that the average gain over the 70-m antenna is about 0.0102 dB, which is 0.67 dB lower than the ideal gain. The loss is mainly due to the downconversion, symbol detection, and the delay. Since the downconversion and symbol detection cause about a 0.3-dB SNR loss, the 0.38-dB SNR loss is due to the delay misalignment.

The above results are shown in Figs. 6 and 7 in terms of symbol error rate versus the symbol signal-to-noise ratio.

IV. Conclusions

This article presents the symbol SNR losses due to the process of arraying of signals with a downconverted sub-carrier. The results show that the losses due to arraying may occur when the carrier phase jitter has a standard deviation greater than 0.1 rad. The loss due to asynchronization between two antennas may cause about a 0.38-dB loss. Under perfect carrier, subcarrier, and symbol alignment, the loss observed is about 0.28 dB, which is mainly due to the downconversion process, and it agrees with the results obtained in a previous study [2].

Acknowledgments

The authors thank Edgar Satorius for his many helpful suggestions and contributions to this work. The authors are also grateful to Sami Hinedi for his help in this work. Special thanks are due to Todd Chauvin, Laura Ekroot, Loretta Ho, and David Watola for their help in the simulation implementation. The JPL CRAY consulting team is also recognized for their very helpful support.

References

- [1] A. Mileant and S. Hinedi, "Overview of Arraying Techniques in the Deep Space Network," *The Telecommunications and Data Acquisition Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 109–139, February 15, 1991.

- [2] Y. Feria and J. Statman, "Symbol Signal-to-Noise Ratio Loss in Square-Wave Subcarrier Downconversion," *The Telecommunications and Data Acquisition Progress Report 42-112*, vol. October-December 1992, Jet Propulsion Laboratory, Pasadena, California, pp. 74-82, February 15, 1993.
- [3] J. S. Ulvestad, "Phasing the Antennas of the Very Large Array for Reception of Telemetry from Voyager 2 at Neptune Encounter," *The Telecommunications and Data Acquisition Progress Report 42-94*, vol. April-June 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 257-273, August 15, 1988.
- [4] W. J. Hurd, J. Rabkin, M. D. Russel, B. Siev, H. W. Cooper, T. O. Anderson, and P. U. Winter, "Antenna Arraying of Voyager Telemetry Signals by Symbol Stream Combining," *The Telecommunications and Data Acquisition Progress Report 42-86*, vol. April-June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 131-142, August 15, 1986.

Table 1. Simulation conditions.

Carrier phase jitter, σ_ϕ	0.1	0.2	0.3
SNR gain, dB, over 1 ideal antenna	2.7241	2.5496	2.4280
SNR gain, dB, over 1 simulated antenna	2.9999	2.8254	2.7247
SNR loss, dB, compared to 2 ideal antennas	0.2759	0.4504	0.5720
SNR loss, dB, compared to 2 simulated antennas	-0.0356	0.1389	0.2605

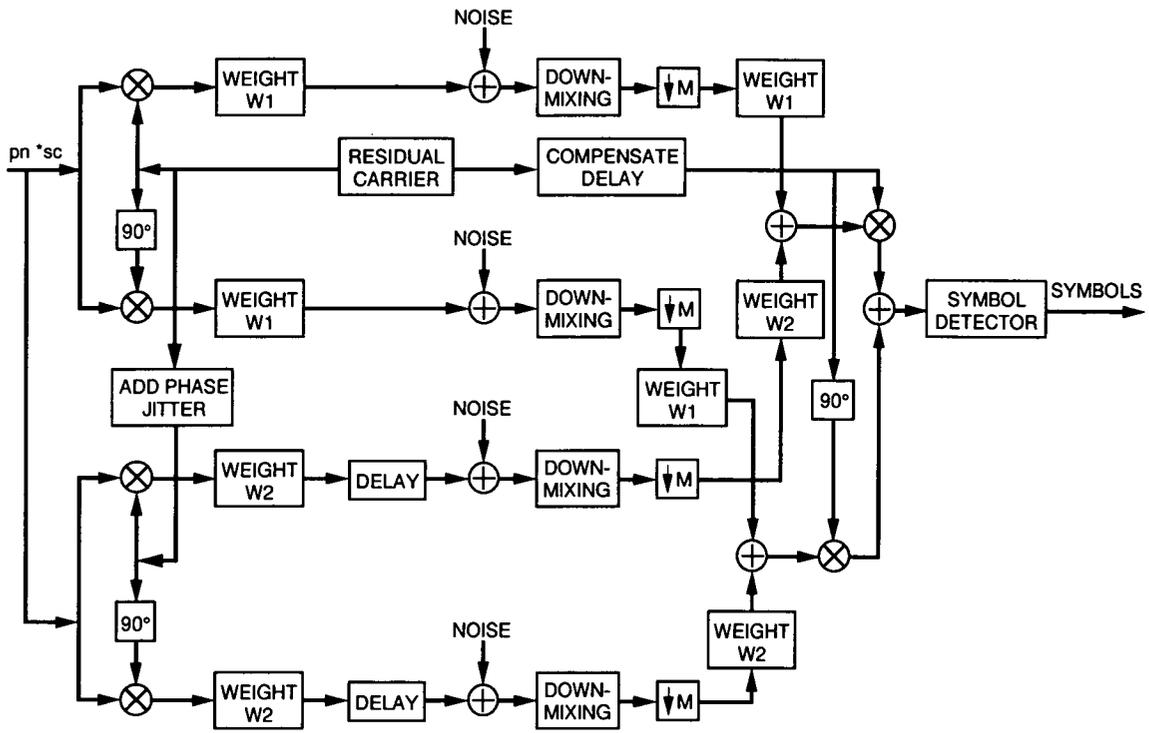


Fig. 1. The general procedure.

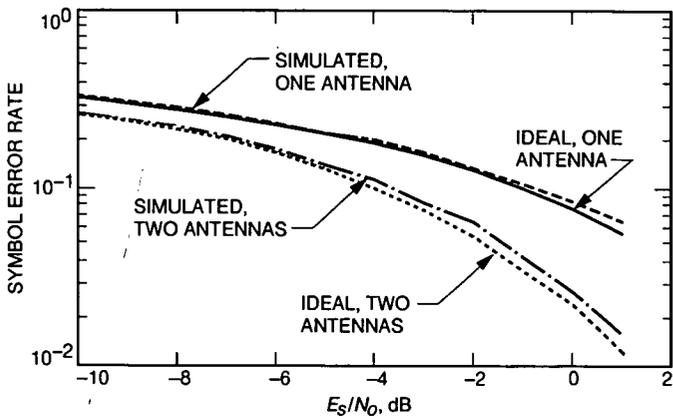


Fig. 2. Combining two identical antennas.

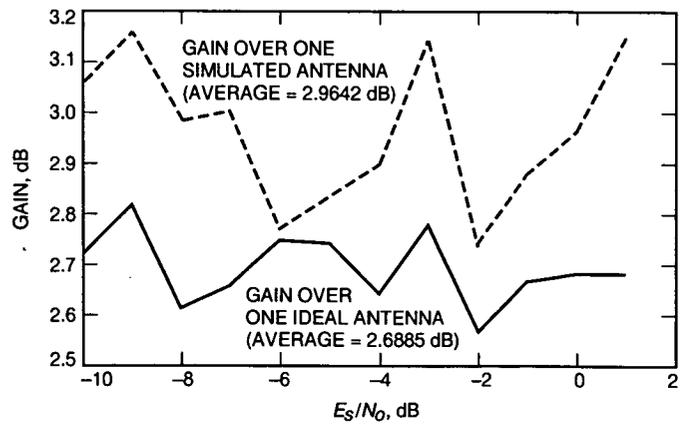


Fig. 3. Gain from combining two identical antennas.

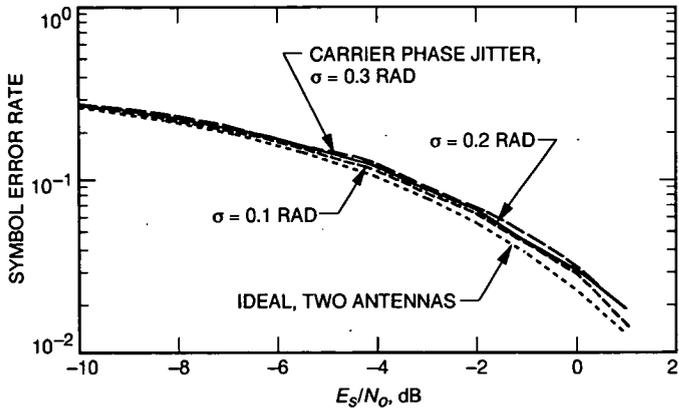


Fig. 4. Combining two identical antennas, one with carrier phase jitter.

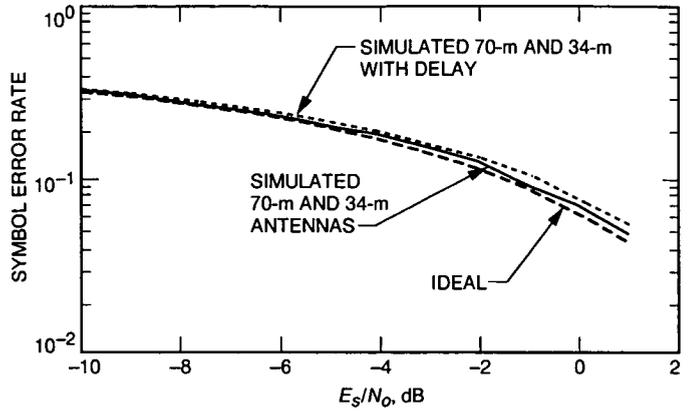


Fig. 6. Combining a 70-m and a 34-m antenna.

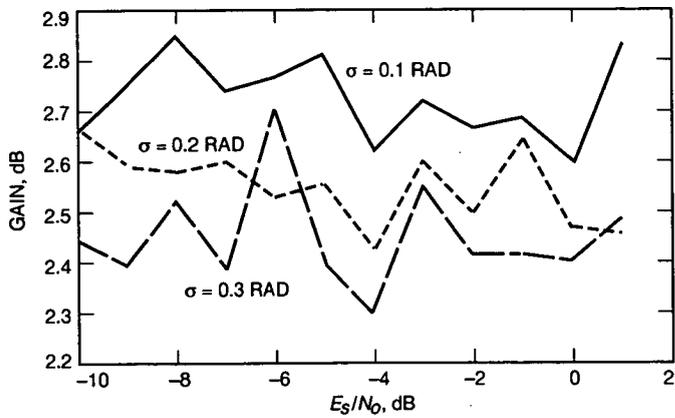


Fig. 5. Gain, dB, over one ideal antenna.

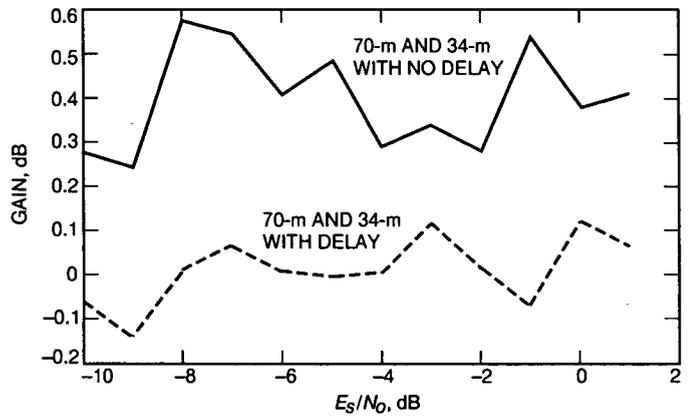


Fig. 7. Gain over the 70-m antenna.

514-46
167934
P-19

500758

DSS-14 Subreflector Actuator Dynamics During the Landers Earthquake

R. Levy and D. Strain

Ground Antennas and Facilities Engineering Section

The June 28, 1992, Landers earthquake ground motion records at the Echo site (DSS-12 antenna) were adjusted to provide a better match with spectra from the measured Mars site (DSS-14 antenna) instrument tower response. A finite-element model of the antenna structural system was analyzed for response to this ground motion. Dynamic forces and displacements were computed in the locality of components that had failed during the earthquake. Calculated forces in the range of from 30,000 to 35,000 lb on failed Y-axis actuator U-joints were consistent with laboratory load tests. The load capacity of these joints was found to be below the range of from 34,000 to 42,000 lb. Dynamic amplification factors of from 6 to 16 were computed for the quadripod apex accelerations with respect to the ground accelerations. The largest amplification factor—25—was found at the outboard end of the X-actuator.

I. Introduction

The June 28, 1992, Landers earthquake shock arrived at the JPL Goldstone Complex at about 5 a.m. PDT, causing failures within components of the Mars antenna subreflector positioner assembly and damage to the subreflector. The antenna is shown in Fig. 1 at a 15-deg elevation, which was the attitude at the time of the earthquake. The epicenter was about 55 kilometers to the southeast of the antenna complex, and the main shock, of magnitude 7.4, occurred at the Camp Rock-Emerson Fault and its southern extension. Strong aftershocks occurred within several hours, but the damage at the antenna is attributed to the primary shock.

A damage investigative and corrective action team issued a report.¹ That report contains a preliminary analysis of the subreflector-positioner assembly dynamic re-

sponse to the earthquake. Since the report, a comprehensive finite-element model of the antenna structural system with subreflector, positioner, and concrete pedestal has been completed. The response derived from this model is contained herein. Small amounts of related coverage and material from the action team report² are reproduced here for convenient referral. The reference, however, covers more of the background, details, and proposed remedial actions.

Figure 2 contains schematics of the subreflector and positioner drive assembly, which were subject to major damage during the earthquake. Figure 3 shows six of the areas

¹ Engineering Analysis and Corrective Action Team, *DSS-14 70-meter Antenna Earthquake Damage Investigation*, vols. 1 and 2, JPL Report 890-251 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 25, 1992.

² Ibid.

Z-axis actuator U-joints failed and have been circled in the figure. Another circle has been placed at the X-axis drive to indicate that the supporting bracket was bent. The X-axis actuator shaft was also bent, and four of the six safety cables failed when the actuator failures allowed the subreflector to drop. The axes denoted in these two figures are local axes fixed on the antenna: the Z-axis is the antenna-pointing axis, the X-axis is parallel to the elevation axis, and the Y-axis is vertical (positive upwards) when the antenna points to the horizon. At the 15-deg elevation angle, in effect at the time of the earthquake, the Y-axis actuators support almost all of the static weight of the positioner and the subreflector.

The subreflector impacted against a quadripod leg, which caused some damage to the subreflector and the leg. The subreflector was eventually wedged against the quadripod leg and supported by one of the intact safety cables, and possibly by the remaining Z-axis actuator and the damaged X-axis actuator. A failure scenario assembled in the action team report³ assigns the first failure to the left side (as oriented in Fig. 3) Y-axis U-joint; the next failure (attributed to effects of transference of gravity loading and a rotation) was at a remaining Y-axis actuator U-joint. Then, it was proposed, the top (Fig. 3) Z-axis actuator U-joint failed, and this was followed by the impact of the falling torus on the X-actuator bracket. After this, the scenario continues to explain further events and observations of additional damage.

The original design called for identical Y- and Z-axis U-joints and a lower capacity design for the X-axis U-joints. The four intact Z-axis U-joints, one Y-axis U-joint, and both X-axis U-joints were returned to the laboratory for failure load tests. The failure loads for the five Z- or Y-axis U-joints were in the range of from 34,350 to 42,290 lb in tension, with a mean of 38,538 and standard deviation of 3,320 lb. The catalog value of the tension-rated load for these U-joints was 54,000 lb, and they were supposedly capable of withstanding temporary loads of 150 percent of the rated load. Therefore, it is evident that these U-joints were substandard. One X-axis U-joint failed in tension at 30,800 lb, and the other failed at 48,830 lb in compression. The foregoing values are all based upon the U-joints that survived the earthquake. Hence it could be concluded that in the field, the U-joints failed from the earthquake at loads less than those of the tests.

II. Ground Motion Time History Records

A California Institute of Technology Seismological Laboratory instrumentation station provided three-axis ac-

³ Ibid.

celerograph records of the ground motion in the north-south, east-west, and vertical directions at the Echo Site, which is about 6 kilometers to the south of the Mars antenna. These records are plotted in Figs. 4(a), (b), and (c). Another set of tri-axial accelerograph records was also obtained at an intermediate-level platform of the Mars antenna instrument tower, which is shown in Fig. 5. Records are available for 40-sec durations at both locations and are digitized at the rates of 100 per sec for the Echo site and 200 per sec for the Mars site.

The maximum accelerations recorded at Echo were 4.8 percent of gravity (0.048 g) for the north-south component, 0.058 g for the east-west component, and 0.041 g for the vertical component. These are not large accelerations in comparison with those of past damage-producing earthquakes. However, the portions of the records showing the strongest motion are of long enough duration to be damaging.

The maximum accelerations recorded at the instrument tower platform were 0.20 g in a "longitudinal" direction, 0.16 g in a "transverse" (lateral) direction, and 0.16 g in the vertical direction. These records provide the responses at the platform, which in fact represent some amplification of the actual earthquake ground motion. Had the records been available for the site ground motion, rather than for the response to the ground motion, they would have served as the input for the analytical dynamics investigation of the earthquake response of the antenna.

III. Adjustments of Ground Motion Records

The Echo records would be an obvious choice for estimating the ground motion at the Mars antenna site, provided that the 6 kilometers of intervening strata to the Mars site were homogenous and not subject to discontinuities in physical properties. To test this, an analytical model of the instrument tower structure was assembled and subjected to the excitation defined by the Echo records. The purpose was to compare the output of the model derived at the platform level with the available records of tower response to the earthquake.

The analytical model was developed from a preliminary 12 degrees-of-freedom representation that was later adjusted to improve the agreement with field modal tests [1]. Since the tower appeared to be essentially symmetrical with respect to any two orthogonal vertical planes, a single two-dimensional model (height dimensions and horizontal displacements) was estimated to be adequate for vibratory motions parallel to the ground. The field testing program excited and measured the tower vibrations in the longitudinal and lateral directions. A separate one-dimensional

model (height dimensions and vertical displacements) was developed for vertical vibration modes, but it was not possible to excite vertical mode vibrations in the field to check or adjust this model. The lowest frequencies (more than 15 Hz) computed from the vertical vibration model appear to be much higher than typical important excitation frequencies of earthquake ground motions.

Figure 6 contains comparisons of frequencies and mode shapes for the first three vibration modes from measurement and from the adjusted model. The frequencies are in agreement, as are the first two mode shapes. The third mode shapes differ, but this is not important because the frequency is above the range of significant earthquake excitation. There is a small lack of symmetry in the as-built tower with respect to the longitudinal and lateral directions. Consequently, the measured frequencies are slightly different for these two directions, and the tabulations in the figure show the ranges.

The recorded Echo north (north-south) and east (east-west) acceleration records were integrated to provide the associated ground displacements, which were then used as input to the model. Differences in the power spectral densities (PSD) of the computed acceleration responses were evident at the 3.5- and 10-Hz first- and second-mode natural frequencies when the spectral densities were compared with the PSD obtained from the direct Mars-site measurement records. The Echo acceleration records were then adjusted to reduce these differences by supplementing them with linear combinations of trial acceleration vectors. The combining factors were determined by a nonlinear least squares optimization to minimize the differences in the 1-Hz to 12-Hz frequency range between the PSD computed from the tower model and the PSD determined from the direct measurements. Figures 7(a) and 8(a) show the PSD curves obtained from the direct measurements and the curves obtained by calculations that used the original Echo ground motions as input. Figures 7(b) and 8(b) show improved agreement of the measurements with the calculations obtained by using the adjusted Echo records as inputs to the tower model. Figures 9 and 10 show the power spectra for the original and adjusted Echo ground accelerations. It can be seen that the adjustments provided only moderate changes to the original Echo data. The assumptions made in processing all of the foregoing data were (1) the ground motions could be assumed to be approximately stationary during the 15-sec to 26-sec time interval (11-sec) of the record, during which time all the motion records were judged to be the most intense, and (2) the effective damping ratio for the tower was 2 percent.

IV. Finite-Element Model

The finite-element method (FEM) design model for the antenna-tipping structure was developed in the mid-1980's. It contained 5500 degrees of freedom and represented one-half of the full structure. More recently, this model was extended to include the full structure, and the original set of three translational degrees of freedom at each node was supplemented by three more rotational degrees of freedom. This model, which now contains 38,000 degrees of freedom, runs on the NASTRAN structural analysis computer program and requires a Cray computer to accommodate the problem size readily and to provide a reasonable turnaround time. This model includes a representation of the subreflector-positioner component with 2500 degrees of freedom and a concrete pedestal component with an additional 11,000 degrees of freedom. These latter two components were successfully incorporated within the primary structure model after the publication of the action team report.⁴

Figure 11 represents a side view of the complete antenna model at the 15-deg earthquake incidence elevation. There are about 8000 nodes and 17,000 elements (13,000 bar elements, 2000 plate elements, and 2000 solid elements) in this model. Figure 12 shows a close-up of the quadripod apex, subreflector-positioner, and subreflector regions of the model. The tilt of the subreflector relative to the apex is standard because of the tri-cone configuration. Figure 13 shows a side-view cross-section of the pedestal component model, which consists entirely of solid elements.

V. Dynamic Analysis

The finite-element model problem (38,000 degrees of freedom) was processed by the NASTRAN structural analysis program on the Cray computer. The NASTRAN processing consisted only of solving the eigenvalue problem for natural frequencies and mode shapes for the 40 lowest modes. The NASTRAN output was post-processed by a FORTRAN program written to compute the elastic-rigid coupling matrix and effective modal weights according to algorithms provided in [2]. Selected components of the eigenvectors and modal forces in the actuator members were also extracted. All of the post-processed data were written in ASCII files and transferred to a 386 PC, where the dynamic response analysis to the earthquake was readily developed.

Conventional modal superposition analysis was used to generate response time histories of actuator forces and a

⁴ Ibid.

subset of displacements at nodes to which the actuators were connected. The physical system was described by the generalized modal masses, elastic-rigid coupling matrices, and natural frequencies. The excitation to the system was provided by the Echo three-axis ground acceleration time histories as described above.

Two sets of ground motion time histories were processed; one set contained the original recorded Echo motions, and the other was the adjusted Echo motions. The model had been rotated about the elevation axis to the 15-deg elevation of the earthquake incident, so that the vertical earthquake records were applied along the model Z -axis. The antenna azimuth at earthquake incidence was 67 deg clockwise from north. Therefore, to invoke a consistent assumption, the Echo north-south and east-west data were rotated about a vertical axis to be aligned with the antenna model's north and east axes. Table 1 summarizes the maximum and minimum ground accelerations that were input to the models in the direction of the antenna axes after rotation.

VI. Computed Dynamic Actuator Forces

Three major assumptions in the computations were

- (1) A power failure at the start of the earthquake caused an automatic application of the antenna brakes. The dynamic effects on the structure were assumed to be sufficiently attenuated at the time of the stronger ground motions to be ignored in the analysis.
- (2) The damping ratio for all modes was 1 percent. This is based upon preliminary field measurements, during which the system was subjected to a sudden imposition of the brakes.
- (3) All components remained intact during the linear-response analysis. Consequently, this assumption implies that the analysis can be meaningful only up to the time of the first component failure or yielding.

Figures 14 (a) through (f) show the dynamic force histories for the six actuators, based upon analysis of the full model with the adjusted Echo ground motions as the input excitations. The static weight loadings at the 15-deg elevation of -670, -589, 7460, 0, 11,750, and 11,420 lb have been included in the figure. Table 2 contains a summary of calculated dynamic forces based upon the conditions used to develop Fig. 14 and also from several other cases.

VII. Computed Dynamic Displacements and Accelerations

A few selected displacement time histories are shown in Fig. 15. All of these are in relative coordinates with respect

to the ground motions. The total displacements are the sum of the relative displacements and the ground motion displacements. Although displacements in Figs. 15(a) and (b) are at the outboard ends of the actuators, the displacements at the inboard ends differ by only small percentages. The apex displacements in Figs. 15(c), (d), and (e) were obtained by averaging the displacements for a pair of nodes to each side of the center of the apex. The peak displacements of the X -axis actuator are about 50 percent greater than for the peak X -axis displacements of the apex. This could be because there is only a single X -axis actuator, which is supported by a relatively flexible bracket at its inboard end. On the other hand, the Y -axis actuator displacements are only slightly more than the apex Y -axis displacements.

The accelerations corresponding to the above relative displacements are shown in Fig. 16. To obtain the absolute accelerations, relative accelerations should be supplemented by adding the ground motion accelerations. Referring back to Table 1, it appears that the accelerations of the X -axis actuator indicate an amplification factor of about 25 with respect to the ground acceleration, and the Y -axis actuator amplification factors are about 17. The apex amplification factors are about 12, 16, and 6 in the X , Y , and Z directions, respectively.

VIII. Summary and Conclusions

The June 28 earthquake ground motion records at the Echo site were adjusted to provide output power spectra that were a close match to the power spectra of measured Mars site instrument tower records at a distance of 6 kilometers. A comprehensive finite-element method analytical model of the Mars antenna structural system was subjected to the excitation of the adjusted Echo ground motion. Dynamic analyses were completed to provide response time histories of forces and displacements associated with subreflector-positioner actuator components, some of which had failed during the earthquake.

The computed peak dynamic loads in the range of 30,000 to 35,000 lb for the two Y -axis actuators for which the U-joints had failed during the earthquake are consistently lower than laboratory failure load tests on surviving U-joints, which were in the range of 34,000 to 42,000 lb. The next largest forces from the analyses, of about 25,000 lb, were for the X -axis actuator. Nevertheless, this was still below the laboratory failure test load of 30,800 lb. Thus, it seems that the X -axis actuator could have survived undamaged from the earthquake except for the subsequent impact of the falling torus structure. In particular, these calculated maximum dynamic

the U-joints had failed during the earthquake are consistently lower than laboratory failure load tests on surviving U-joints, which were in the range of 34,000 to 42,000 lb. The next largest forces from the analyses, of about 25,000 lb, were for the X-axis actuator. Nevertheless, this was still below the laboratory failure test load of 30,800 lb. Thus, it seems that the X-axis actuator could have survived undamaged from the earthquake except for the subsequent impact of the falling torus structure. In particular, these calculated maximum dynamic forces are consistent with the failure scenario of the action team report.⁵ All of the foregoing calculated results are sensitive to the structural damping. Here the damping ratio was assumed to be 1 percent based upon a limited amount of measured data.

Although the X-axis actuator could have survived the earthquake, the analysis for the acceleration amplification factors shows that the X-axis actuator is excessively flexible. The amplification factor for the X-axis actuator was about 25 times with respect to the input X-axis ground

acceleration. This was about twice the amplification factor for the quadripod. The remaining Y- and Z-axis amplification factors for the quadripod were about 16 and 6, respectively. The Y-axis actuators, although deficient in load-carrying capacity, did not produce significantly different acceleration amplification factors from the quadripod apex.

The loads and displacement analyses here concentrated on the actuators and U-joints of the subreflector positioner. A separate and lengthier study would be needed to examine the effects of this earthquake on the complete structure. Nevertheless, it is possible that the examination here of the quadripod apex and subreflector positioner components may have considered the most vulnerable regions of the antenna. Future analysis of the seismic safety of the antenna would, of course, require an evaluation of the effects possible for a much broader range of seismic excitations than the single event considered here.

⁵ Ibid.

Acknowledgments

JPL Engineering Analysis and Corrective Action (EACA) team members and contributors who provided background and evaluations are acknowledged specifically in the action team report⁶ and its distribution list. We also wish to specifically recognize California Institute of Technology staff members who participated in the study—in particular, Professors J. Hall and W. Iwan for the suggestion that the Mars instrument tower response records be employed to particularize the site ground motion; Professor G. Housner, who proposed that the Echo ground motion could be used as a basis with supplemental adjustments to provide consistent tower response spectra; and Senior Research Fellow E. Hauksson, who provided the Echo ground motion records.

References

- [1] *Experimental Modal Analysis of the DSS-14 Instrument Tower*, Report R9227-6405, AGBABIAN Associates, Pasadena, California, September 8, 1992.
- [2] R. Levy, "Computation of Vibration Mode Elastic-Rigid and Effective Weight Coefficients from Finite-Element Computer Program Output," *The Telecommunications and Data Acquisition Progress Report 42-105*, vol. January-March 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 13-20, May 15, 1991.

⁶ Ibid.

Table 1. FEM ground accelerations after axis rotations.

Measurements		Accelerations, g		
		X-axis	Y-axis	Z-axis
Original Echo data	maximum	0.041	0.054	0.027
	minimum	-0.037	-0.053	-0.041
Adjusted Echo data	maximum	0.044	0.059	0.027
	minimum	-0.036	-0.058	-0.041

Table 2. Summary of extreme forces in the subreflector-positioner actuators.

Configuration	Actuator loads, kilo-pounds					
	Right-Z ^a	Left-Z ^a	Top-Z ^a	X	Left-Y	Right-Y
With adjusted Echo motion and pedestal:						
maximum	8.3	10.0	9.0	25.5	34.6	31.5
minimum	-8.6	-9.7	-11.3	-25.3	-12.7	-7.2
With original Echo motion and pedestal:						
maximum	7.5	9.3	8.9	23.7	32.8	30.4
minimum	-8.3	-9.2	-10.5	-23.6	-11.3	-6.3
With adjusted Echo motion but no pedestal:						
maximum	7.7	10.9	9.1	25.5	33.5	29.7
minimum	-6.7	-11.0	-11.0	-25.6	-11.5	-7.4
With original Echo motion and no pedestal:						
maximum	7.6	10.2	8.6	23.9	31.8	29.1
minimum	-6.7	-9.8	-10.1	-24.0	9.4	-6.0

^a Does not include static weight loading at 15-deg elevation.

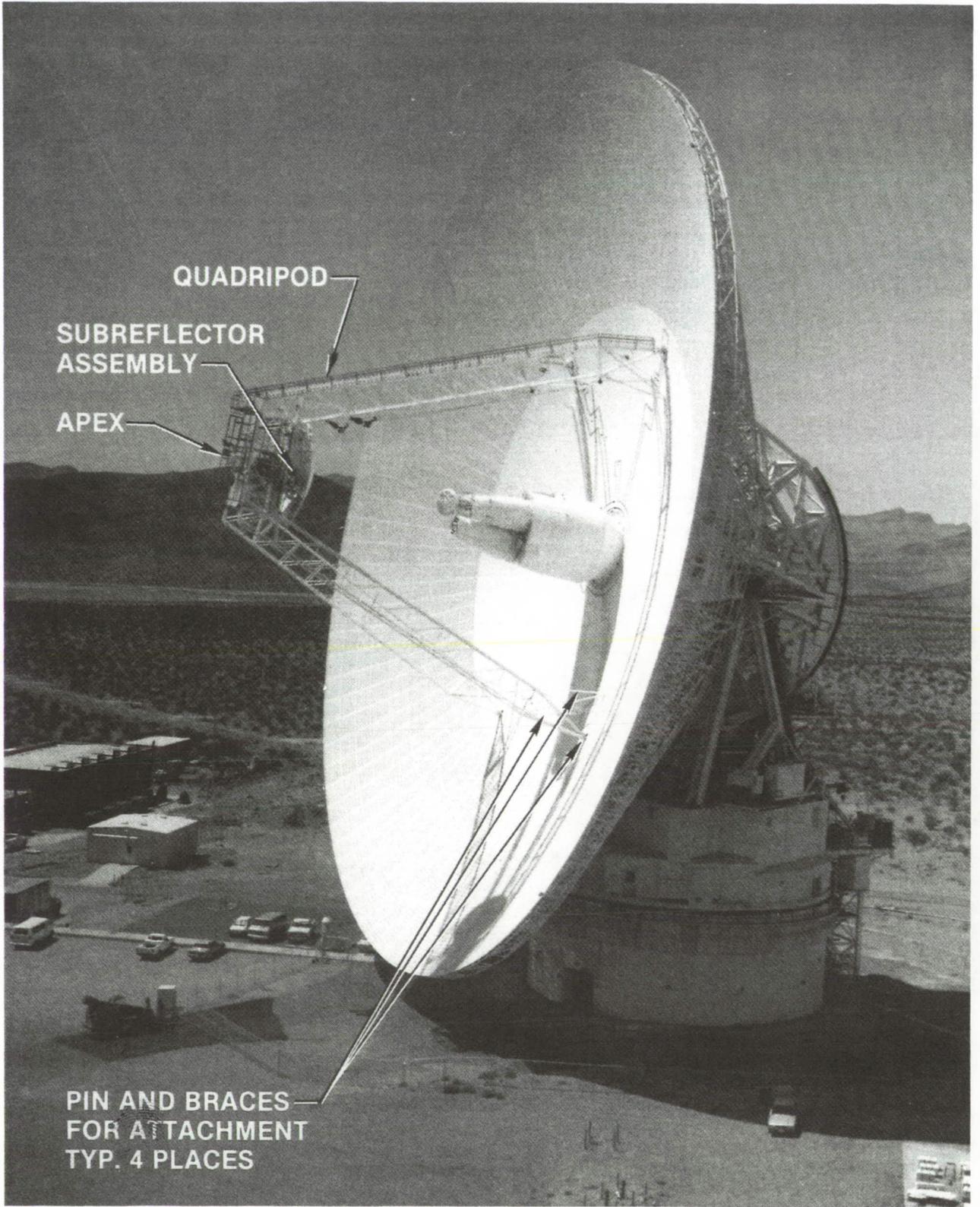


Fig. 1. DSS-14 (Mars) antenna at earthquake position.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

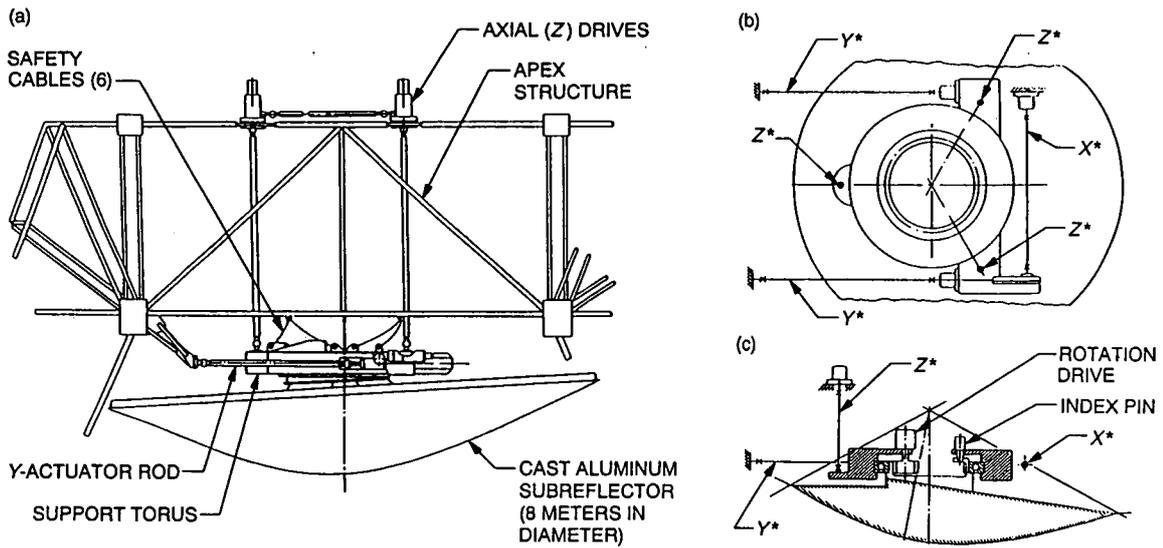


Fig. 2. Subreflector and positioner assembly: (a) elevation view, (b) plan view showing rotation bearing and actuator rods (with asterisks), and (c) elevation section showing rotation bearing and actuator rods (with asterisks).

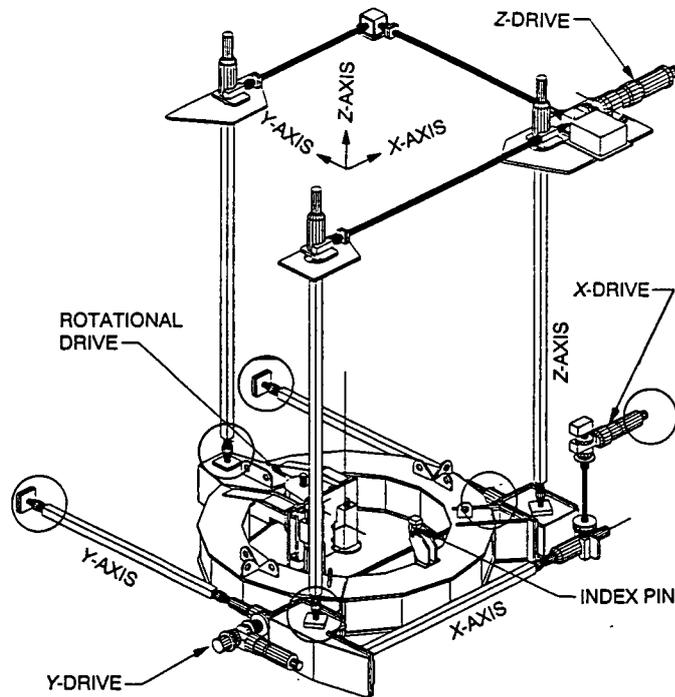


Fig. 3. Subreflector positioner damage points.

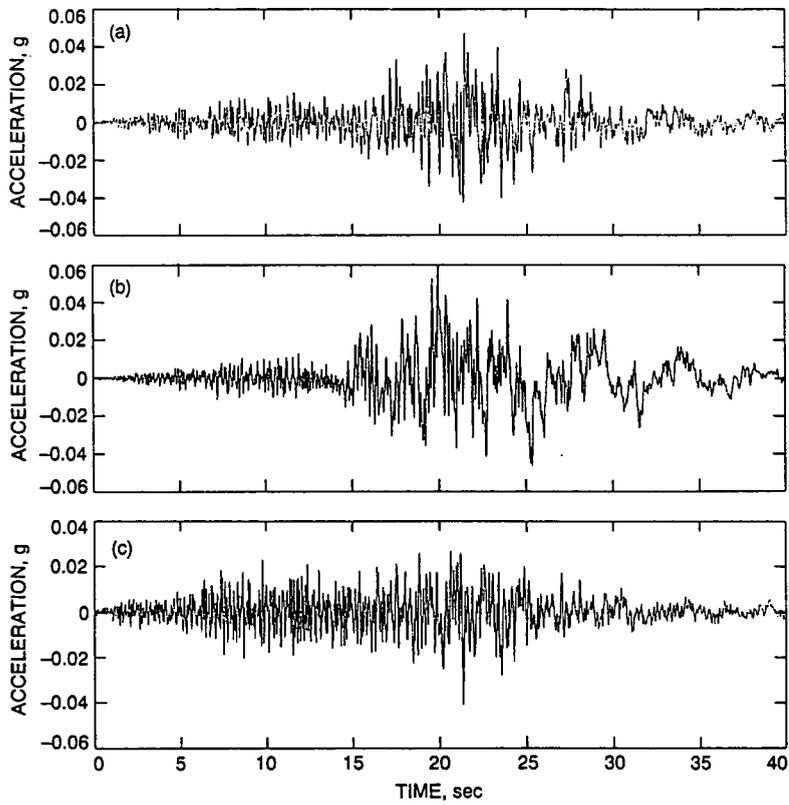


Fig. 4. Echo recorded ground motion for (a) north-south, (b) east-west, and (c) vertical.

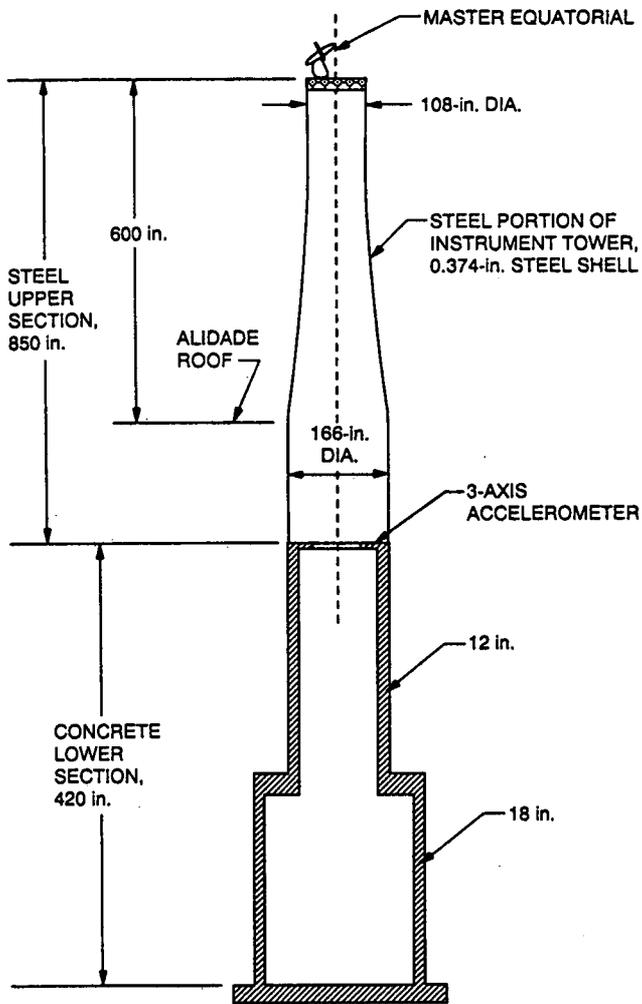


Fig. 5. Mars antenna instrument tower.

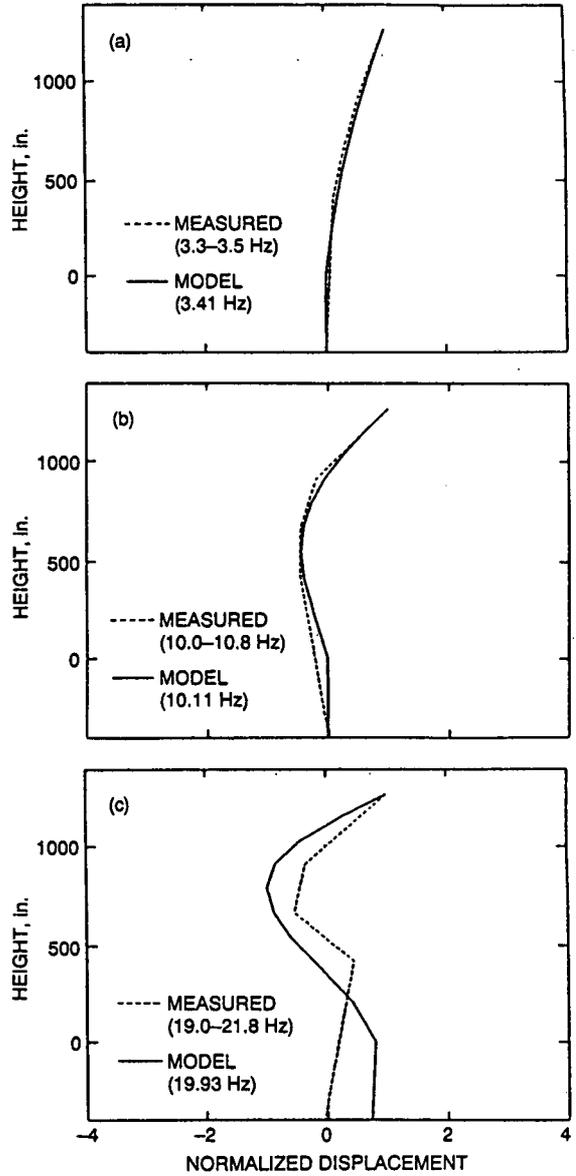


Fig. 6. Instrument tower measured and modeled vibration modes: (a) first mode, (b) second mode, and (c) third mode.

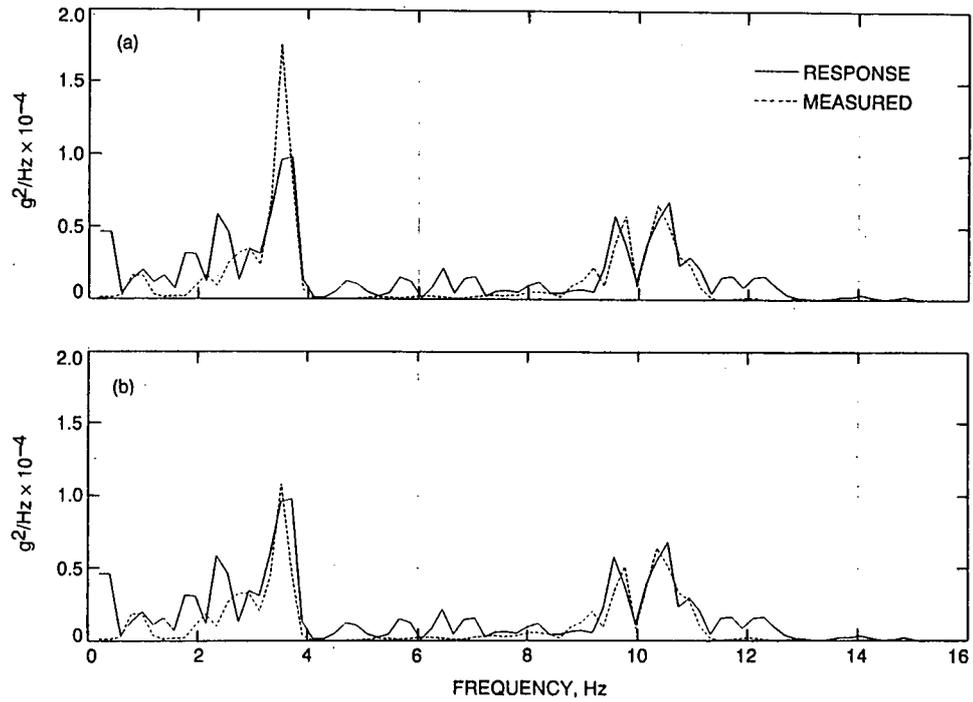


Fig. 7. Tower lateral response spectral densities: (a) original Echo north and (b) adjusted Echo north.

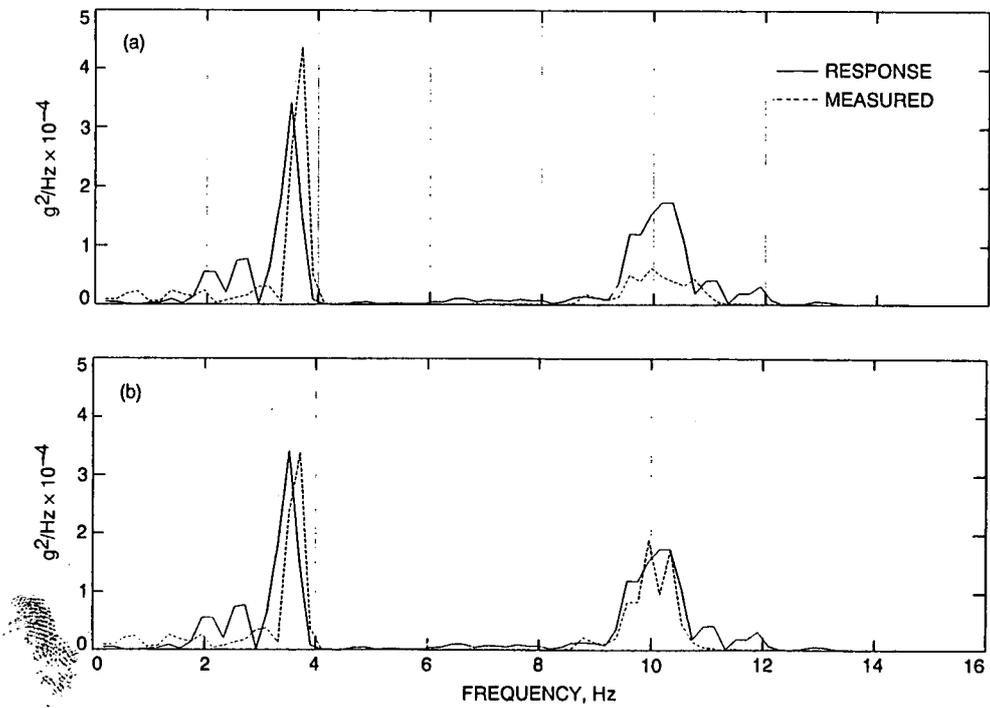


Fig. 8. Tower longitudinal response spectral densities: (a) original Echo east and (b) adjusted Echo east.

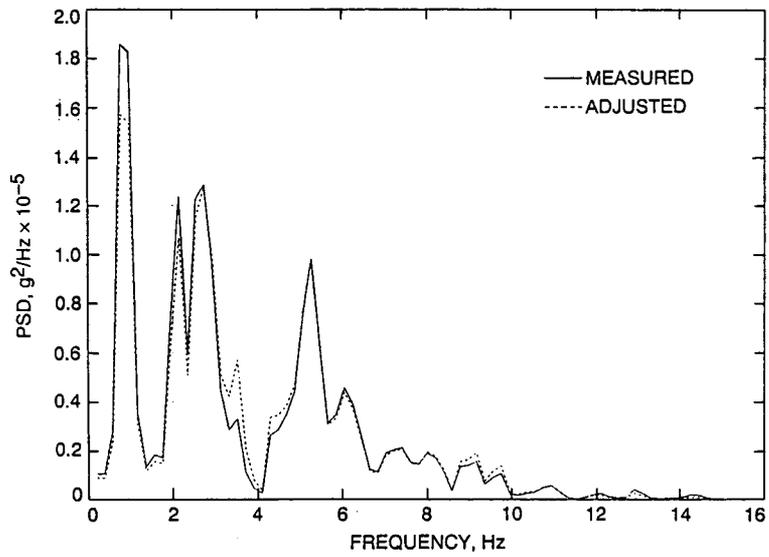


Fig. 9. Power spectra, Echo north record.

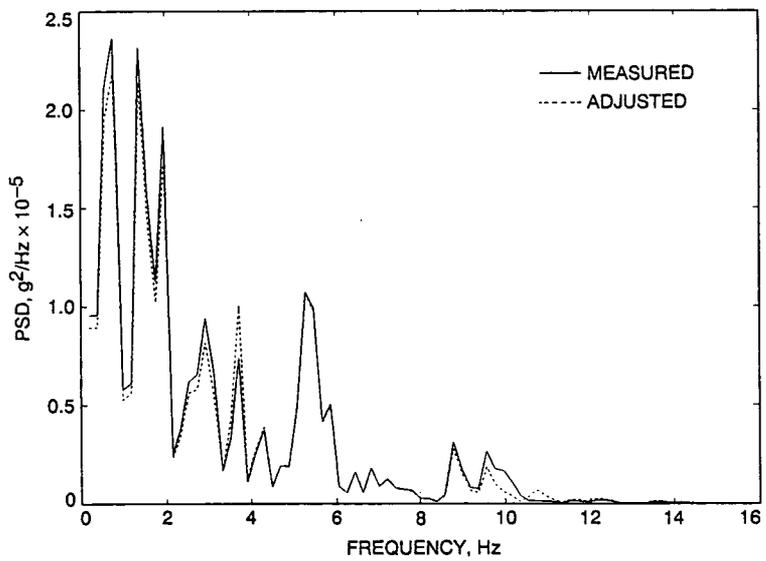


Fig. 10. Power spectra, Echo east record.

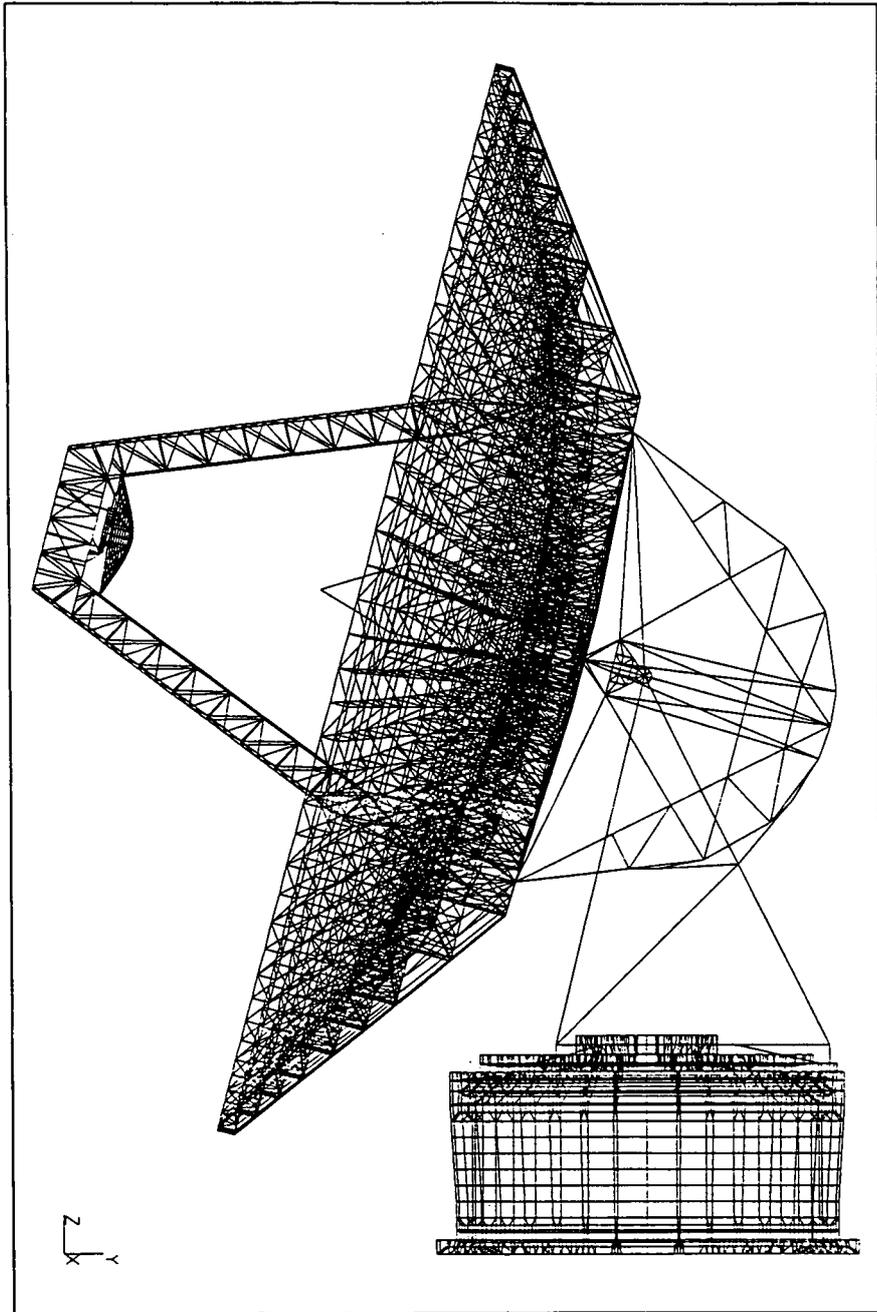


Fig. 11. Mars antenna finite-element model.

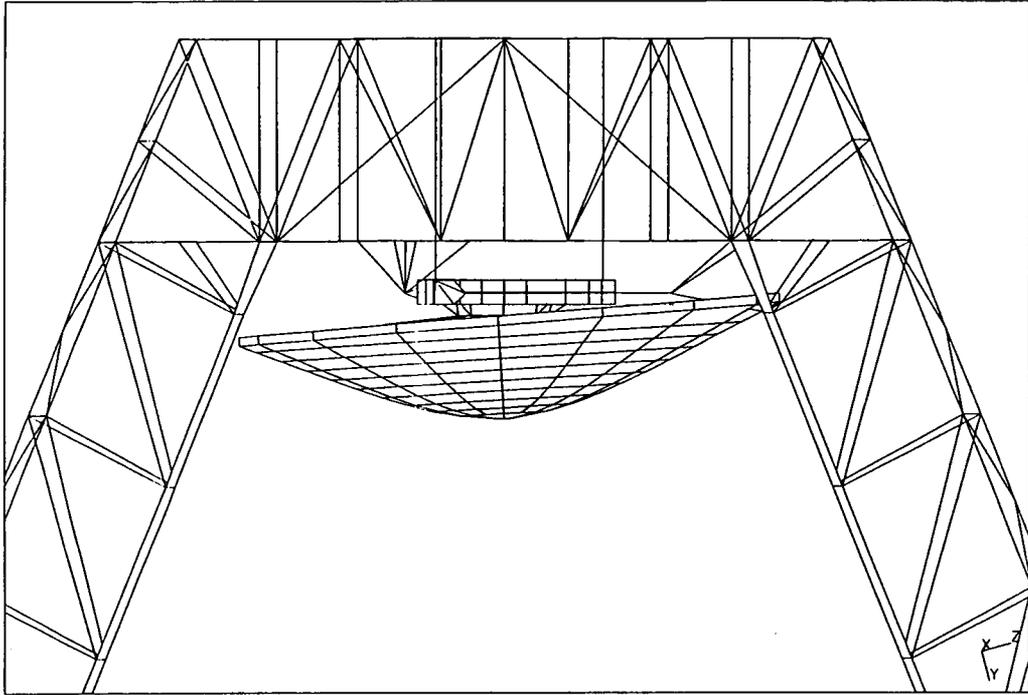


Fig. 12. Quadripod apex and subreflector components of the model.

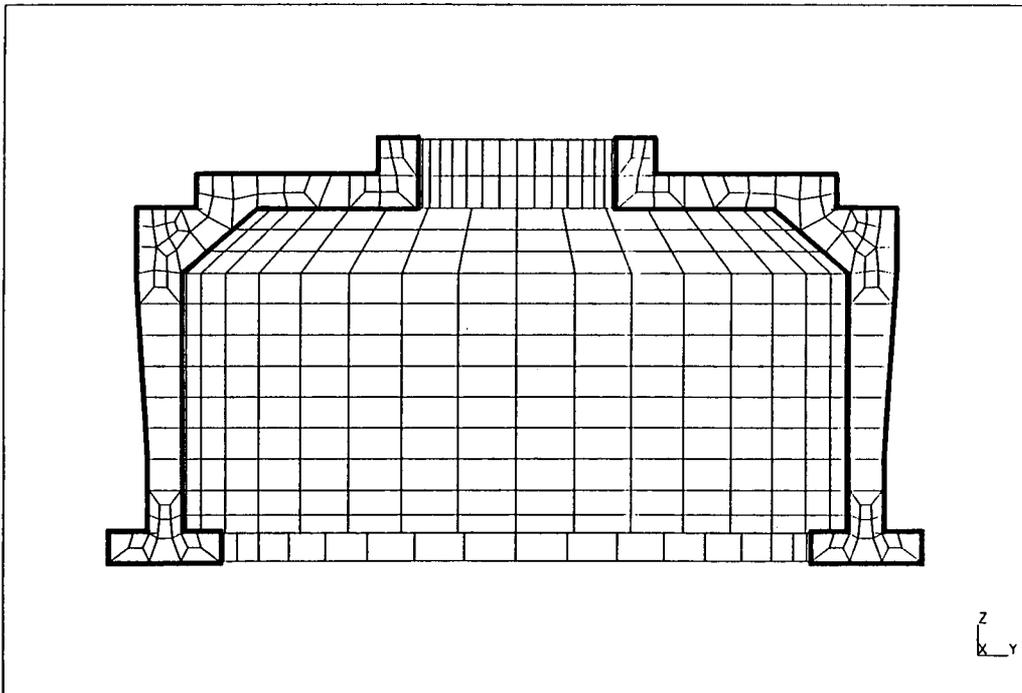


Fig. 13. Pedestal component of the model.

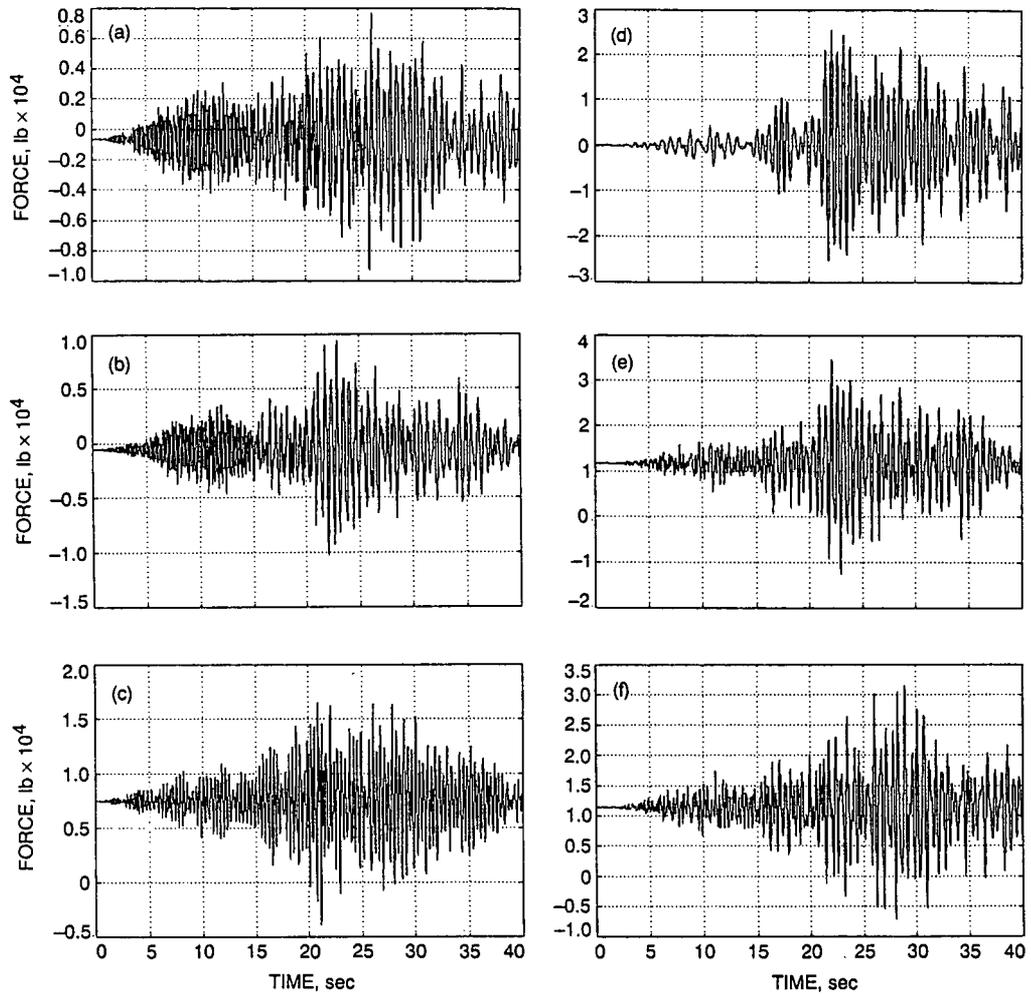


Fig. 14. Dynamic forces (a) on the right Z-actuator, (b) on the left Z-actuator, (c) on the top Z-actuator, (d) on the X-actuator, (e) on the left Y-actuator, and (f) on the right Y-actuator.

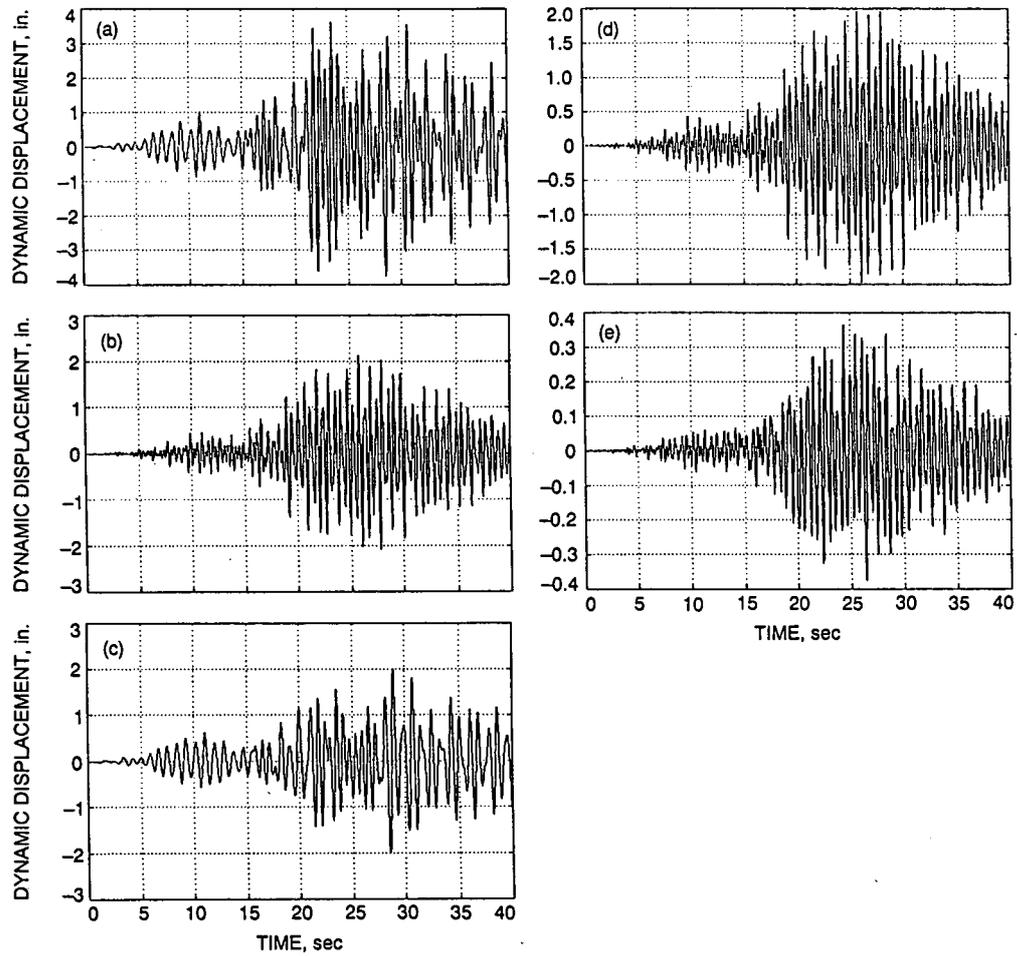


Fig. 15. Dynamic displacements: (a) outboard end—X-actuator, (b) left side of the outboard end—Y-actuator, (c) average apex—X direction, (d) average apex—Y direction, and (e) average apex—Z direction.

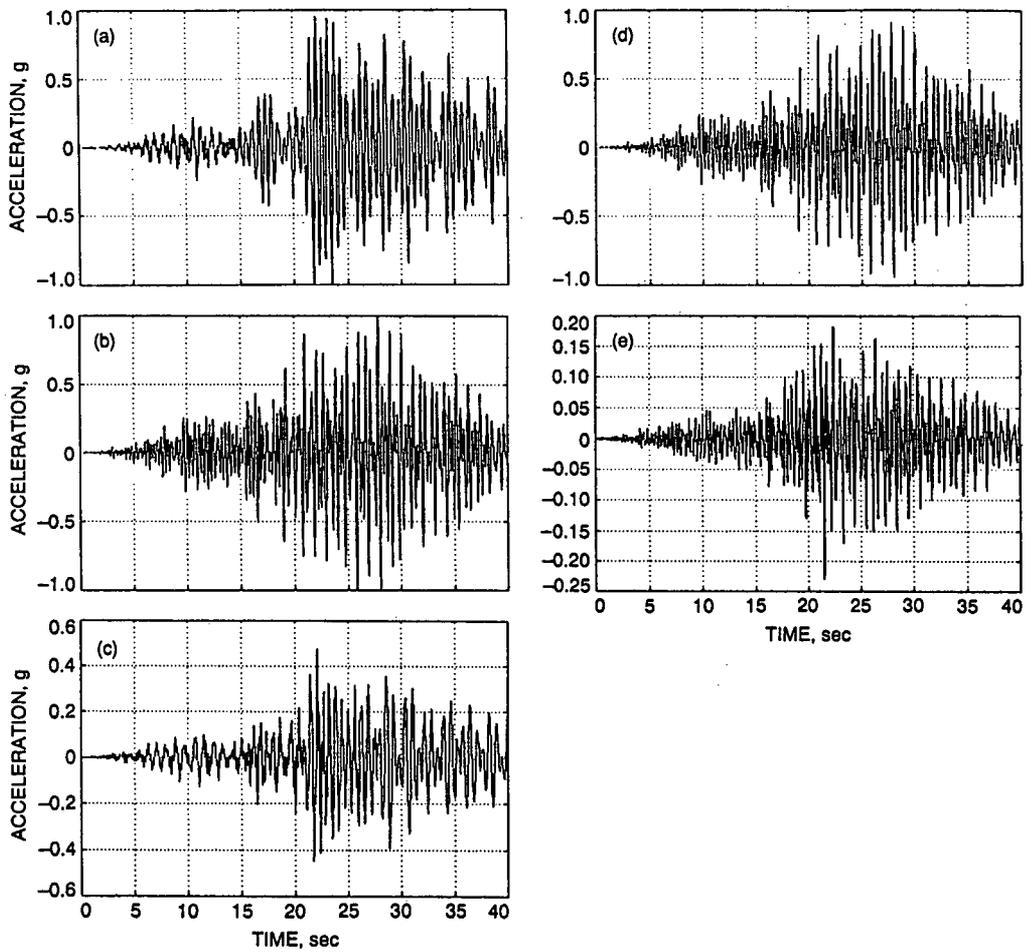


Fig. 16. Dynamic accelerations: (a) X-actuator, (b) left-side Y-actuator, (c) apex—X direction, (d) apex—Y direction, and (e) apex—Z direction.

515-37
167935

p. 27

An Analysis of the 70-Meter Antenna Hydrostatic Bearing by Means of Computer Simulation

R. D. Bartos

Ground Antennas and Facilities Engineering Section

Recently the computer program "A Computer Solution for Hydrostatic Bearings with Variable Film Thickness," used to design the hydrostatic bearing of the 70-meter antennas, was modified to improve the accuracy with which the program predicts the film height profile and oil pressure distribution between the hydrostatic bearing pad and the runner. This article presents a description of the modified computer program, the theory upon which the computer program computations are based, computer simulation results, and a discussion of the computer simulation results.

I. Introduction

A significant level of engineering effort was expended in the early 1960s to develop a design for the hydrostatic bearing to support the alidade structure of the 64-meter DSN antennas. The hydrostatic bearing was designed using a computer program entitled "A Computer Solution for Hydrostatic Bearings with Variable Film Thickness," written by Dr. Vittorio Castelli while employed at the Franklin Institute. Until 1992 this program was used to estimate the oil film height that could be expected between the bearing pad and the runner under various operating conditions and did not consider pedestal deflections, pad deflections, or imperfections in the flatness of the runner when making oil film height predictions. Because the computer program did not consider the deflection and flatness issues, JPL engineering has not had complete confidence in the program's ability to accurately predict the oil film height of the hydrostatic bearing. In an effort to improve JPL's capability to accurately estimate the hydrostatic bearing oil film height, JPL contracted Dr. Castelli to modify the original computer program. The modified program allows

JPL to assess the effects of oil flow through the recesses, oil temperature, pedestal deflection, pad deflection, runner flatness, and pad loading. The additional program capabilities were determined to be necessary so that JPL can accurately assess the effects of adding additional weight to the antenna bearings as was done in 1987 when the 64-meter antennas were upgraded to 70-meter antennas. This article outlines the capabilities of the modified hydrostatic bearing program and presents the results of the program under various operating conditions of the hydrostatic bearing.

II. Hydrostatic Bearing Theoretical Model

The hydrostatic bearing of the DSN 70-meter antennas consists of a rectangular steel bearing pad connected to the antenna alidade frame by a self-aligning spherical joint. The pad rests on a relatively flat steel runner that is attached to the antenna pedestal. The bearing functions by pumping oil at a constant flow rate into each of six recesses located on the bottom of the pad so that the

antenna floats on a thin oil film as the oil flows out of the recess between the pad and the runner. The computer program developed by Dr. Castelli predicts the film height profile and pressure distribution between the bearing pad and the runner by finding, through an iterative process, the film height profile and pressure distribution that satisfy a set of governing equations. The governing equations are presented in this section.

A. Oil Film

The equation which governs the pressure distribution and oil film height between the pad and the runner is the Reynolds equation. The Reynolds equation is given by the expression¹

$$\frac{\delta}{\delta x} \left(p h^3 \frac{\delta p}{\delta x} \right) + \frac{\delta}{\delta y} \left(p h^3 \frac{\delta p}{\delta y} \right) = 0 \quad (1)$$

where p is the pressure, h is the film height, and x and y are Cartesian coordinates (see Appendix for glossary). The program assumes that oil is provided to each of six recesses at a constant flow rate and oil viscosity. The oil from the recesses flows in the clearance between the pad and the runner back to the reservoir. The behavior of the oil flow through the clearance between the bearing pad and the runner is governed by the equations [1]

$$q_x = -\frac{h^3 \delta y}{12\mu} \left(\frac{\delta p}{\delta x} \right) \quad (2)$$

and

$$q_y = -\frac{h^3 \delta x}{12\mu} \left(\frac{\delta p}{\delta y} \right) \quad (3)$$

where μ is the absolute viscosity of the oil, q_x is the flow rate in the x -direction at a point beneath the pad through a slot of height h and width δy , and q_y is the flow rate in the y -direction at a point beneath the pad through a slot of height h and width δx . The kinematic viscosity of the oil is determined for a particular oil using Walther's Equation given by²

$$\log(\log(\nu + a)) = m \log(T) + b \quad (4)$$

¹ Vittorio Castelli, *Report on the Development of a Computer Program for the Simulation of the Hydrostatic Bearing Problem of the 70 Meter Antenna Systems*, JPL D-10449 (internal document), Jet Propulsion Laboratory, Pasadena, California, p. 3, June 1992.

² *Ibid.*, p. 18.

where ν is the kinematic viscosity in centistokes, a is a constant equal to 0.6 centistokes, T is the absolute temperature, and m and b are constants of a particular oil. The constants m and b for an oil are determined by measuring the viscosity of the oil at two different temperatures and solving the two equations derived from Eq. (4) simultaneously for m and b . The absolute viscosity used in Eqs. (2) and (3) is found through the relationship [1]

$$\mu = \rho \nu \quad (5)$$

where ρ is the oil density. The computer program written by Dr. Castelli requires that the absolute viscosity be input at two different temperatures and assumes that the oil density is constant at 900 kg/m³ such that Eqs. (4) and (5) can be used to compute the absolute viscosity at any operating temperature.

The alidade structure is supported by the hydrostatic bearing in a state of static equilibrium; that is, the acceleration of the bearing pads is zero in the vertical direction. Hence the load supported by the bearing must be equal to the pressure forces under the pad. The total load supported by the pad is related to the pressure forces under the pad through the equation³

$$W = \int \int p \delta x \delta y \quad (6)$$

where the integrations are taken over the area on the bottom of the bearing pad. Since the bearing pad is mounted on a spherical grease-lubricated joint, the total moments acting on the pad as produced by the pressure forces must be zero. The moments acting on the pad about the x - and y -axes are computed using the equations⁴

$$W_x = \int \int p y \delta x \delta y \quad (7)$$

and

$$W_y = \int \int p x \delta x \delta y \quad (8)$$

respectively, where the integrations are taken over the area on the bottom of the bearing pad.

Equations (2) through (8) provide a set of constraint equations for the Reynolds equation which governs the

³ *Ibid.*, p. 3.

⁴ *Ibid.*, pp. 3-4.

pressure distribution and oil film height between the pad and the runner.

B. Deflection of the Pedestal and Pad

The deflections of the antenna pedestal and the bearing pad are computed by the modified hydrostatic bearing program using a numerical Green's function method. The use of the Green's function method to simulate the deflection of the pedestal and the pad is advantageous because it allows the deflection problem to be solved once using finite-element analysis for a given set of loading conditions. The solution for a given set of loading conditions can subsequently be scaled within the hydrostatic bearing program for other loading conditions with significantly less computation and without coupling a finite-element analysis to the hydrostatic bearing program. A Green's function gives the deflection of a structure at a point B on the structure resulting from a unit load applied at point A on the structure. Because Green's functions are based on linear elastic theory, the deflection of a point B due to a load K times the unit load applied at point A will be K times the deflection that exists with the unit load. In addition, the total deflection at point B will be the summation of the deflections at point B from the loads applied at all other points on the structure. The Green's functions for the various loading and deflection points were computed by Ed Solcz using the ALGOR finite-element analysis package for the DSS-14 pad and pedestal.⁵ Although the basic principles upon which a Green's function is based are the same when analyzing either the pedestal or pad deflection, the way in which a Green's function is implemented within the computer program is slightly different as described below.

1. Pedestal Deflection. The Green's functions describing the deflection of the pedestal have the form⁶

$$G(r_L, r_d, \theta_d, E, \nu) = \text{the deflection at point B as the result of a unit load at point A} \quad (9)$$

where E is the modulus of elasticity of the pedestal material, ν is Poisson's ratio, and r_L , r_d , and θ_d are geometrical dimensions as defined in Fig. 1. As a result of the symmetry of the antenna pedestal, the Green's functions were computed for unit loads applied to different radii on the bearing runner. The pedestal deflection at a coordinate (x, y) beneath the bearing pad is computed using the equation⁷

$$e(x, y) = \int \int p(\xi, \eta) G(r_L(\xi, \eta), r_d(x, y), \theta_d(x, y), E, \nu) \delta\xi \delta\eta \quad (10)$$

where the integrations are taken over the area on the bottom of the bearing pad.

2. Pad Deflection. The pad deflection is computed using the equation⁸

$$g(x, y) = P_{avg} d_u(x, y) + P_o d_o(x, y) + P_o^* d_o^*(x, y) \quad (11)$$

where

$$P_{avg} = \frac{1}{6} (P_1 + P_2 + P_3 + P_4 + P_5 + P_6) \quad (12)$$

$$P_o = \frac{1}{2} (P_1 + P_6 - P_2 - P_5) \quad (13)$$

$$P_o^* = \frac{1}{2} (P_3 + P_4 - P_2 - P_5) \quad (14)$$

and where $d_u(x, y)$ is the Green's function describing the deflection of the pad due to a uniform pressure distribution beneath the pad of unit magnitude, $d_o(x, y)$ is the Green's function describing the pad deflection when uniform pressure distributions of unit magnitude are present in areas 1 and 6 according to Fig. 2, $d_o^*(x, y)$ is the Green's function when uniform pressure distributions of unit magnitude are present in areas 3 and 4 according to Fig. 2, and P_i is the average pressure over area number i according to Fig. 2.

C. Oil Film Height Profile

Once the pedestal and pad deflections have been determined, the oil film height profile between the pad and the runner is computed using the equation⁹

$$h(x, y) = h(0, 0) + \phi_x y + \phi_y x + [e(x, y) - e(0, 0)] + [g(x, y) - g(0, 0)] + [b(x, y) - b(0, 0)] \quad (15)$$

where

⁵ Ibid., pp. 65-87.

⁶ Ibid., p. 8.

⁷ Ibid., pp. 8-9.

⁸ Ibid., pp. 9-10.

⁹ Ibid., p. 4.

- $h(x, y)$ = oil film height
 $e(x, y)$ = absolute deflection distribution of the pedestal positive up
 $g(x, y)$ = deflection distribution of the pad positive up
 $b(x, y)$ = distribution of the bump on the runner positive up
 ϕ_x = tilt angle of the pad about the x -axis in radians
 ϕ_y = tilt angle of the pad about the y -axis in radians

The computer program arrives at the final film height profile through an iterative process that converges on the final solution where all of the governing equations presented in Eqs. (1) through (15) are simultaneously satisfied.

D. Pump Flow Curves

One of the limitations of the modified hydrostatic bearing program developed by Dr. Castelli is that the program assumes that oil is provided to the bearing recess at a constant specified flow rate. However, in practice, this is not the case because the De Laval Turbine, Inc., IMO B12-LBSX-118 pumps which are used to supply oil flow to the bearing recesses actually provide a flow rate according to the equation

$$Q = 0.4S - 466.823 \sqrt{\frac{\Delta p}{\nu}} \text{ ml/sec} \quad (16)$$

where S is the pump speed in rpm, Δp is the differential pressure across the pump in MPa, and ν is the kinematic viscosity in centistokes. Figure 3 shows how the output flow of the pump varies with oil kinematic viscosity and differential pressure across the pump using Eq. (16). The curves shown in Fig. 3 should distort downward as a pump wears because the coefficient 466.823 is expected to increase in magnitude as the internal pump clearances become larger.

III. Computer Program

A. Description

The original computer program used to analyze the hydrostatic bearing was modified by Dr. Castelli to expand the program's capabilities. The modified program allows the user to input a wide variety of parameters related to

the operation of the hydrostatic bearing. This provides JPL with a great deal of flexibility in simulating the hydrostatic bearing under many operating conditions. The computer program allows the program user to specify the geometry of the bearing pad, input a constant oil flow rate to each recess, input the viscosity-to-temperature relationship of the oil, specify the operating temperature of the oil, input the load acting on the pad, input the resultant moments acting on the pad, input the Green's functions for the pedestal and pad deflections, input the modulus of elasticity of the pedestal and pad material, and input a runner profile to simulate imperfections in the flatness of the runner surface. Once provided with all of the required inputs, the computer program computes the film height profile, pressure distribution, recess pressures, and minimum film height through an iterative process.

B. Computer Simulation

Several computer simulation runs were made in order to determine the sensitivity of the hydrostatic bearing to various parameters. A Lahey Compiler was used to process all of the Fortran 77 source code, and the program was run using a 386 microprocessor. An average of 30 min of computation time was required to provide results. This section presents the results of the hydrostatic bearing analysis performed using the modified hydrostatic bearing computer program.

1. Simulation Conditions. Due to the large number of input parameters which influence film height, it would not be feasible to run the hydrostatic bearing program for every possible operating condition. To keep the amount of computer output and computation time to a minimum while still obtaining useful information, a small number of computer runs were made to determine how sensitive the hydrostatic bearing performance is to changes in the input parameter values. This was accomplished by running the simulation program using parameter values that are higher and lower than the value believed to currently exist under normal operation at DSS 14. The largest pad load at DSS 14 is believed to be approximately 1.3×10^7 newtons.¹⁰ The geometry of the bearing pad used to support the DSN 70-meter antennas is shown in Fig. 4. ISO 150 oil was supplied to each of the six recesses at a temperature of 308 K. The flow rate was assumed to be a constant flow rate of 550 ml/sec to each recess. This flow rate was selected for simulation purposes because the value is conservative since the flow is not expected to be less

¹⁰ B. Saldua, *Baseline of DSS-14 Hydrostatic Pad Loads*, JPL Interoffice Memorandum 3323-92-001 (internal document), Jet Propulsion Laboratory, Pasadena, California, January 2, 1992.

than this with an unworn pump based upon expected recess pressures, oil temperature estimates, and the pump performance curves shown in Fig. 3. Unless specified in the results, all computer runs were made using the runner profile shown in Fig. 5. The data presented in Fig. 5 are based upon actual sensor data collected on October 16, 1992, and the profile calculation was made according to the currently authorized DSN method.¹¹ The locations of the sensors used to obtain the runner profile are shown in Fig. 4.

2. Oil Temperature. The temperature of the oil supplied to the recesses of the hydrostatic bearing has a major effect on the performance of the bearing. The performance of the bearing is affected by oil temperature because the oil properties change with temperature. Figures 6, 7, and 8 show how the oil properties of absolute viscosity, kinematic viscosity, and density vary with oil temperature for ISO 150 oil and ISO 220 oil. A graph of the minimum film height as a function of oil temperature for ISO 150 oil and ISO 220 oil is shown in Fig. 9. Figures 10 and 11 show the recess pressures as a function of oil temperature for ISO 150 oil and ISO 220 oil, respectively. Currently all of the 70-meter antennas operate the hydrostatic bearing using ISO 150 oil. The behavior curves for ISO 220 oil have also been included because some individuals have suggested improving film height by using a more viscous oil. A complete discussion of the benefits and problems of using a more viscous oil in the hydrostatic bearing is provided later in this article.

3. Recess Flow. The oil flow rate provided to a particular recess is dependent upon the oil viscosity and differential pump pressure as depicted in Fig. 3. The flow rate provided by a pump is also related to the degree of pump wear. Therefore, the sensitivity of the bearing performance to oil flow rate through the recess is of significant interest. Figure 12 shows the film height as a function of oil flow through the recess assuming all recesses receive the same flow. The recess pressures are shown as a function of oil flow rate in Fig. 13. Because the pumps operate at different pressures and are worn to different degrees, it is unlikely that every pump provides exactly the same flow rate. To determine what effect this would have on the performance of the hydrostatic bearing, several runs were performed while providing different oil flow rates to each of the recesses. The results of these runs are presented in Table 1.

¹¹ W. Kuehn, *Software Definition Document for Hydrostatic Bearing Instrumentation*, SDD-DOA-5534-SP (internal document), Jet Propulsion Laboratory, Pasadena, California, pp. 10-12, March 8, 1984.

4. Load. Modifications to the 70-meter antennas are frequently made in order to improve their capabilities. These modifications generally require additional weight to be added to the alidade structure, which increases the load that must be supported by the hydrostatic bearing. Therefore there is a great deal of interest at JPL in determining the performance effects of altering the load supported by the hydrostatic bearing. Figures 14 and 15 show the minimum film height and recess pressures, respectively, as a function of pad load.

5. Pedestal and Pad Deflections. The hydrostatic bearing program originally developed in the early 1960s did not take into consideration the deflection of the pedestal and the pad when predicting the film height and pressure distributions. Hence, the effect of deflection on hydrostatic bearing performance has been unknown and of great concern to JPL for many years. In order to identify the effects of pedestal and pad deflections, four computer runs were made while assuming a flat runner profile. The four computer runs consisted of one run where the pedestal and the pad were infinitely stiff, one run where the pad deflected and the pedestal was infinitely stiff, one run where the pad was infinitely stiff and the pedestal deflected, and one run where both the pedestal and the pad deflected. The minimum film height and recess pressures for each of these runs are presented in Table 2; graphs of the film height profiles for these cases are shown in Figs. 16 through 23.

6. Runner Profile. DSN operations have expended a large degree of effort to shim the runner to a flat surface in order to avoid low film height alarms. As a result of the large quantity of resources used to perform the shimming operations, JPL is interested in determining the effects that imperfections in runner flatness have upon the hydrostatic bearing performance. Table 3 shows the minimum clearance and recess pressure that arise when the hydrostatic bearing is operated with a perfectly flat runner and when it is operated with the DSS-14 runner profile shown in Fig. 5. The graphs of the clearance profiles with a flat runner surface and with the DSS-14 runner profile are shown in Figs. 22 through 25.

IV. Discussion of Computer Program Results

This section presents a discussion of the results gathered from the computer analysis.

A. Oil Temperature

An examination of the relationship between oil temperature and hydrostatic bearing film height shown in Fig. 9

reveals that the oil temperature and the type of oil used in the hydrostatic bearing have a significant effect on the minimum film height present between the bearing pad and the runner. Currently the hydrostatic bearing instrumentation group provides a low film height warning when the clearance between the pad and the runner is less than 127 μm (0.005 in.) and sets the antenna brake when the clearance is less than 76.2 μm (0.003 in.) while the antenna is moving. DSS 14 currently operates the hydrostatic bearing using ISO 150 oil at a temperature of approximately 308 K (94.73 deg F). At this oil temperature and under DSS-14 operating conditions the minimum film height is expected to be approximately 151.6 μm (0.00597 in.). Hence DSS 14 has approximately a 24.6- μm (0.00097-in.) margin of safety against a warning. This is not very much when one considers the uncertainty in the operating conditions of the bearing. It is clear from Fig. 9 that the minimum film height could be significantly increased by operating the hydrostatic bearing at a lower temperature or using ISO 220 oil instead of ISO 150 oil in the bearing. In practice there may be some difficulty in lowering the oil temperature to a desired level to obtain better film heights because lowering the oil temperature decreases the differential temperature between the bearing oil and the cooling water. This, in turn, decreases the rate at which heat can be removed from the oil through the heat exchanger. If the heat exchanger cannot remove the heat generated in the system at a desired oil temperature, the heat exchanger cannot supply oil at that desired temperature. To obtain better film heights without having this difficulty, ISO 220 oil could be used instead of ISO 150 oil. However, this method of improving oil film height may create problems also because the temperature at which the pumps have enough horsepower to perform a cold start during the winter months is increased and because the seals in the high pressure pumps can be damaged during a cold start if the viscosity of the oil passing through the pump is greater than 630 centistokes. Modifications to the existing system may be required to implement either of these methods of improving hydrostatic bearing film height. The identification of any required modifications requires additional study.

Figures 10 and 11 show that the recess pressures do not change significantly as the temperature of the oil changes.

B. Recess Flow

The results of the computer analysis performed to determine the sensitivity of the minimum bearing film height to the recess flow rate shown in Fig. 12 indicate that the flow rate has a moderate effect on the oil film height. The effect of recess flow is considered moderate because relatively large changes in flow rate are required to affect the

minimum film height by an amount that is considered significant. Although the effect of oil flow is moderate, it should not be ignored, because large variations can occur since pump flow rates can vary by large amounts as the result of variations in oil viscosity and differential pressure across the high pressure pumps, as indicated in Fig. 5. Pump wear will also cause variations in flow rate because clearances are increased within the pump. The extent to which pump wear affects the output flow rate of the high pressure pumps is not known because instrumentation to measure the output flow rate of the pumps while under load does not currently exist.

Based upon Fig. 12, it would be logical to consider improving the bearing film height by increasing the flow rate provided to each recess. However, this is not a cost-effective method of improving film height for the following reasons:

- (1) All 18 high pressure pumps and the electric motors which drive the pumps would have to be replaced in order to provide the increased flow rate.
- (2) Increased electric power consumption would be required to drive the pumps since the output pressure of the pumps would not change with an increase in flow, as indicated by Fig. 13.
- (3) The precharge pumping system would have to be completely redesigned to accommodate the higher flow rates and to provide an oil cooling system capable of dissipating the additional thermal energy generated by the system.

C. Load

Occasionally it is desirable to add additional equipment to the antenna in order to improve its performance. The addition of equipment adds to the weight which must be supported by the hydrostatic bearing. Figure 14 shows the extent to which the minimum film height decreases with an increase in the load supported by the bearing. This graph can be used to estimate the decrease in film height resulting from any future additional weight to the antenna in order to determine if the additional weight could result in low film height alarms. An increase in bearing pad load also results in an increase in recess pressures, as shown in Fig. 15. The pumps used on the 70-meter antenna are capable of operating at pressures up to 27.58 MPa. When a proposal is made to add weight to the antenna, an estimate of the high pressure pump output pressure should be made using Fig. 15 to determine if the maximum pump operating pressure is exceeded and if the electric motors are capable of providing the additional power to rotate the pumps. Since additional weight results in an increase in

pump pressure, an increase in antenna weight also results in an increase in the thermal energy generated by the system. The cooling system must be checked to ensure that it is capable of dissipating the additional thermal energy from the oil at the desired operating temperature or low film height alarms may occur due to the inability of the system to maintain the correct oil temperature. The increase in pump pressure resulting from increased pad load also results in a reduced output flow from the pumps as indicated in Fig. 3. The effects of the reduction in pump flow are expected to be moderate for the reasons stated in Section IV.B but should be considered when planning increases in pad loading.

It is important to realize that Fig. 15 can be used to make rough estimates of antenna pad loads. Pad load estimates can be made by measuring the pump output pressures at the stations and using Fig. 15 to correlate the pump pressures to the pad load.

D. Pedestal and Pad Deflection

The design of the hydrostatic bearing in the 1960s was based upon the original computer program "A Computer Solution for Hydrostatic Bearings with Variable Film Thickness." This program did not consider pad or pedestal deflections when making film height profile or pressure distribution predictions. The designers of the hydrostatic bearing attempted to remove the effect of pedestal and pad deflections by trying to design the pedestal and pad such that the deflections of the pedestal and pad would be matched. By matching the pedestal and pad deflections the bearing designers hoped to obtain a uniform film height beneath the pad as would be the case if the pedestal and pad did not deflect at all.¹² The modified computer program assesses the effects of both pedestal and pad deflections. An examination of the computer simulation results from the modified program shown in Table 2 and Figs. 16 through 23 reveal that the deflections of the pedestal and the pad cause the minimum film height to be significantly lower than was expected when the hydrostatic bearing was designed without consideration of the pad and pedestal deflections. Table 2 and Figs. 16, 17, 22, and 23 show that the expected minimum film height decreases by 114.6 μm (0.0045 in.) when both pedestal and pad deflections are considered compared to when no deflections are considered as was the case when the hydrostatic bearing was designed. Hence, the deflection of both the pedestal and the pad are

¹² *The NASA/JPL 64-Meter-Diameter Antenna at Goldstone, California: Project Report, Technical Memorandum 33-671 (internal document), Jet Propulsion Laboratory, Pasadena, California, pp. 63-65, July 15, 1974.*

believed to contribute significantly to the low film height problems experienced by the 70-meter antennas.

E. Runner Profile

Operations personnel have expended a great deal of effort over the years in order to shim beneath the hydrostatic bearing runner in order to obtain a "flat" runner so that low film height alarms do not occur. The shimming is necessary to keep the runner flat because corrosion of the runner, corrosion of the sole plates, and degradation of the grout cause the runner profile to change over time. To assess the effects of a nonflat runner, a comparison of the minimum film height and recess pressures was made between a computer simulation performed with a flat runner and a simulation performed with the DSS-14 runner profile shown in Fig. 5. The results of the comparison provided in Table 3 show that there is a 7.08- μm (0.0003-in.) decrease in minimum film height when the DSS-14 profile is used. This decrease in minimum film height due to the imperfect DSS-14 runner is not considered to be significant. The fact that the DSS-14 runner profile did not produce a significant degradation in film height does not indicate that the runner profile does not have a significant effect on bearing performance but rather that the DSS-14 profile used for the computer simulations does not have a significant effect. In practice operations personnel have found that the runner profile has a large effect on film height. Further investigation is required to identify the kinds of runner profiles that cause film height problems.

F. Recess Pressure Variations

Field measurements of the recess pressures at DSS 14 show a variation among the pressures of different recesses. Measurements taken from pad 3 on December 10, 1991, revealed the pressures of recesses 1 through 6 as defined in Fig. 4 to be 9.65 MPa, 10.34 MPa, 7.58 MPa, 8.62 MPa, 9.65 MPa, and 11.38 MPa, respectively.¹³ The difference between the highest and lowest pressure is 3.8 MPa. This variation in recess pressures is much larger than the variations of the recess pressures predicted by the computer simulations as shown in Figs. 10, 11, 13, and 15, and Tables 1 through 3. An examination of Table 2 shows that some variations among recess pressures result from the deflection of the pad and runner. Table 3 shows that variations in recess pressures can be caused by the runner profile. Although the deflections of the pad and pedestal and the runner profile result in some variation of the recess pressures, neither produced a variation as high as the one observed in practice. Several computer simulations were performed with different flow rates to the recesses to see if

¹³ B. Saluda, *op. cit.*

the large variations in recess pressures could be attributed to variations in pump flow rate. The results of these simulations presented in Table 1 indicate that only small variations in recess pressures can occur from differences in the flow rates provided to each recess. Further investigation is required to determine the reasons for the variations among the recess pressures observed on the 70-meter antennas.

G. Measurement of the Runner Profile and Bearing Film Heights

The effects of the runner profile on bearing performance were assessed using a runner profile computed according to the method currently approved by the DSN¹⁴ from measurements taken at DSS 14 with the hydrostatic bearing instrumentation group. The results of the computer simulations presented in this article and discussions with station personnel at DSS 14 bring in to question the accuracy to which this runner profile is known with respect to a plane perpendicular to the gravity vector. If the runner profile cannot be determined accurately then it is difficult to try to assess the actual effects of the runner profile on the bearing film height.

The runner profile is currently measured using the hydrostatic bearing instrumentation group. This instrumentation group consists of five linear variable displacement transducer (LVDT) sensors mounted to a bar which is attached to one edge of bearing pad 3 as shown in Fig. 4. Two tipping probes, probes 6 and 7, are mounted over the pad to measure the motion of the top of the pad relative to the alidade frame. Probe 6 measures the vertical motion of the pad directly over probe 1 while probe 7 measures the vertical motion of the pad directly over probe 5. The vertical motion of the alidade frame relative to the azimuth bull gear is measured using LVDT probe number 9, mounted to the alidade frame with the tip of the probe riding on the top of the bull gear. The profile of the bull gear is determined by using a 10-degree level bar which determines the difference in elevation with respect to gravity of points at 10-degree intervals along the bull gear as shown in Fig. 26. The differences are added together to obtain the profile of the bull gear with respect to any measurement point defined to be zero on the bull gear. The profile of the bull gear between the 10-degree measurement points is interpolated using a 3-1/3-degree level bar mounted to the alidade frame that has three LVDT probes, probes 9, 15, and 16, mounted on it 3-1/3 degrees apart with respect to the antenna center. The bull gear probe previously mentioned is the center probe on this level bar. A diagram showing the

location of all the probes is shown in Figs. 27 and 28. It is clear from Fig. 27 that

$$Z + G + Y = X + T + h \quad (17)$$

Differentiating Eq. (17) yields

$$dZ + dG + dY = dX + dT + dh \quad (18)$$

Rearranging Eq. (18) produces an equation for the differential change in runner profile given by

$$dZ = dX + dT + dh - dG - dY \quad (19)$$

where dh is the change in film height as measured by probes 1 through 5, dT is the change in distance between probes 1 through 5 and the line extended between the pad contact points of probes 6 and 7, dX is the change in elevation between the top of the bearing pad and the alidade frame, dY is the change in elevation between the alidade frame and the bull gear measured using probe 9, and dG is the change in the bull gear profile as measured by the 10-degree level bar and interpolated using probes 9, 15, and 16. The value of dT is assumed to be zero by the current measurement scheme. The profile of the runner is determined by computing the change in runner height from the location on the runner that is defined as having zero runner height with respect to a plane perpendicular to gravity. The equation used to compute the runner profile under probes 1 through 5 at a particular azimuth angle is¹⁵

$$Z_i = P_9 + G - P_i - \frac{P_6 + P_7}{2} - \left(\frac{P_6 - P_7}{2} \right) a_i \quad (20)$$

$$a_1 = 1.0 \quad (21)$$

$$a_2 = 0.5 \quad (22)$$

$$a_3 = 0 \quad (23)$$

$$a_4 = -0.5 \quad (24)$$

$$a_5 = -1.0 \quad (25)$$

where P_i is the measurement of probe i and G is the interpolated bull gear profile.

¹⁴ W. Kuehn, op. cit., pp. 10-12.

¹⁵ W. Kuehn, op. cit., p. 12.

As a result of the bearing pad deflection, inaccuracies exist with either the measurement of the bearing film height and/or the measurement of the runner profile. The uncertainty exists in the film height and runner profile measurements because the deflection of the bar to which probes 1 through 5 are attached is unknown. If one assumes that the probe bar is attached in such a way that its deflection curve is exactly that of the pad as shown in Fig. 28, then the measurements taken by probes 1 through 5 represent the true film height. However, this results in inaccuracies because the distance between probes 1 through 5 and the line extended between the pad contact points of probes 6 and 7 is not the same for each probe as is assumed when computing the runner profile using Eq. (20). If one assumes that the probe bar is attached in such a way that it does not deflect at all as shown in Fig. 29, then the distance between probes 1 through 5 and the line extended between the pad contact points of probes 6 and 7 are the same for each probe, but probes 1 through 5 do not all measure the true film height. If the probe bar deflects differently than the bearing pad then there are inaccuracies in both the film height measurements and the runner profile measurements.

During construction, the hydrostatic bearing runner was installed such that the surface of the runner was flat in an unloaded state as intended by the bearing design. Because probes 1 through 5 are at the edge of the pad, they measure the profile of the runner in its deflected state. Hence, the current means of measuring the runner profile does not establish if the runner is flat in an unloaded state.

Station personnel have claimed that antenna shut downs have occurred in the past because oil contaminants have become stuck beneath the sensor probes. Since station personnel have found contaminant particles beneath the probes in the past it is clear that contaminant particles present in the hydrostatic bearing reservoir are another source of uncertainty when measuring the bearing film height and runner profile.

As the antenna rotates there is metal-to-metal rubbing contact between the tips of the probes measuring film height and the runner. Over time the tips of the probes and the runner surface will wear, resulting in uncertain measurement of the bearing film height and the runner profile.

The accurate determination of the runner profile is very dependent on knowing the profile of the bull with respect to an arbitrary plane perpendicular to gravity as evident from Eq. (20). As previously stated the bull gear profile is determined using a bar leveled with respect to gravity to

measure the differences in bull gear elevation at 10-degree intervals. An electronic level is used to level the bar and an LVDT transducer is used to measure the difference in bull gear elevations at 10-degree intervals. There is some uncertainty in the measurement which is unknown. Because the bull gear profile at a point on the bull gear is determined by adding the elevation differences between a reference point *A* and point *B* where the bull gear profile height is desired, the error in the measurement of the bull gear profile is the sum of the measurement errors of all 10-degree intervals between point *A* and point *B*. Because there are 36 10-degree intervals, the uncertainty in the bull gear profile increases rapidly as one goes around the bull gear. Evidence of this is provided by station personnel at DSS 14 who claim that it is difficult to sum the elevation differences between all 36 intervals and obtain a sum within 254 μm (0.01 in.) of zero. Station personnel at DSS 14 have also indicated that it is impossible to obtain repeatability of any bull gear profile measurements if the antenna brakes are set during the measurement process. This indicates that the bull gear is not a constant reference point as it is currently treated for runner profile measurement and shimming purposes. Visible cracking of the antenna pedestal around the bull gear sole plates at DSS 14 provides further evidence supporting the claim that the bull gear profile is not constant. The movement of the bull gear at DSS 14 is expected to be a major source of runner profile measurement error. As of December 1992, a bull gear survey had not been performed at DSS 14 since February 1991, and since 1986, the data collected from a bull gear survey had not been entered into the computer program used to calculate the runner profile.

Inspection of the bull gear surface upon which measurements are taken by probe 9, which is used to reference the alidade frame to the bull gear, reveals that the top of the bull gear is a greased surface with many deep scratches and exposed to the atmosphere. Because the surface is exposed to the atmosphere, there is no guarantee that contaminants present on the bull gear do not affect the accuracy of the measurement of the position of the alidade frame with respect to the bull gear as measured by probe 9.

As the result of the numerous sources of measurement and computational errors presented above, there is uncertainty in just how accurately the hydrostatic bearing instrumentation group is able to measure bearing film height and runner profile. The inability to accurately measure the runner profile and film height can only create problems for the ongoing effort to maintain the 70-meter antenna runner flat in an unloaded state.

V. Conclusion

The program "A Computer Solution for Hydrostatic Bearings with Variable Film Thickness" originally used to design the 64-meter antenna hydrostatic bearing was modified in order to improve the ability of the program to accurately predict the film height profile between the hydrostatic bearing pad and the runner. The modified computer program is an improvement over the original because it considers pedestal deflections, pad deflections, and imperfections in the runner flatness when making predictions of the film height profile and oil pressure distribution. The results of several computer simulation runs were presented to show the sensitivity of the minimum bearing film height and recess pressures to oil temperature, oil type, recess flow rate, bearing load, pedestal deflection, pad deflection, and runner profile. This information can be used to assess the effects on bearing performance of any proposed additional pad loads or modifications to the hydrostatic bearing system. The computer analysis indicated that the minimum film height is significantly less than predicted during the 1960s' design analysis because pedestal

and pad deflections were not considered during the original analysis. The analysis also revealed that a 6-K (10.8-deg-F) decrease in oil temperature from the normal operating temperature results in a 37.5- μm (0.0014-in.) increase in minimum film height, a 15.38-percent increase in pad load results in a 29.8- μm (0.0012-in.) decrease in minimum film height, and an 18.18-percent decrease in recess flow results in a 12.9- μm (0.0005-in.) decrease in minimum film height. The effect on bearing performance of the runner profile at DSS 14 as measured by the hydrostatic bearing instrumentation group was presented. The current value of using the modified hydrostatic bearing computer program to evaluate the effect of actual runner profiles, such as the one from DSS 14 used in this study, is questionable because the uncertainty of the runner profile and film height measurements made by the hydrostatic bearing instrumentation group is not well defined for the reasons stated in this article. The accuracy of the measurements made by the hydrostatic bearing instrumentation group should be well defined since these are the measurements upon which the hydrostatic bearing runner is shimmed and low film height alarms are made.

Acknowledgments

The author thanks Bill Almassy, Jim Bunce, John Cucchissi, Steve Paine, and Chris Yung for their useful discussions and assistance.

Reference

- [1] J. Halling, *Principles of Tribology*, London, Great Britain: Macmillan Education LTD, 1978.

Appendix

Glossary

a	constant
a_i	constant
b	constant
$b(x, y)$	runner profile under pad coordinates (x, y) positive up
$d_o(x, y)$	Green's function at pad coordinates (x, y)
$d_o^*(x, y)$	Green's function at pad coordinates (x, y)
$d_u(x, y)$	Green's function at pad coordinates (x, y)
$e(x, y)$	pedestal deflection positive up at coordinates (x, y)
E	modulus of elasticity
$g(x, y)$	pad deflection positive up at pad coordinates (x, y)
G	bull gear profile with respect to a reference plane perpendicular to gravity
h	film height
h_{\min}	minimum film height
$h(x, y)$	film height at pad coordinates (x, y)
m	constant
p	pressure
Δp	differential pressure across the pump
P_{avg}	average pressure
P_i	pressure in recess number i , reading of probe number i , average pressure of area number i
P_o	pressure
P_o^*	pressure
q_x	oil flow rate at a point under the pad in the x direction through an area of height h and width dy
q_y	oil flow rate at a point under the pad in the y direction through an area of height h and width dx
Q	pump flow rate
Q_i	oil flow into recess number i
r_d	radius where the deflection is desired
r_L	radius at which a load is applied
S	pump rotational speed in rpm
T	absolute temperature, vertical distance between probes 1 through 5 and the line extended between the pad contact point of probes 6 and 7
W	load on the bearing pad
W_x	moment on the bearing pad about the x -axis
W_y	moment on the bearing pad about the y -axis
x	Cartesian coordinate

x	Cartesian coordinate
X	vertical distance between the top of the bearing pad and a reference point on the alidade frame
y	Cartesian coordinate
Y	vertical distance between the top of the bull gear and a reference point on the alidade frame
Z	runner profile from the reference plane
Z_i	runner profile from the reference plane as measured by probe number i
η	dummy variable
θ_d	angle between the applied load and the point where the deflection is desired
μ	absolute viscosity
ν	kinematic viscosity, Poisson's ratio
ξ	dummy variable
ρ	oil density
ϕ_x	pad tilt angle about the x -axis
ϕ_y	pad tilt angle about the y -axis

Table 1. Recess pressures and minimum film height when different oil flow rates are provided to each recess. Each computer simulation was performed using ISO 150 oil at a temperature of 308 K and a pad load of 1.3×10^7 newtons. Pedestal and pad deflections were considered with the pad centered at 135 deg trough azimuth over the runner profile shown in Fig. 5.

Pump flow rate, ml/sec						Recess pressure, MPa						h_{\min} , μm
Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	P_1	P_2	P_3	P_4	P_5	P_6	
550	550	550	550	550	550	9.88	10.06	9.96	10.24	10.03	10.33	151.6
350	550	550	550	550	550	9.23	10.38	9.89	10.37	10.21	10.00	132.6
550	350	550	550	550	550	10.06	9.75	10.16	10.28	10.02	10.37	143.5
650	550	550	550	550	550	10.03	10.01	9.96	10.18	10.03	10.37	153.0
550	650	550	550	550	550	9.63	10.40	9.72	10.09	10.22	10.17	154.9
350	550	650	350	550	650	9.53	10.33	9.79	9.88	10.32	10.17	119.1

Table 2. Effect of pedestal and pad deflections on recess pressures and minimum film height. Each computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

Pedestal deflection	Pad deflection	Recess pressure, MPa						h_{\min} , μm
		P_1	P_2	P_3	P_4	P_5	P_6	
No	No	9.96	11.49	9.96	9.96	11.49	9.96	273.4
No	Yes	7.13	15.02	7.12	7.13	15.03	7.11	199.3
Yes	No	10.65	9.27	11.00	11.01	9.42	11.13	102.68
Yes	Yes	10.14	10.03	10.23	10.29	10.01	10.38	158.72

Table 3. A comparison of the recess pressures and minimum film height when a flat runner is used and when the runner profile shown in Fig. 5 is used with the pad centered at 135 deg trough azimuth. Each computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, and recess flows of 550 ml/sec. Pedestal and pad deflections were considered.

Flat runner	Recess pressure, MPa						h_{\min} , μm
	P_1	P_2	P_3	P_4	P_5	P_6	
Yes	10.14	10.03	10.23	10.29	10.01	10.38	158.72
No	9.88	10.06	9.96	10.24	10.03	10.33	151.6

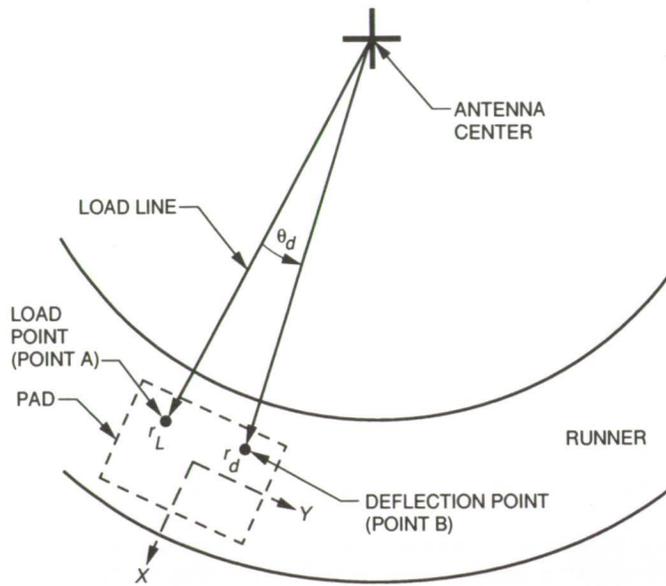


Fig. 1. Geometry and coordinate system of the antenna pedestal and pad as viewed from the top.

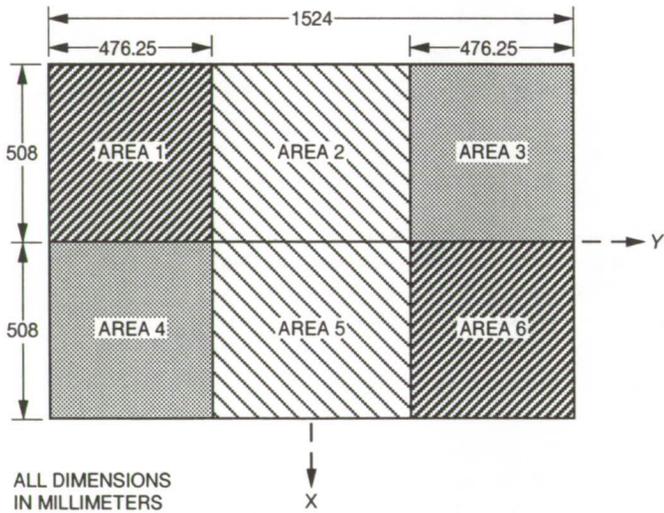


Fig. 2. Component pressure distributions used to compute the deflection of the hydrostatic bearing pad.

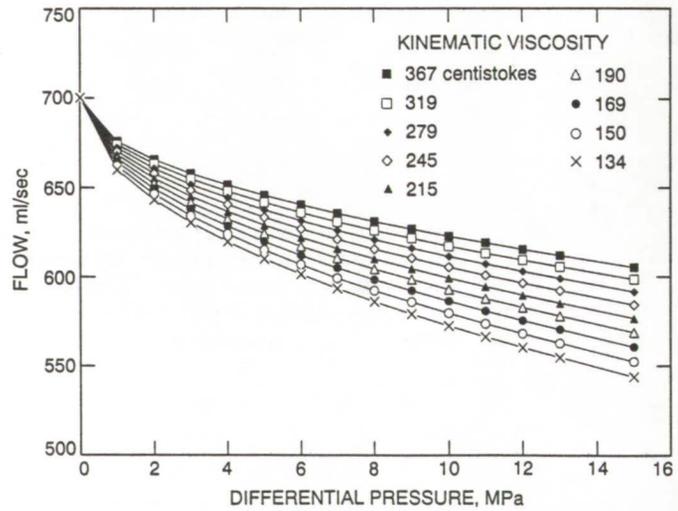


Fig. 3. Output flow rate of an unworn De Laval Turbine, Inc., IMO B12-LBSX-118 pump operating at 1750 rev/min as a function of the differential pressure across the pump and the kinematic viscosity of the oil.

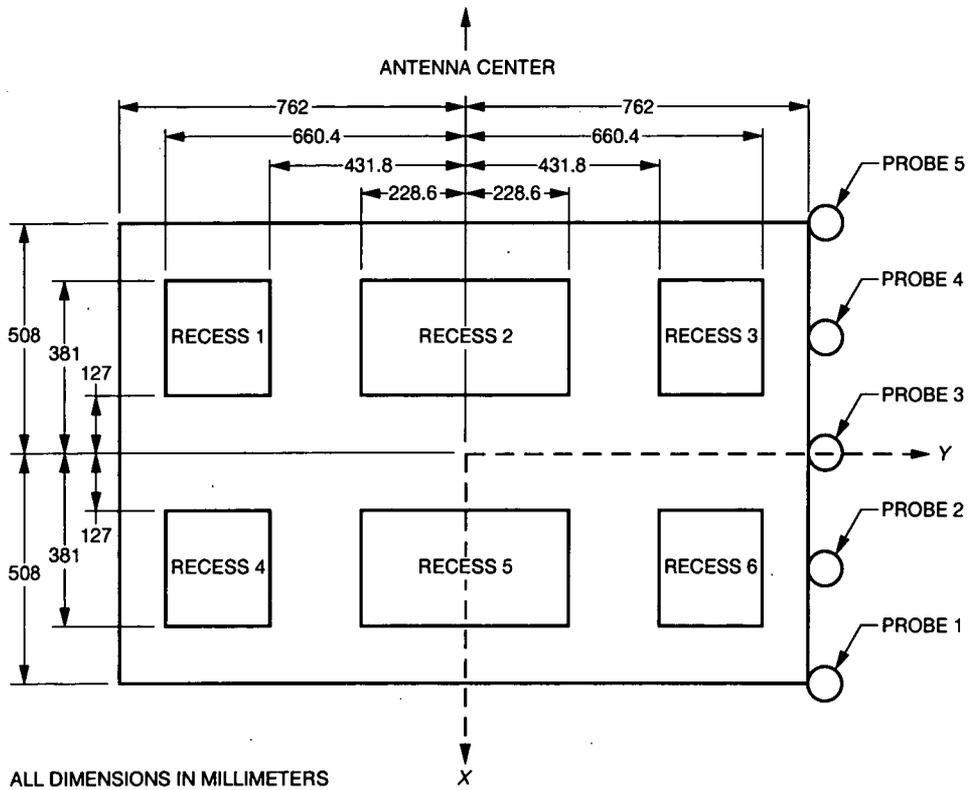


Fig. 4. Geometry and coordinate system of the DSN 70-meter antenna hydrostatic bearing pad number 3 as viewed from the top.

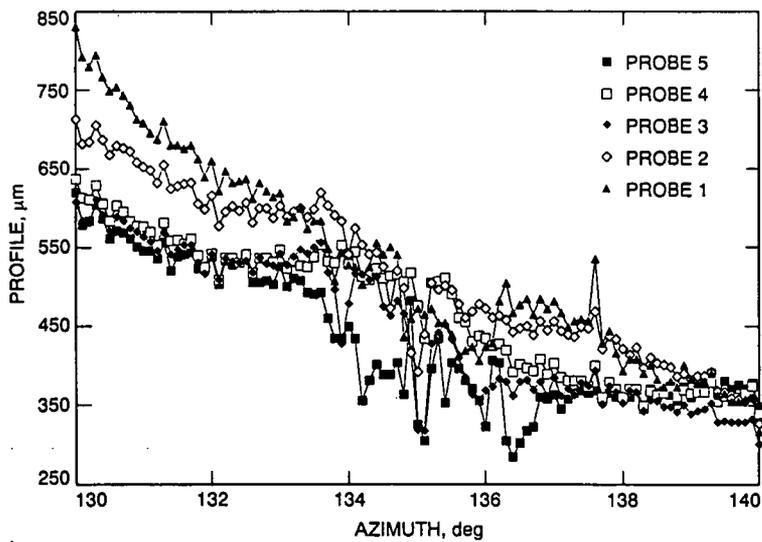


Fig. 5. DSS-14 hydrostatic bearing runner profile data collected on October 16, 1992, between 130 deg and 140 deg trough azimuth.

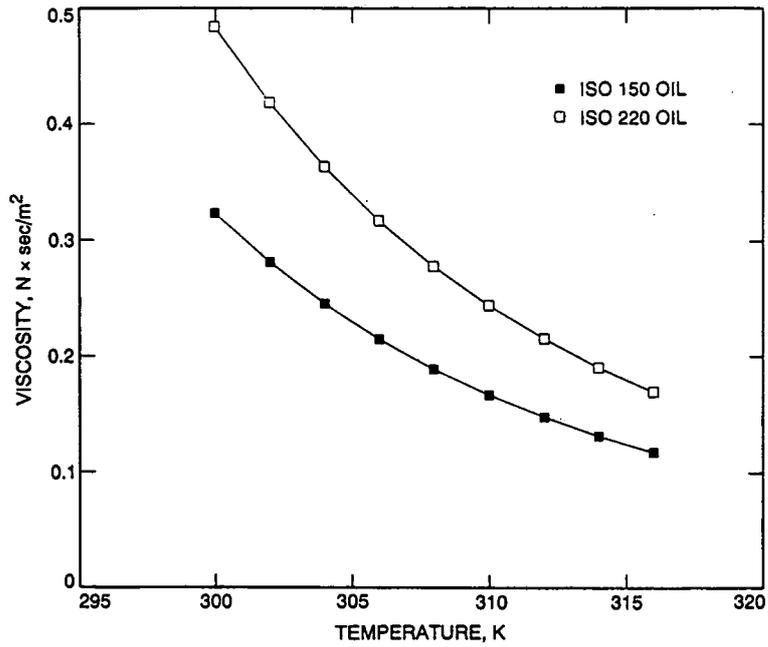


Fig. 6. Absolute viscosity of ISO 150 oil and ISO 220 oil as a function of temperature.

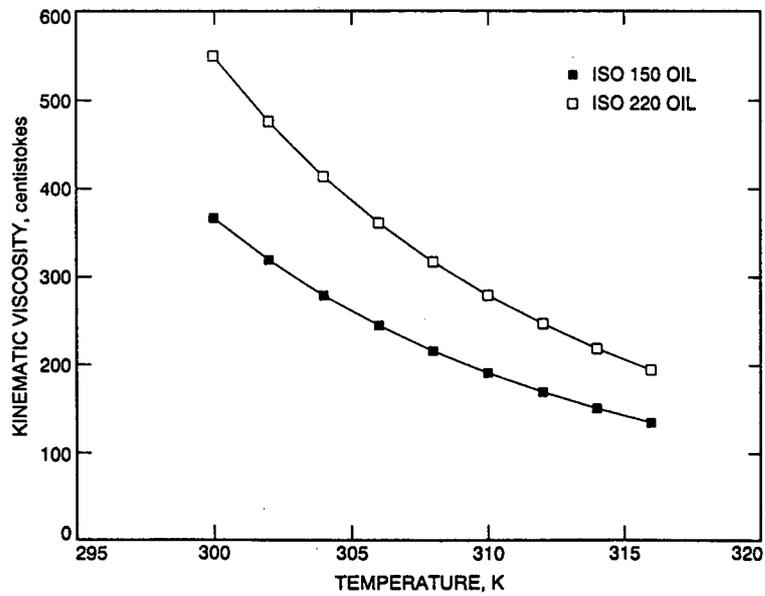


Fig. 7. Kinematic viscosity of ISO 150 oil and ISO 220 oil as a function of oil temperature.

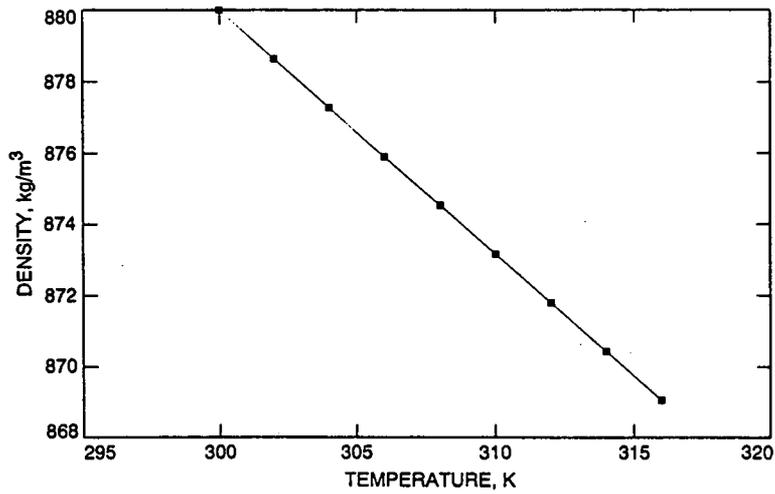


Fig. 8. Density of ISO 150 and ISO 220 oil as a function of oil temperature.

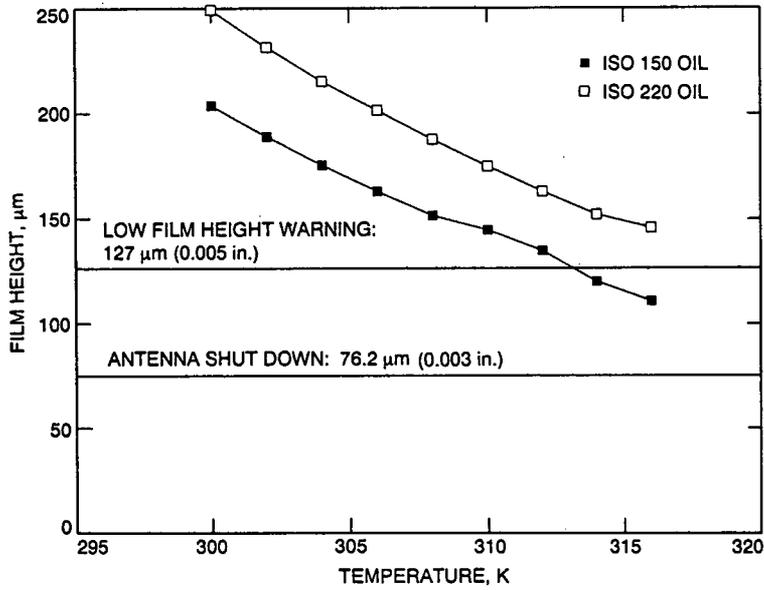


Fig. 9. Minimum film height as a function of temperature for ISO 150 oil and ISO 220 oil. All of the hydrostatic bearing simulations used to produce this graph were performed using a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and the runner profile shown in Fig. 5 with the pad centered at 135 deg trough azimuth. Pedestal and pad deflections were considered during each simulation.

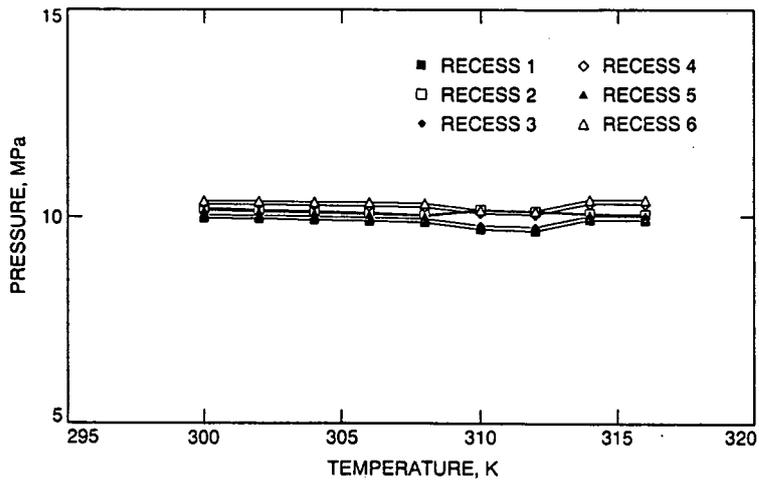


Fig. 10. Recess pressures as a function of temperature when using ISO 150 oil. All of the hydrostatic bearing simulations used to produce this graph were performed using a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and the runner profile shown in Fig. 5 with the pad centered at 135 degrees trough azimuth. Pedestal and pad deflections were considered during each simulation.

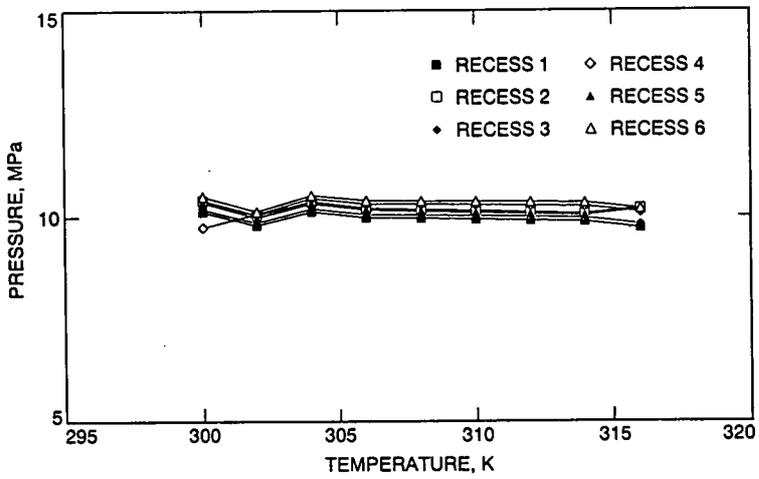


Fig. 11. Recess pressures as a function of temperature when using ISO 220 oil. All of the hydrostatic bearing simulations used to produce this graph were performed using a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and the runner profile shown in Fig. 5 with the pad centered at 135 degrees trough azimuth. Pedestal and pad deflections were considered during each simulation.

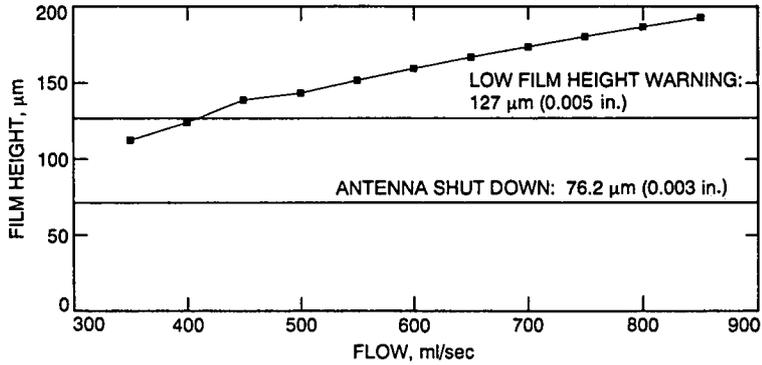


Fig. 12. Minimum film height as a function of recess flow rate when using ISO 150 oil. All of the hydrostatic bearing simulations used to produce this graph were performed using a pad load of 1.3×10^7 newtons, an oil temperature of 308 K, and the runner profile shown in Fig. 5 with the pad centered at 135 degrees trough azimuth. Pedestal and pad deflections were considered during each simulation.

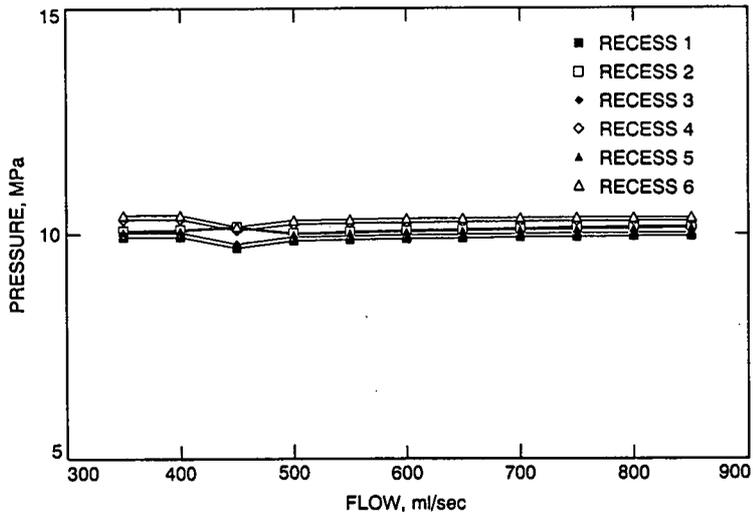


Fig. 13. Recess pressures as a function of recess flow rate when using ISO 150 oil. All of the hydrostatic bearing simulations used to produce this graph were performed using a pad load of 1.3×10^7 newtons, an oil temperature of 308 K, and the runner profile shown in Fig. 5 with the pad centered at 135 degrees trough azimuth. Pedestal and pad deflections were considered during each simulation.

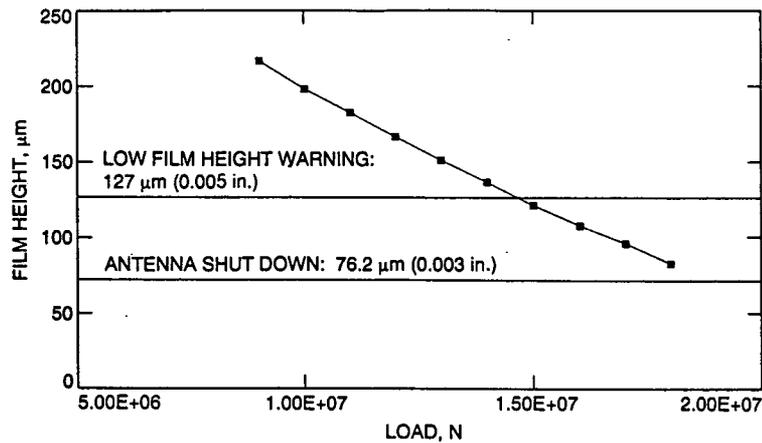


Fig. 14. Minimum film height as a function of pad load when using ISO 150 oil. All of the hydrostatic bearing simulations used to produce this graph were performed using an oil temperature of 308 K, recess flows of 550 ml/sec, and the runner profile shown in Fig. 5 with the pad centered at 135 degrees trough azimuth. Pedestal and pad deflections were considered during each simulation.

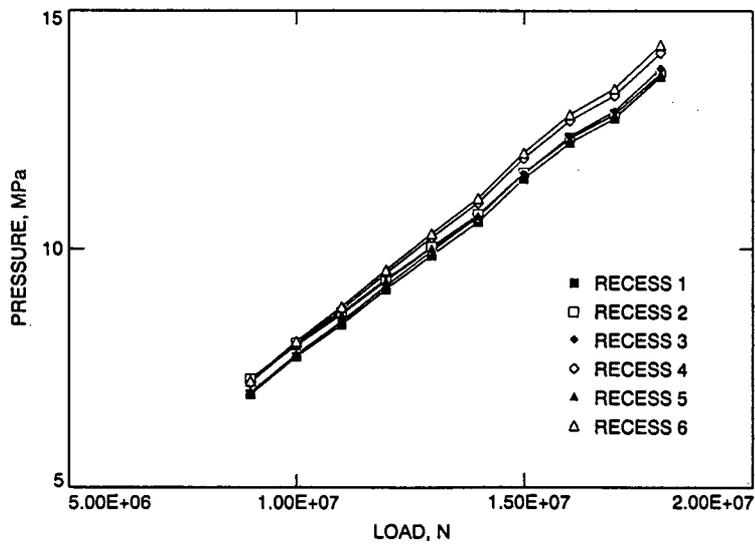


Fig. 15. Recess pressures as a function of pad load when using ISO 150 oil. All of the hydrostatic bearing simulations used to produce this graph were performed using an oil temperature of 308 K, recess flows of 550 ml/sec, and the runner profile shown in Fig. 5 with the pad centered at 135 degrees trough azimuth. Pedestal and pad deflections were considered during each simulation.

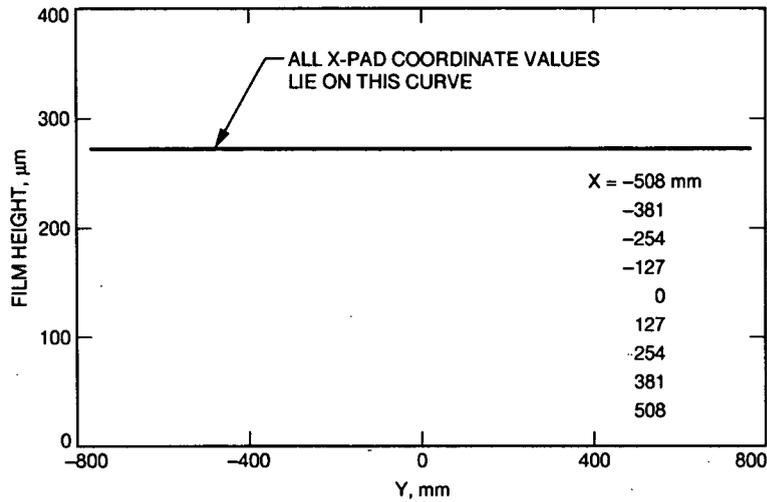


Fig. 16. Oil film height profile as a function of the Y-pad coordinate when the pedestal and pad do not deflect. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

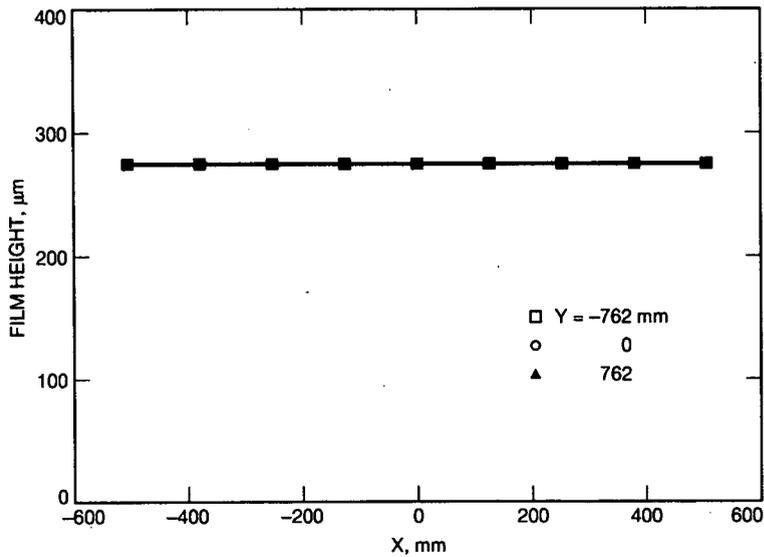


Fig. 17. Oil film height profile as a function of the X-pad coordinate when the pedestal and pad do not deflect. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

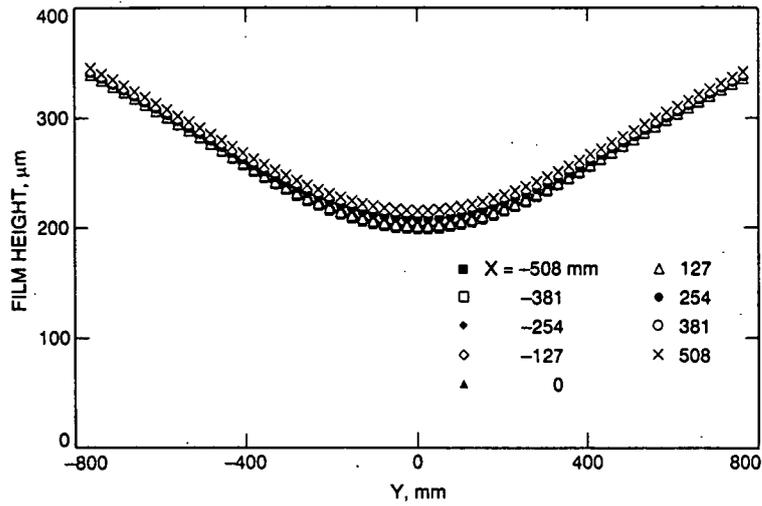


Fig. 18. Oil film height profile as a function of the Y-pad coordinate when the pedestal does not deflect and the pad deflects. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

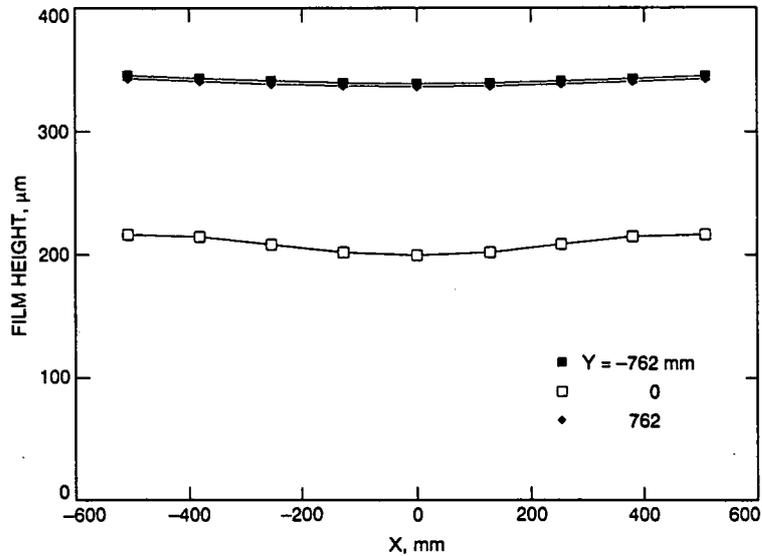


Fig. 19. Oil film height profile as a function of the X-pad coordinate when the pedestal does not deflect and the pad deflects. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

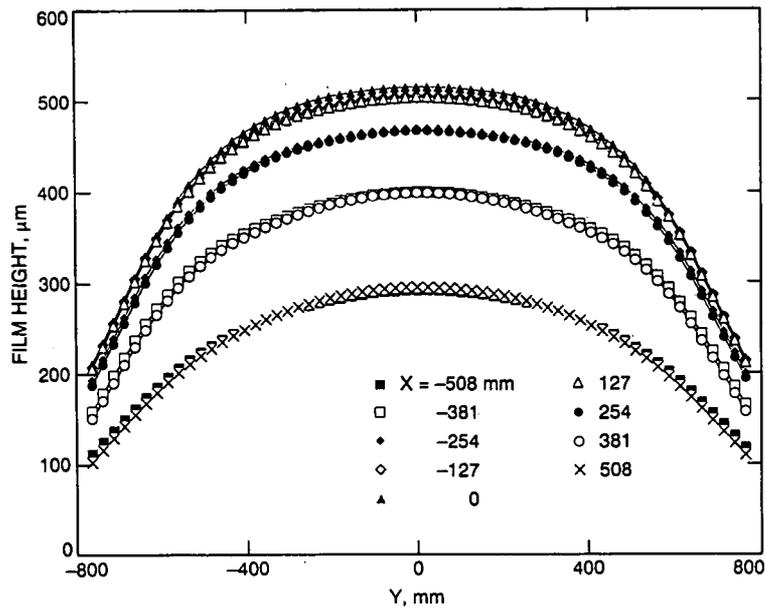


Fig. 20. Oil film height profile as a function of the Y-pad coordinate when the pedestal deflects and the pad does not deflect. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

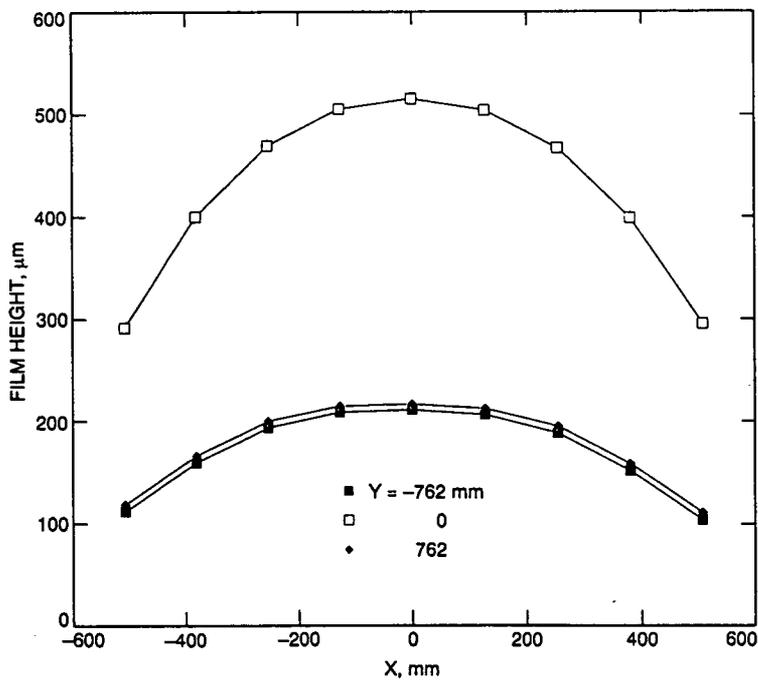


Fig. 21. Oil film height profile as a function of the X-pad coordinate when the pedestal deflects and the pad does not deflect. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

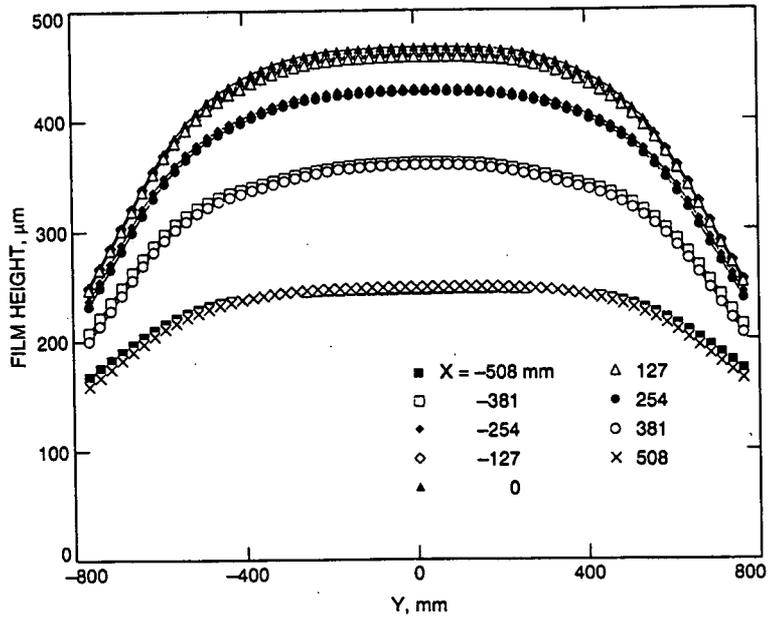


Fig. 22. Oil film height profile as a function of the Y-pad coordinate when the pedestal and the pad deflect. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

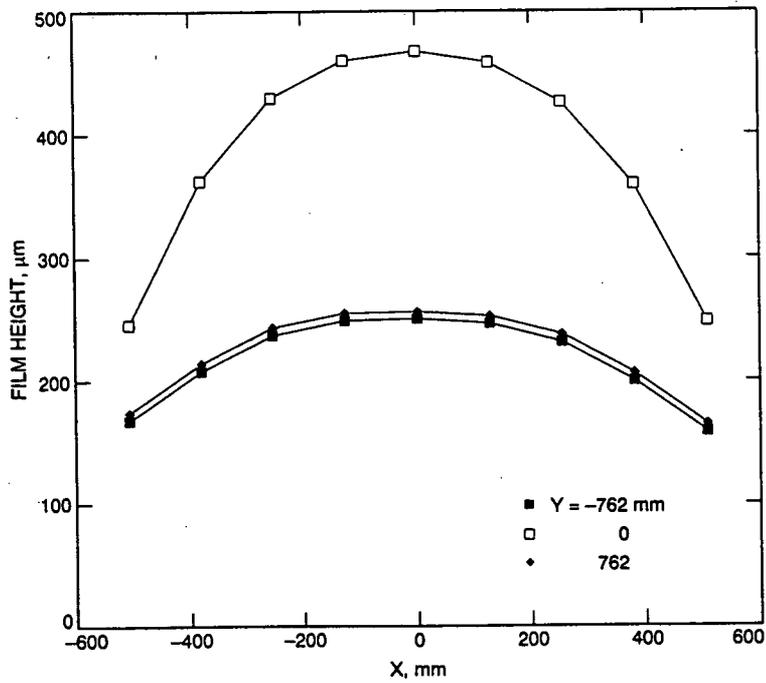


Fig. 23. Oil film height profile as a function of the X-pad coordinate when the pedestal and the pad deflect. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and a flat runner profile.

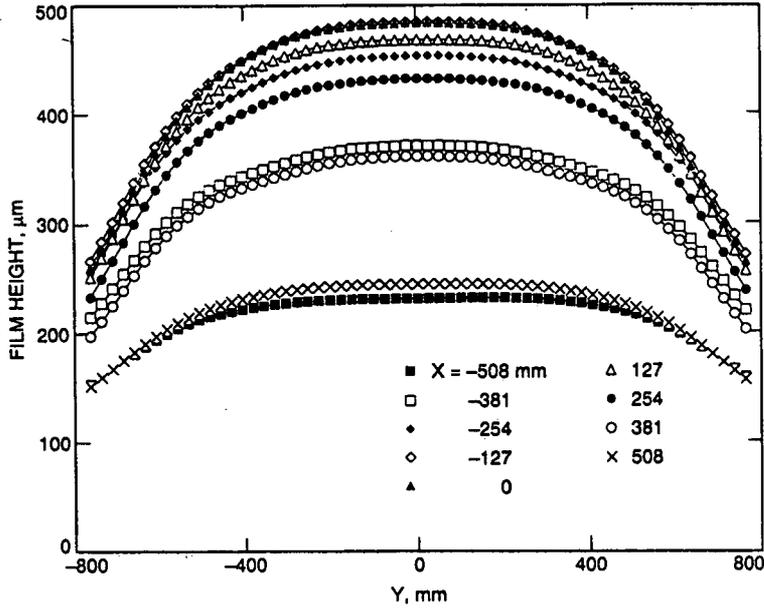


Fig. 24. Oil film height profile as a function of the Y-pad coordinate when the pedestal and the pad deflect. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and the runner profile shown in Fig. 5 with the pad centered at 135 degrees trough azimuth.

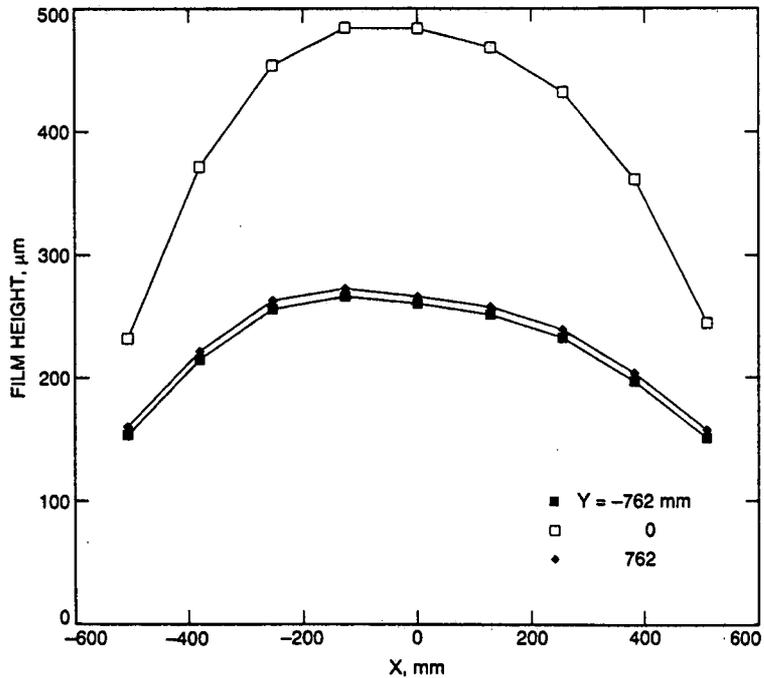


Fig. 25. Oil film height profile as a function of the X-pad coordinate when the pedestal and the pad deflect. The computer simulation was performed using ISO 150 oil at a temperature of 308 K, a pad load of 1.3×10^7 newtons, recess flows of 550 ml/sec, and the runner profile shown in Fig. 5 with the pad centered at 135 degrees trough azimuth.

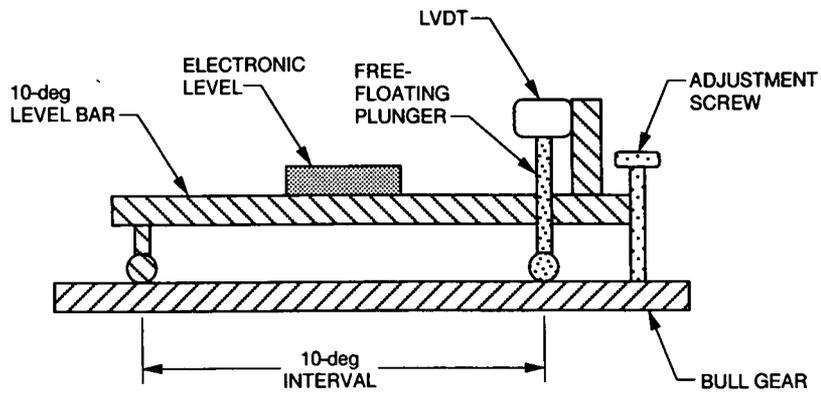


Fig. 26. Ten-degree level bar used to measure the bull gear profile.

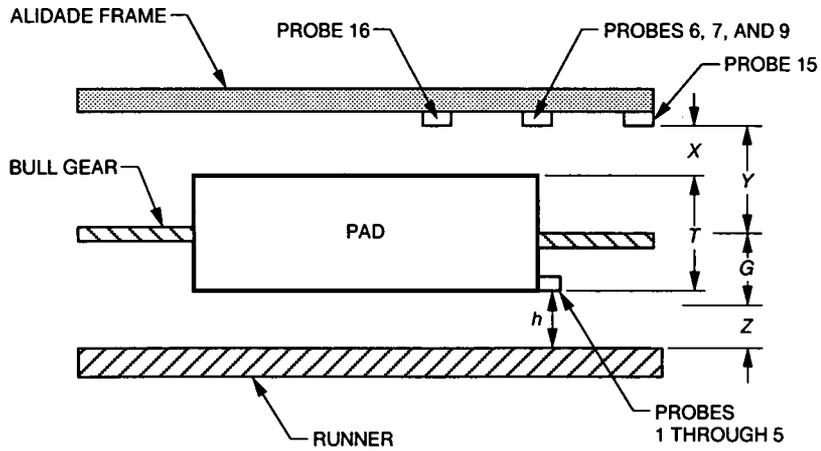


Fig. 27. Hydrostatic bearing pad, runner, bull gear, and alidade frame viewed radially while facing toward the antenna center.

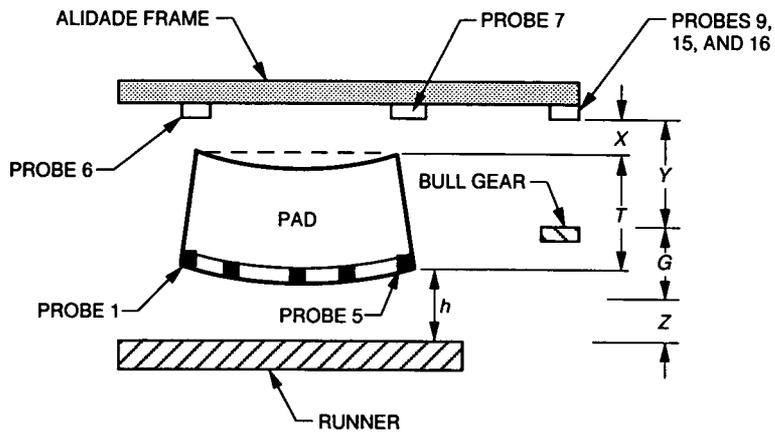


Fig. 28. Hydrostatic bearing pad, runner, bull gear, and alidade frame viewed tangentially while facing toward the probe bar on pad 3. The probe bar is assumed to deflect with the pad.

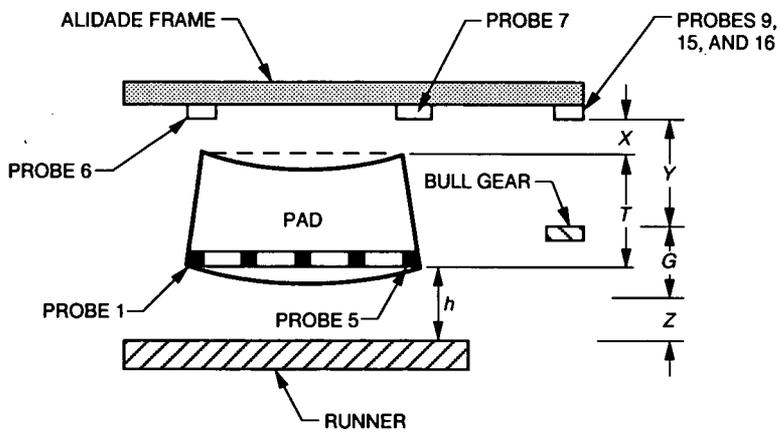


Fig. 29. Hydrostatic bearing pad, runner, bull gear, and alidade frame viewed tangentially while facing toward the probe bar on pad 3. The probe bar is assumed not to deflect.

516-32
167 936

500771

1-10

Acoustic Charge Transport Technology Investigation for Advanced Development Transponder

S. Kayalar

Spacecraft Telecommunications Equipment Section

Acoustic charge transport (ACT) technology has provided a basis for a new family of analog signal processors, including a programmable transversal filter (PTF). Through monolithic integration of ACT delay lines with GaAs metal semiconductor field effect transistor (MESFET) digital memory and controllers, these devices significantly extend the performance of PTFs. This article introduces the basic operation of these devices and summarizes their present and future specifications. The production and testing of these devices indicate that this new technology is a promising one for future space applications.

I. Introduction

The acoustic charge transport (ACT) processor is a new class of signal-processing devices that enables the user to sample a time signal (digital) without quantizing the sample amplitudes (analog) [1]. A comparison of signal-processing architectures using conventional digital processors and ACT processors is shown in Fig. 1. The ACT processor, which eliminates the use of analog-to-digital (A/D) and digital-to-analog (D/A) converters, offers a capability that can be advantageous in a wide variety of applications, such as equalizers, programmable IF filters, and numerically controlled oscillators.

Two programmable transversal filter (PTF) modules, which use ACT technology, and data acquisition and analysis software (WAVE) were procured from Electronic Decisions Incorporated (EDI). This article summarizes the investigation and the testing performed on the ACT PTF

module, and considers possible functions that can be implemented at different stages in the advanced development transponder (ADT).

II. Significance of ACT in the Advanced Development Transponder Design

The PTF uses an ACT processor that enables the user to process analog signals without quantizing them. A benefit of using this device in the ADT is the ability to control independently the center frequency, bandwidth, amplitude, and phase of the PTF. A filter with variable center frequency allows the designer to use an optimum IF frequency or, if needed, a variable IF frequency to minimize the effects of mixing products and other unwanted signals. By varying the bandwidth of the ACT PTF filter, the receiver bandwidth may be matched closely to the information bandwidth of the desired signal, thus minimizing

the channel noise. Figure 2 shows a block diagram of the ADT with the shaded blocks indicating the possible usage of the ACT processor. Included among these blocks are the bandpass filters (BPFs), downconverters, and voltage-controlled oscillators (VCOs). At present, the ACT PTF filter realizations are limited in that these filters do not have high Q factors that are needed in the transponder. Thus, the most likely application of ACT PTF in the ADT design is the ranging channel bandpass filter.

III. ACT Principles of Operation

The acoustic charge transport technique of the ACT processor is drawn from a technology mix of surface acoustic waves (SAWs), charge-coupled devices (CCDs), and GaAs [2]. The ACT device comprises a single GaAs integrated circuit (IC) that requires minimal support circuitry and uses an on-chip delay line that can be tapped at several points without distorting the signal.

Rather than converting signals to SAWs, the ACT devices use acoustic waves to transport sampled signals in a bucket-brigade fashion, much like CCDs. Unlike CCDs, however, ACT channels transport charge packets using SAW-induced electric fields, and signals can be sensed non-destructively at several points.

An input signal, shown as $x(t)$ in Fig. 3, is applied to an ohmic input contact (IC) on a depleted n -type epitaxial layer on a semi-insulating GaAs substrate. At one end of the channel, acoustic waves are generated by a high-Q piezoelectric transducer in a proton-isolated region. These waves propagate across the device at the speed of sound (2,864 m/sec). Since GaAs is a piezoelectric material, the propagating surface wave induces within the epitaxial layer a series of electric fields that may be thought of as traveling potential wells. As each well passes under the input contact, a number of electrons proportional to the input voltage at that instant are injected into the well. Through this process the analog input signal has been converted into a discrete-time signal. The signal is thus sampled in time, but not quantized in amplitude. This key feature of the ACT device is also indicated in Fig. 1.

Because the amount of charge is proportional to the input electrode's voltage, each packet represents a sample of the input signal. The fields associated with the surface wave then transport these samples through the ACT channel. The inherently smooth and continuous nature of the traveling wave transport accounts for the processor's high efficiency and low noise.

The sampling rate of an ACT channel, which is equal to the frequency of the acoustic waves, typically ranges from

300 to 600 MHz. High device speeds are possible because of the high electron mobility of GaAs. Acoustically transported charge packets are sensed by one or more output electrodes placed on the surface of the delay line. Charge packets in the channel induce image charges on the sensing electrodes, creating a voltage at each electrode that is proportional to the amount of charge in its vicinity. Electrode voltages are created without affecting the charge packets by the nondestructive sensing (NDS) circuitry. At the end of the ACT channel, an ohmic extraction contact (EX) is used to remove the charge from the wells.

A standard ACT chip, shown in Fig. 7, consists of a delay line, a SAW drive, and tap-weight and tap-address registers.

IV. The EDI Programmable Transversal Filter

The ACT tapped delay line has been used to implement a complete and compact PTF through the addition of on-chip tap-weighting and memory circuits, as shown in Fig. 4 [3,4]. The tap weights are used to set the magnitude of the coefficients of the desired impulse response. The accumulation function required by the transversal filter is performed by two summing buses and an operational amplifier.

The presently available ACT PTFs have 128 taps that may be set to any of 31 values between +1 and -1 (5-bit tap weighting). The input sampling rate is 360 MHz, and the tap spacing is 5.6 nsec. This gives a single-tap bandwidth of 180 MHz and a filter Nyquist interval of 90 MHz. A single tap may be programmed in 1 μ sec, and the entire device may be programmed in under 100 μ sec.

Table 1 shows a comparison between the EDI's current design and projected future designs of the ACT PTF.

V. The EDI PTF Module

The EDI PTF module, shown in Fig. 8, consists of an ACT processor, digital control and interface circuits, and a chassis with connectors for an external RF signal, DC power supplies, and a digital I/O port.

The ACT processor, shown in Fig. 9, is the physical and functional center of the module. It contains the ACT PTF device, I/O RF interface amplifiers, a hybrid SAW drive amplifier, bias and decoupling circuitry, and digital connections. The ACT PTF device includes a 128-tap ACT delay line, 128 5-bit (4-bit magnitude and sign) attenuators, a 128 \times 5 RAM architecture, and address decoding

circuitry. This chip provides a complete PTF in a monolithic form. An input amplifier and a differential output amplifier are added to reduce insertion loss and to provide matching to a 50- Ω system. The hybrid SAW drive amplifier is used in conjunction with SAW transducers on the PTF chip to form a delay-line oscillator which supplies the 1-W, 360-MHz signal needed to transport the charge.

VI. ACT Testing and Applications

A. Space Qualification

EDI has performed some preliminary radiation and temperature tests on the ACT device. The neutron radiation hardness tests were conducted using four neutron doses of 5×10^{11} , 2×10^{12} , 1×10^{13} , and 1×10^{14} per cm^2 . These tests indicate that in spite of relatively large losses in the sheet conductance of the epitaxial layer which forms the transport channel, the key operational parameters of the ACT device—namely charge injection transconductance, charge capacity, and Nyquist bandwidth frequency response loss—are barely affected. This performance is attributed to the advanced architecture of the ACT channel structure which uses a graded doping profile to maintain an extremely narrow charge packet cross section. Thus the ACT device exhibits excellent neutron radiation hardness. The total dose gamma radiation hardness test has not yet been conducted, but it is believed that this type of radiation is expected to present less risk than neutron radiation effects because the ACT device is implemented using GaAs technology—GaAs devices are inherently radiation hard.

The ACT device has also been tested by EDI at temperatures 0, 25, and 50 deg C, with limited testing at -20 deg C. It has performed satisfactorily at all temperatures. The Missile Division of the Hughes Technology Center at Carlsbad, California, is currently working towards the space qualification of the ACT chips.

B. ACT Foundry

EDI is transferring the ACT technology to Hughes for a large-scale production of the standard ACT PTF. One of the immediate applications of this standard ACT PTF will be as matched filters in commercial disk drives. Three-inch GaAs wafers are being used for the fabrication process. Hughes is implementing qualified GaAs processes for the fabrication of ACT chips. The current processes are qualified for level 3. They need to be brought up to level 5 for space qualification. Hughes is maintaining 3-sigma performance up to the end-product, including the fabrication processes.

C. Performance Verification

The ACT PTF module test setup (Fig. 5) was configured in the Transponder Development Laboratory by using an IBM computer and WAVE data acquisition and analysis software. An IEEE-488 data acquisition card was installed in the IBM computer to have the capability of observing, storing, and analyzing the I/O RF signals.

The WAVE software allows the user to specify the desired response, to program the PTF modules, and to access an extensive set of tools for analyzing, manipulating, and displaying waveforms. Some of the prominent features of WAVE software are: built-in signal-processing functions and an editor, built-in interface functions for an IEEE-488 card, a finite-impulse-response (FIR) filter-design function, PTF tap-weights loading utility, individual/group PTF tap control, and interactive graphics.

D. Filtering Application

A series of tests were performed at JPL to verify the operating characteristics of the ACT PTF module when it is used as a filter. The test equipment was configured as shown in Fig. 5.

Figure 10 shows the frequency response of the ACT PTF when a single tap is turned on (i.e., set for a tap weight of +1) with all other taps turned off in a manner to balance the capacitive loading equally between the positive and negative bus bars. The frequency response is quite flat, with the exception of the slope caused by the ACT element factor. The small peak at 89 MHz is due to the summation of the off (but not exactly zero) taps.

Figure 11 shows the frequency response when the taps are set alternatively to +1 and -1. The highest peak is centered at about 89 MHz, with tapering peaks on both sides.

Figure 12 shows multiple traces of a Hanning-weighted bandpass filter having a 2-MHz passband with center frequencies at 41, 46, 51, and 56 MHz. The roll-off in the passband gain is due to the ACT element factor. The passband gain can be adjusted to compensate for the element factor or other system gain variations.

Figure 13 shows the filter at 56 MHz with its bandwidth adjusted to 2, 4, 6, and 8 MHz. The increase in gain as the bandwidth decreases is expected since a processing gain is realized for narrower passbands which have longer impulse responses and thus make use of a large number of taps.

Passband phase control is also easily achieved as shown in Fig. 14. Here the phase is adjusted in steps of 45 deg,

from -180 to 180 deg. The passband shape remains undisturbed by the phase adjustments as shown in Fig. 15.

E. Oscillator Application

A few tests were performed to verify the operating characteristics of the ACT PTF module when it is used as an oscillator. The digitally controlled oscillator (DCO) configured with the ACT PTF is shown in Fig. 6. The output of the PTF is fed back to the input, such that the loop gain is approximately 1, causing the circuit to oscillate. The loop gain is greater than 1 only near the center of the passband, giving coarse selection of frequency. The frequency that causes the loop phase shift of 0 deg is the frequency of oscillation.

The ACT DCO has a programmable range from 15 to 80 MHz. It can shift relative frequencies by steps as small as 1 kHz for use in frequency tracking applications. It can be set to an absolute frequency with a precision of ± 2.5 kHz if module temperature is considered in the tap settings. Figure 16 shows a 1 -kHz-per-division superposition of the DCO output when it was programmed to 56.010 and 56.011 MHz. The verification of phase noise and stability of the ACT DCO is included as part of the future plan.

F. Items Remaining

The following tasks need to be performed to more fully characterize the ACT PTF module for possible inclusion and implementation in the ADT:

- (1) Verification of the phase noise and stability characteristics of the ACT DCO.
- (2) Verification of whether the ACT PTF module can be used as a downconverter.

- (3) Replacement of the second IF bandpass filter in the breadboard transponder and verification of its operation in terms of: I/O impedance, I/O signal levels, insertion loss, power consumption, temperature stability, roll-off factors, passband gain, and out-of-band rejection.
- (4) Replacement of one of the reference oscillators used in the breadboard transponder and verification of its operation in terms of: I/O impedance, I/O signal levels, insertion loss, power consumption, temperature stability, and frequency stability.
- (5) In cooperation with the Electronic Decisions Division of Comlinear Corporation, assessment of the viability of implementing the functions of the ACT processor discussed above for the ADT. It may be possible to custom design an ACT processor which meets the requirements stated in Table 1, specifically a filter with a narrower bandwidth, either by cascading two ACT filters to extend the channel length or by using heterojunction ACT (HACT) devices.

VII. Conclusion

In this article the principal characteristics and the unique features of the Acoustic Charge Transport Programmable Transversal Filter have been reviewed. Tests have been performed to characterize some of the applications, and this filter's possible use in the Advanced Development Transponder has been considered. The principal advantages of using such a device are its size, weight, power consumption, and programmability. In addition, it offers digital signal-processing techniques without A/D and D/A conversions.

References

- [1] A. Landrie, "ACT Technology bestows accelerated signal processing," *Microwaves and RF*, vol. 28, no. 8, pp. 147-150, August 1989.
- [2] M. J. Hoskins, J. Morkoc, and B. J. Hunsinger, "Charge transport by surface acoustic waves in GaAs," *Applied Physics Letters*, vol. 41, no. 4, pp. 332-334, August 1982.
- [3] D. A. Fleish and G. C. Pieters, "The ACT Programmable Transversal Filter," *Microwave Journal*, vol. 34, pp. 284, 289-292, May 1991.
- [4] F. M. Fliegel, R. L. Martin, and F. Guediri, "Performance of acoustic charge transport filters," *IEEE Trans. on Microwave Theory and Techniques*, vol. 35, no. 12, pp. 1164-1168, December 1987.

Table 1. Programmable transversal filter specifications.

Specification	EDI Present design	EDI Future design	Transponder design
Sampling rate, MHz	360	1000	360-2000
Number of taps	128	1024	128-1024
Tap spacing, nsec	5.6	<3	0.5-5.6
Tap weights, bits (bipolar)	5	8	5-10
Programming time	<1 μ sec per tap	—	—
Dynamic range			
Spurious free, dB	40	50-60	40-70
Blocking, dB	55	>60	>60
Max. input level, dBm	-10	-10	-10
Insertion loss, dB	20	0	0
I/O impedance, Ω	50	50	50
Temperature range, deg C	0-70	0-70	-30 to +85
Power dissipation, W	<3	<1	<1
Radiation hardness, kilorads	—	—	100
Package size, cm x cm	3 x 3	3 x 3	3 x 3

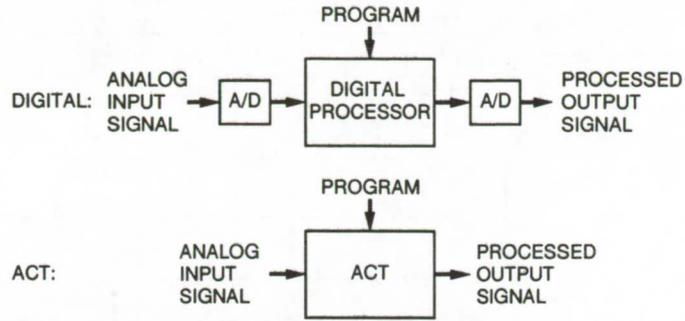


Fig. 1. Digital processing versus ACT processing.

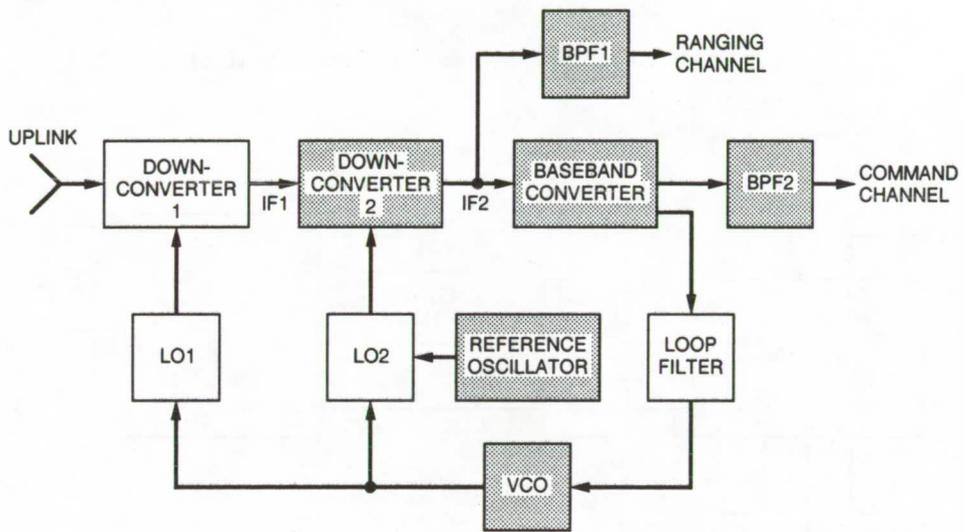


Fig. 2. Possible ACT applications in the advanced development transponder.

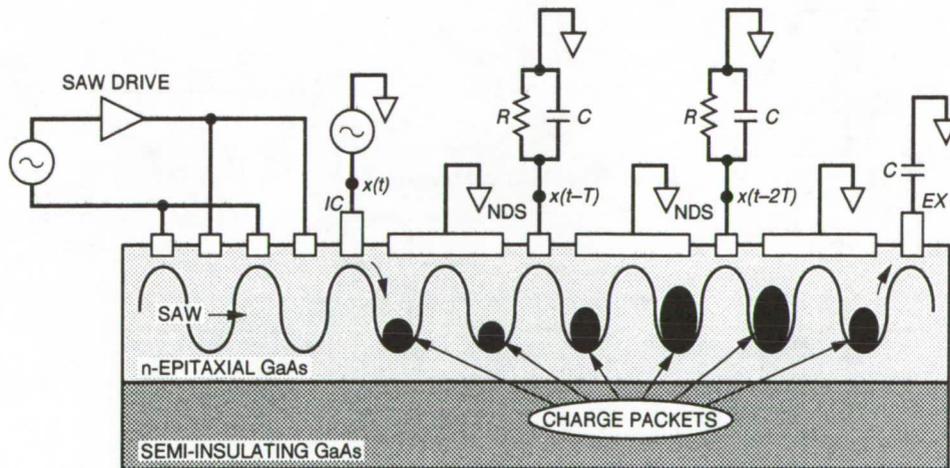
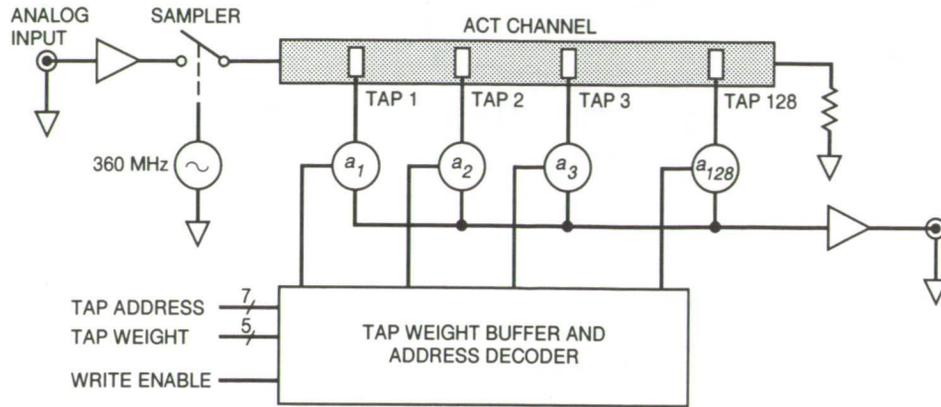


Fig. 3. ACT delay-line structure.



IMPULSE RESPONSE:
$$h(t) = \sum_{n=1}^{128} a_n \delta(t - nT)$$

TRANSFER FUNCTION:
$$H(\omega) = \sum_{n=1}^{128} a_n e^{-jn\omega T}$$

Fig. 4. EDI ACT programmable transversal filter (PTF).

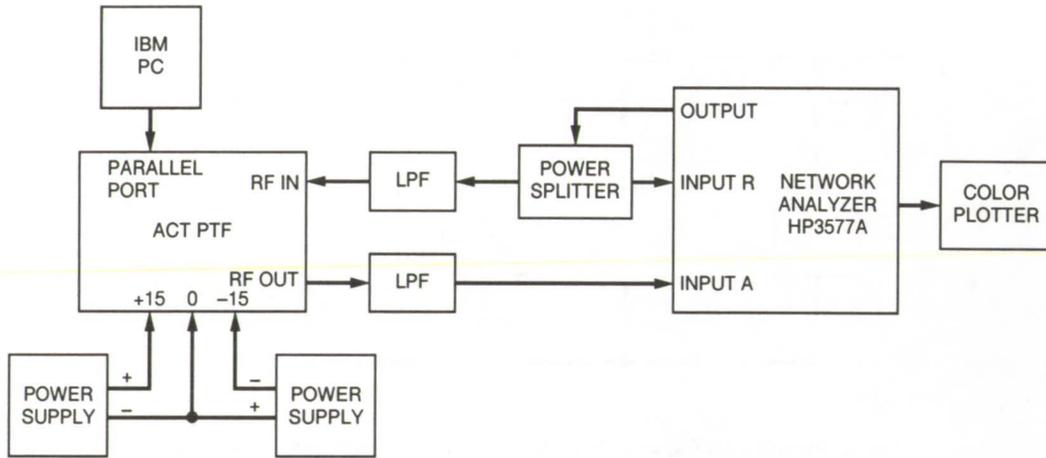


Fig. 5. Test configuration for the ACT PTF.

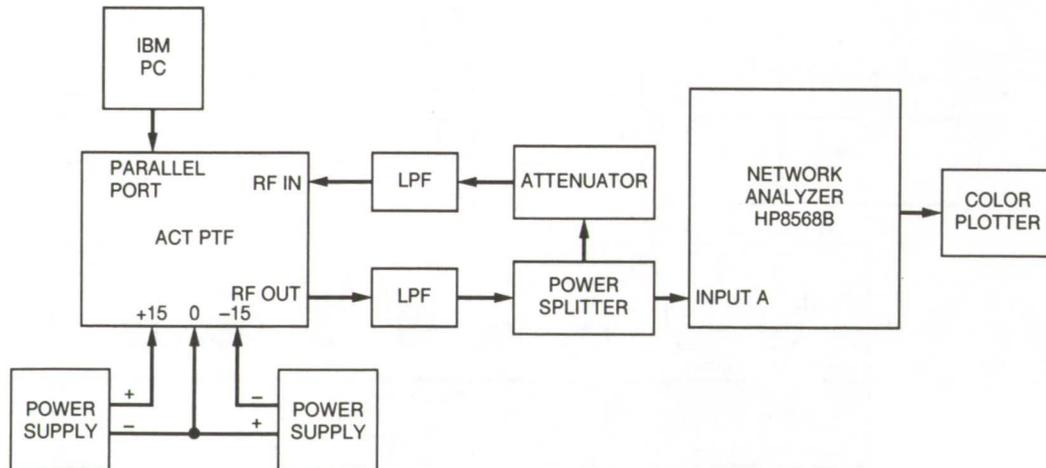


Fig. 6. Test configuration for the digitally controlled oscillator.

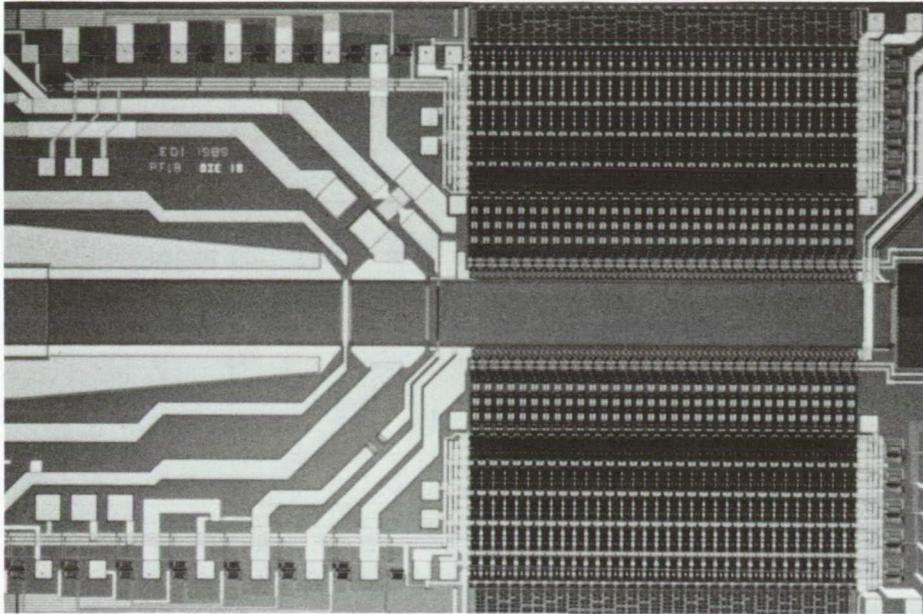


Fig. 7. ACT chip.

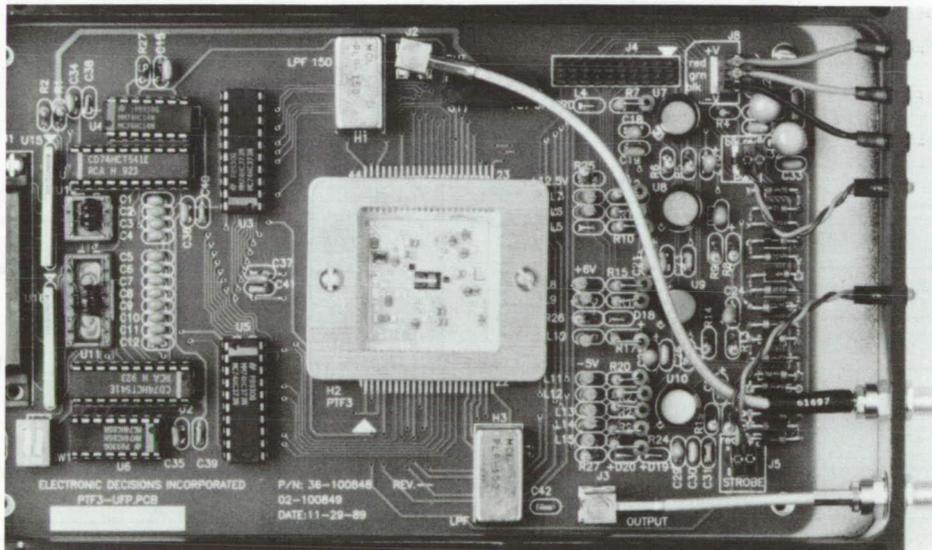


Fig. 8. ACT PTF module.

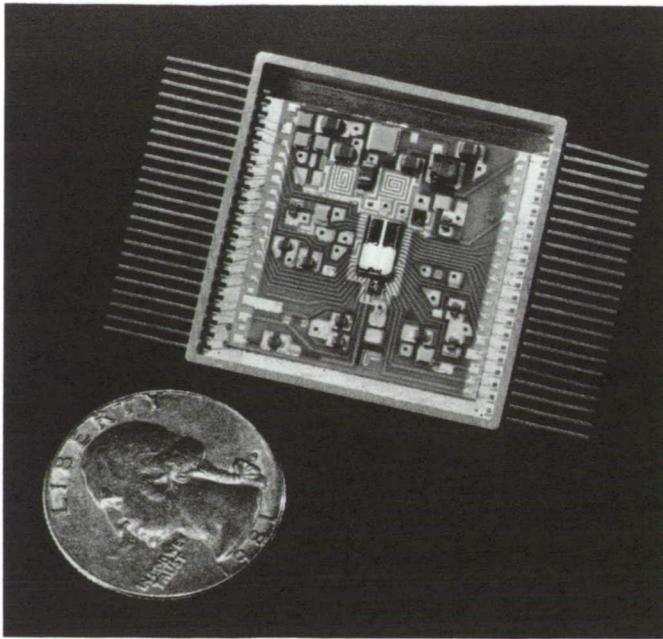
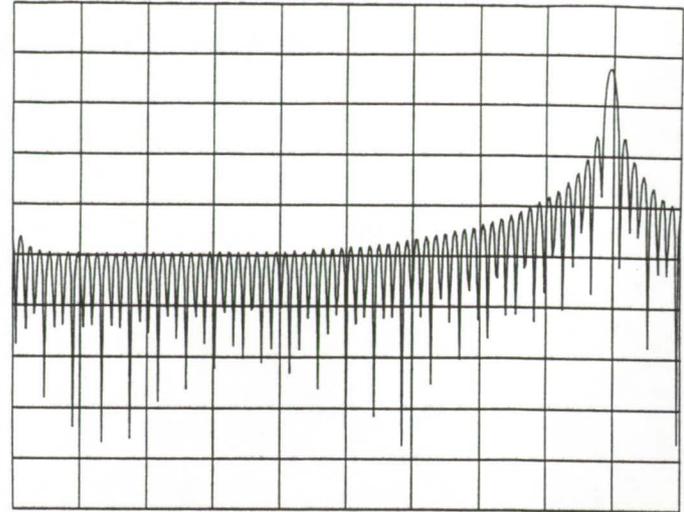


Fig. 9. ACT processor.

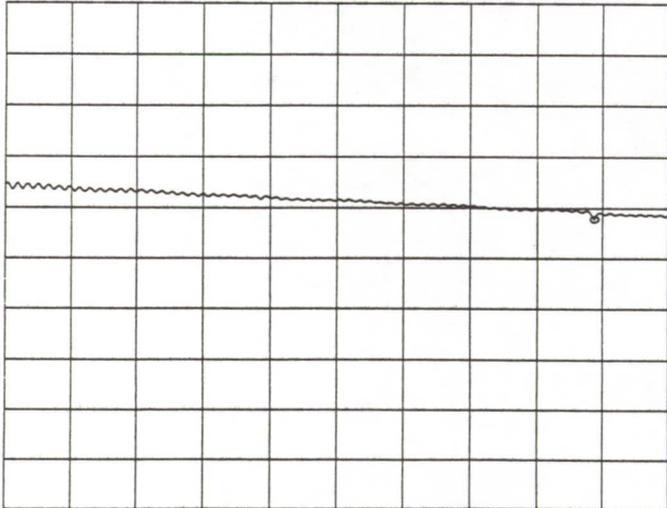
REF LEVEL /DIV
10.000dBm 10.000dB



START 1 000 000.000Hz STOP 100 000 000.000Hz
AMPTD -3.0dBm

Fig. 11. ACT PTF frequency response with alternating +1 and -1 taps.

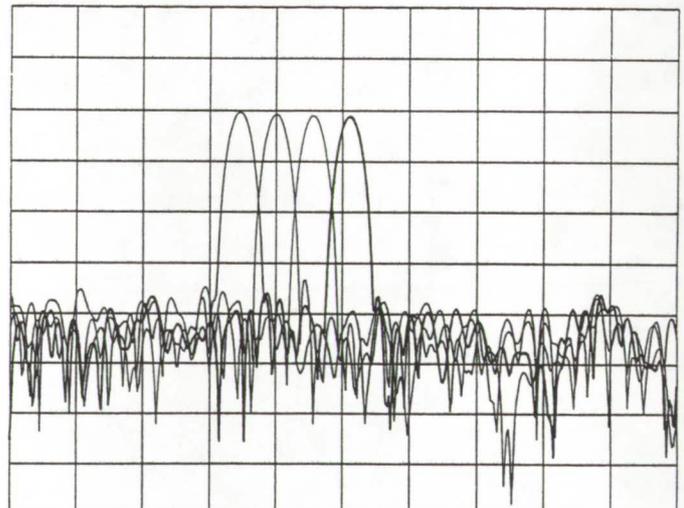
REF LEVEL /DIV MARKER 89 875 000.000Hz
20.000dB 10.000dB MAG (A/R) -22.070dB



START 10 000 000.000Hz STOP 100 000 000.000Hz
AMPTD -3.0dBm

Fig. 10. ACT PTF single-tap frequency response.

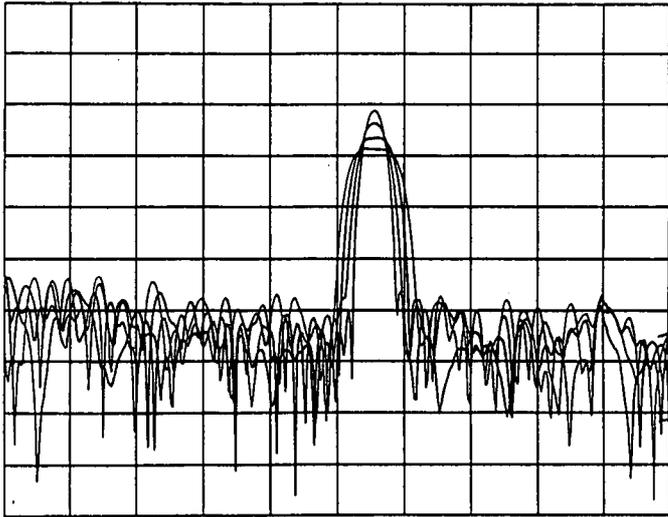
REF LEVEL /DIV
20.000dB 10.000dB



START 10 000 000.000Hz STOP 100 000 000.000Hz
AMPTD -3.0dBm

Fig. 12. ACT PTF center frequency control.

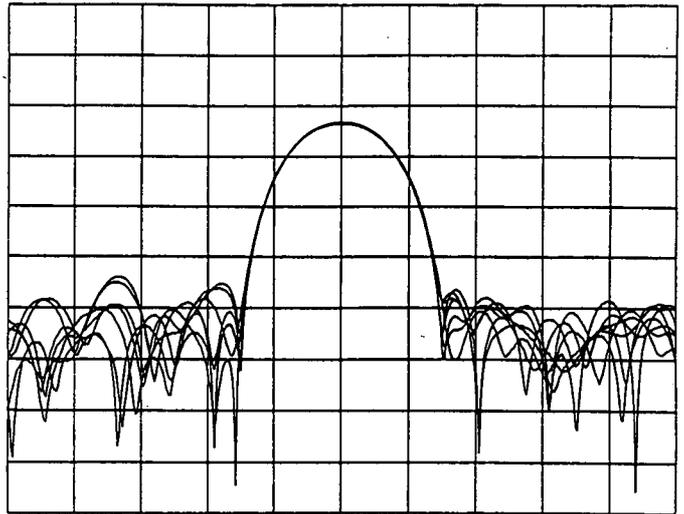
REF LEVEL /DIV
 10.000dBm 10.000dB



START 1 000 000.000Hz STOP 100 000 000.000Hz
 AMPTD -3.0dBm

Fig. 13. ACT PTF bandwidth control.

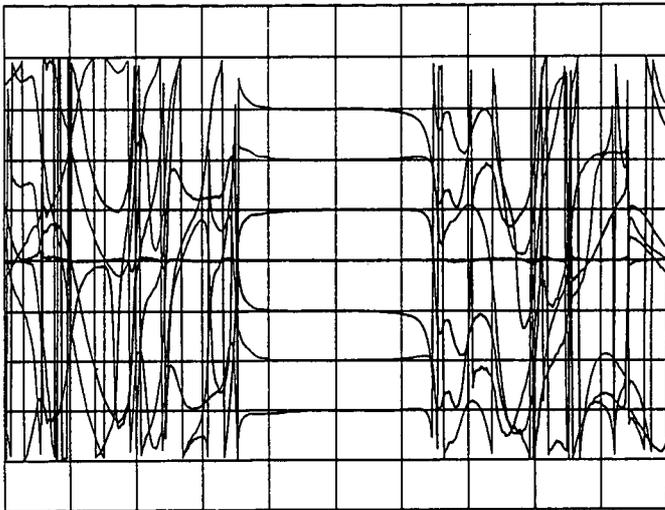
REF LEVEL /DIV
 20.000dB 10.000dB
 0.0deg 45.000deg



START 42 000 000.000Hz STOP 70 000 000.000Hz
 AMPTD -3.0dBm

Fig. 15. ACT PTF passband constant amplitude during phase control.

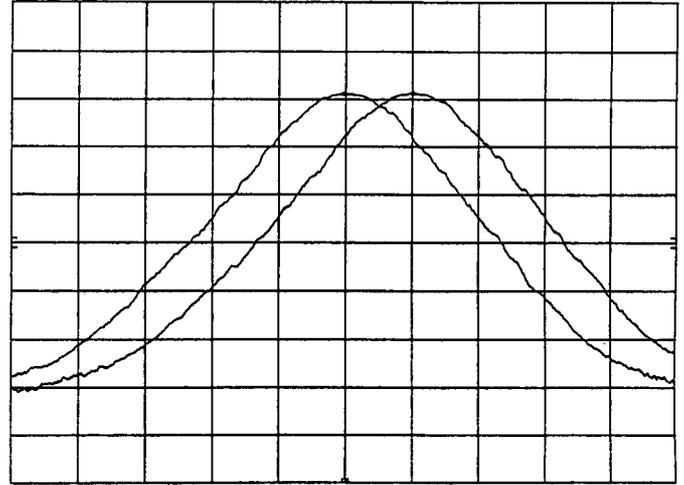
REF LEVEL /DIV MARKER 60 130 000.000Hz
 20.000dB 10.000dB MAG (A/R) -37.370dB
 0.0deg 45.000deg



START 42 000 000.000Hz STOP 70 000 000.000Hz
 AMPTD -3.0dBm

Fig. 14. ACT PTF passband phase control.

REF .0 dBm ATTEN 10 dB MKR Δ 1.04 kHz
 .10 dB



CENTER 56.00950 MHz SPAN 10.00 kHz
 RES BW 1 kHz VBW 3 Hz SWP 15 sec

Fig. 16. ACT PTF oscillator relative frequency stability.

500782

517-32

167937

p-19

The Flight Performance of the Galileo Orbiter USO

D. D. Morabito

Telecommunications Systems Section

T. P. Krisher

Navigation Systems Section

S. W. Asmar

Telecommunications Systems Section

Results are presented in this article from an analysis of radio metric data received by the DSN stations from the Galileo spacecraft using an Ultrastable Oscillator (USO) as a signal source. These results allow the health and performance of the Galileo USO to be evaluated, and are used to calibrate this Radio Science instrument and the data acquired for Radio Science experiments such as the Redshift Observation, Solar Conjunction, and Jovian occultations. Estimates for the USO-referenced spacecraft-transmitted frequency and frequency stability have been made for 82 data acquisition passes conducted between launch (October 1989) and November 1991. Analyses of the spacecraft-transmitted frequencies show that the USO is behaving as expected. The USO was powered off and then back on in August 1991 with no adverse effect on its performance. The frequency stabilities measured by Allan deviation are consistent with expected values due to thermal wideband noise and the USO itself at the appropriate time intervals. The Galileo USO appears to be healthy and functioning normally in a reasonable manner.

I. Introduction

This is an article on the efforts to characterize the instrument used for the Galileo Radio Science investigation. It discusses the performance of the Ultrastable Oscillator (USO), which, when the spacecraft and ground elements of the instrument are configured in the one-way mode, is the limiting error source observed in the received Doppler data at time intervals where there is sufficient signal power.

Galileo was originally scheduled to be launched in May 1986, but was delayed due to the Challenger accident. Galileo was launched on October 18, 1989, and on December 5, 1989, the USO was turned on. Galileo acquired gravity assists from Venus (February 1990) and Earth (December 1990, December 1992) and will go into orbit around Jupiter in December 1995. Between December 1989 and November 1991, 94 one-way passes were scheduled for 2-hr data acquisition periods. However, not all scheduled

passes resulted in valid data; twelve passes were lost for various reasons (six were aborted after a spacecraft safing anomaly early in 1990, four were lost due to a crashed sequence in March–April 1991, and two were lost due to tape problems). Data for the remaining 82 passes were acquired by the DSN and delivered to the Galileo Radio Science Team (RST) for analysis in the form of Archival Tracking Data File (ATDF) tapes. Estimates of the spacecraft-transmitted frequency and frequency stability were made for the 82 passes conducted during these first two years of cruise. Each pass consists of about 2 hr of Doppler data sampled at 1/sec, and hydrogen masers served as the frequency and timing references. The received Doppler data were converted into estimates of spacecraft-transmitted frequencies and frequency residuals after accounting for spacecraft trajectory and other effects.

II. Purpose

The USO-referenced data were acquired for two purposes: (1) the scientific investigation of the redshift phenomenon, which is the frequency shift as the spacecraft moves in and out of the gravitational fields of massive bodies in the solar system as predicted by Einstein's theory of general relativity, and (2) the engineering evaluation of the USO frequency and frequency stability for calibration purposes and to evaluate the health and performance of the USO. These calibration data serve as a baseline for Radio Science experiments such as the Redshift Observation, Solar Corona Experiment, and occultations of Jupiter and its satellites. This article will focus only on the engineering aspects of the USO data. The scientific results are discussed in [1]. For discussions of the expected scientific results of the Galileo Radio Science investigations, the reader is referred to [2,3].

The goal of the USO analysis is thus to establish the USO-referenced spacecraft-transmitted frequency and the frequency stability associated with it, as well as phase noise and spectral purity (when using open-loop data) and to build a database containing the statistics and all parameters relevant to the measurements.

III. Spacecraft Configuration

The Galileo spacecraft configuration for these Radio Science tests is the normal cruise configuration, with the exception of a "TWNC-ON" command which enables the USO to be the radio downlink reference; this is the mode required for the redshift experiment and future occultation investigations. No other changes (e.g., a modulation index change) were requested for this purpose. The majority of

the tests were performed on a quiet spacecraft where no maneuvers or other motion were permitted to occur.

The USO, a science payload instrument integrated with the spacecraft's telecommunications subsystem, was manufactured by Frequency Electronics Inc., New York, between 1975 and 1976. The USO which resides in the Radio Frequency Subsystem (RFS) of the Galileo orbiter is serial number 4, from the same lot as the USO's flown on Voyagers 1 and 2. The Voyager 1 USO failed in November 1992 (after 15 yr of continuous service), engendering concerns about the survival lifetime of the Galileo USO. The USO is a dual-oven-controlled device with an AT-cut quartz-crystal (SiO_2) resonator. The design output frequency is 19.125000 MHz corresponding to channel 14. When the USO is the downlink signal source, it drives the 2.3-GHz exciter. The output frequency is multiplied in the transponder by a factor of 120 to produce the 2.3-GHz transmitted signal (~ 2294.9976 MHz). The 2.3-GHz Traveling Wave Tube Amplifier (TWTA) amplifies the signal it receives from the 2.3-GHz exciter to either one of two power levels: the high-power mode (27 W) or the low-power mode (9 W). The TWTA has routinely been configured to the high-power mode. The 2.3-GHz TWTA provides the amplified RF output to the High-Gain Antenna (HGA)/Low-Gain Antenna (LGA) transmit switch of the 2.3-GHz antenna switch, which connects the outputs of the 2.3-GHz TWTA's to either the HGA or the LGA. If the HGA is ever successfully deployed, an 8.4-GHz transmitted signal will also become available and will be 11/3 of the 2.3-GHz signal frequency.

All of the data acquired were at 2.3 GHz and right circularly polarized. The signal was transmitted from the spacecraft via either LGA-1 or LGA-2. LGA-1 is located on the spacecraft spin-axis in front of the HGA tip sunshade. LGA-2 is located on a boom 3.58 m away from the spin axis, and is pointed in the direction opposite that of HGA and LGA-1. Both LGAs work at 2.3 GHz only. LGA-2 was utilized for the period right after launch and the period right after the Earth 1 encounter. The time periods of the spacecraft LGA configuration for the data set analyzed in this article are:

10/18/89 to 03/15/90	LGA-2
03/15/90 to 12/08/90	LGA-1
12/08/90 to 01/31/91	LGA-2
01/31/91 to 11/30/91	LGA-1

to the line-of-sight elevation angle. A spacecraft spin model correction is applied to the estimated downlink frequency and residuals. The magnitude of this correction depends upon whether the spacecraft was in all-spin mode (0.0481 Hz) or dual-spin mode (0.0525 Hz). The sign of this correction depends upon whether the signal source was LGA-1 or LGA-2. RESID writes the residual frequencies and related information to the F52 output file.

Relativistic effects, including the gravitational redshift, are modeled in the analysis, and removed from the data. The results of the scientific analysis of these effects, the first test of the solar redshift with an interplanetary spacecraft, indicate that the total frequency variation as predicted by general relativity has been verified to an accuracy of 0.5 percent and the solar gravitational redshift to 1 percent [1]. Therefore, it can be stated with a reasonable degree of confidence that the redshift effect is removed to the stated accuracy.

STBLTY: Reads the RESID output file (F52), computes and applies a bias correction to the spacecraft-transmitted-frequency estimate relative to the center of weight of the residuals over the pass, and computes phase, Allan variance, frequency and phase power spectral densities (PSD's), and other relevant statistics, and writes the relevant information for the current pass to a database summary file. STBLTY also produces plots of relevant data quantities.

Once all of the USO passes have been processed through the above programs, the database file will contain a set of records for each processed pass. This file is then processed through the following two programs:

FITUSO: Reads the USO frequencies from the database, and fits and removes an aging model. The resulting post-fit residuals are written to a previously blank field in the database. The user can specify how many passes to skip, how many passes to accept and where the break between logarithmic and linear aging behavior occurs.

USOSMRY: Reads the USO database and displays graphically any requested quantities from a menu of available data types.

VII. Analysis Results

A. General Single Pass Results

Of 94 scheduled USO-referenced data acquisition passes, 82 were processed through the STBLTY program

set. Each pass was typically 2 hr in duration. All data processed were closed-loop data acquired by the Deep Space Stations (DSS's) sampled at 1/sec. Table 1 displays the year, day number, UTC start time, UTC end time, DSS station identification, signal level (AGC), and estimated spacecraft-transmitted frequency for each pass. Figure 2 displays a typical plot of the residuals after the removal of the spacecraft trajectory, the gravitational redshift, a simple troposphere model, and the effect of the spacecraft spin when LGA-1 is the signal source. Figure 3 displays the frequency residuals after every 60 points have been averaged, allowing long-period trends to be examined easily. Figure 4 displays the reconstructed phase for the residuals of Fig. 2. Figure 5 is the log of the Allan deviation for the residuals of Fig. 2. Figures 6 and 7 display the logs of the frequency spectral density and phase spectral density, respectively, for the residuals of Fig. 2. The spikes at about 0.05 Hz are related to the 3-rpm spacecraft spin.

B. LGA-2-Induced Spin Doppler Example

When LGA-2 is the signal source, there is a significant sinusoidal signature present in the received Doppler due to LGA-2 being mounted on a boom located 3.58 m from the spacecraft spin axis. Figure 8 displays a typical plot of frequency residuals for a pass where LGA-2 was used, after the trajectory and other effects were removed. Figure 9 clearly illustrates the sinusoidal signature for a selected 200-sec period.

In order to remove this signature from the data, a three-parameter sinusoidal model was iteratively fit to the Doppler residuals of Fig. 8. This model includes an amplitude, a frequency and a phase offset. After the model was successfully fit and removed, the resulting post-fit residuals (see Fig. 10) yielded Allan deviations consistent with those of LGA-1 passes, suggesting that the three-parameter sinusoidal model is sufficient for removal of the spin-induced Doppler signature.

For three passes where LGA-2 was the signal source, 91-01-14 (91-014), 90-01-09 (90-009) and 90-12-10 (90-345), dynamic activity occurred on board the spacecraft which introduced phase shifts into the data. The result is that the fit of the sinusoidal model failed to remove all of the induced off-axis LGA-2 Doppler signature, resulting in residuals as exemplified in Fig. 11 for pass 91-014. In this specific case, tape-recorder motion was known to occur, where the envelope changes in Fig. 11 correlate with the tape-recorder start and stop times. Figure 12 illustrates the resulting degradation to the Allan deviation (compare with Fig. 5). Pass 90-345 occurred after the Earth 1 flyby where several activities occurred on board

The Galileo spacecraft is a spin-stabilized spacecraft which rotates at about 3 rpm.

IV. Ground Data System Configuration

The DSN configuration was that of normal cruise tracking with the addition of the Radio Science subsystem to acquire open-loop data for certain passes. The radio metric data were sampled at a rate of 1/sec using a loop filter bandwidth of 10 Hz. For all of the passes, hydrogen masers were the frequency and timing references at the ground stations. The overall stability of the ground system (frequency reference, receiver, cables, etc.) is expected to be at about 1×10^{-15} at 1000 sec, which is far more stable than the USO.

The open-loop receiver utilized a 100-Hz bandwidth filter and a sampling rate of 200 per sec. Data were recorded on 1600-bpi or 6250-bpi Original Data Record (ODR) tapes which were delivered to the Galileo RST.

V. Data Products

The following data products are received for each pass assuming they were requested and no failures occurred: (1) an ATDF tape containing the closed-loop Doppler data, (2) an ODR tape containing the open-loop data records (when applicable), (3) a pass folder from the DSN containing copies of frequency predictions, operator logs and related material, and (4) a spacecraft trajectory vector file from the Galileo Navigation group. References on interface agreements and software interface specifications for the data products are available from the Radio Science System's Group Library. The data products are received, logged, validated (tapes only), and archived by the Radio Science Data Production System.

VI. Analysis Software and Techniques

The analysis was performed on the Radio Science Support System Radio Occultation Data Analysis Network (RODAN) computer which includes a Prime 4050 computer, a Floating Point Systems (FPS) array processor, two magnetic tape drives and other peripherals. The system is accessible by a set of IBM PS/2 terminals as well as two Sun workstations.

Although the analysis tools used for this work were inherited from the Voyager Project, there were enough differences in configuration and procedures in the Galileo Project that nontrivial modifications were made to

the software and the techniques. The analysis software was upgraded to estimate more accurately the spacecraft-transmitted frequency and the frequency residuals used for the stability analysis, including installation of code to model the effects introduced by a spinning spacecraft.

The data were processed by the STBLTY software program set. Figure 1 is a block diagram displaying the interconnection of the component programs with the various input data types, intermediate files, and output files. The functions of each of the component programs are described below.

PLLDEC: If open-loop data are to be processed, the phase-lock-loop program PLLDEC is utilized to perform signal detection on the digitized open-loop samples read from the input ODR or Intermediate Data Record (IDR) tapes. The detected frequencies and time tags are written to an output file (F36), and the open-loop receiver tuning (POCA) frequencies are written to a separate file (F33).

OCEP: In the case of closed-loop data, OCEP reads the Doppler counts and Doppler extractor reference frequency from an input ATDF tape, converts these to sky frequencies, and writes the time-tagged sky-frequency data and related information to the F50 disk file. In the case of open-loop data, OCEP reads the PLLDEC output files (F36, F33), converts to sky frequencies and writes the time tags, sky frequencies and other related information to the F50 disk file.

GETTRAJ: Reads the spacecraft state vectors from the Navigation Team-provided Celestial Reference Set (CRS) file, performs vector manipulations and light-time solutions, and writes a disk file (F45) containing the time-tagged spacecraft position and velocity vectors relative to the Sun and the observing DSN station. Details of the orbit-determination solutions determined by Galileo navigation used to generate the trajectory files are given elsewhere [4].

RESID: Reads the sky frequencies from the OCEP output file (F50), and the trajectory vectors from the GETTRAJ output file (F45). Model downlink sky frequencies are estimated from the trajectory vectors, and are corrected for troposphere, spacecraft spin, gravitational redshift and, if applicable, the spinning off-axis LGA-2-induced Doppler signature. RESID estimates the spacecraft-transmitted frequency at the time tag of the first data point, and computes the frequency residuals by differencing the observed sky frequencies from the estimated sky frequencies for each data point. The troposphere model is a simple zenith path delay translated

the spacecraft which introduced dynamic motion. Stator repositioning commands introduced dynamic motion on the spacecraft during the data acquisition period for pass 90-009. In these cases, the dynamic activity was assumed to be symmetric about the center of weight of the spacecraft and thus appeared not to have biased the estimated spacecraft-transmitted frequencies which are determined at the first data point, and then corrected to the center of weight of the full data span. The Allan deviations for these passes were, however, degraded.

C. Solar Interference Example

Figure 13 displays the residuals for pass 91-03-05 (91-064) where several solar disturbances occurred during that day including the period of the data acquisition. The degradations in the observed Doppler noise measurements from the closed-loop system are consistent with those expected from solar interference.¹ This was the only pass in this data set where this behavior was observed. The estimated spacecraft-transmitted frequency appears reasonable; however, the Allan deviations were degraded as expected.

D. Stability Analysis Results

The Allan deviation is the recommended measure of oscillator stability for time-domain signal processing. For all 82 passes, the Allan deviations are displayed in Figs. 14-17 for 1-, 10-, 100- and 1000-sec time intervals, respectively. Outlier points are annotated with the year and day number of the pass (YR-DOY). Note that outlier points are especially apparent for time intervals of 10, 100 and 1000 sec. These include (1) the three LGA-2 passes, 900109 (90-009), 901210 (90-345), and 910114 (91-014), where dynamic events occurred on the spacecraft, (2) the suspected solar interference pass, 910305 (91-064), (3) three passes, 900330 (90-089), 911019 (91-292) and 911130 (91-334), where the cause of the degradation is not known, but is possibly due to ground equipment problems, and (4) the first two passes after initial turn-on, 89-341 and 89-350, where the significant increasing frequency dynamics adversely dominated the Allan deviation measurements at 1000 sec (see Fig. 17).

When the outlier passes are removed, the resulting Allan deviation plots for the remaining passes at 1-, 10-, 100- and 1000-sec time intervals are displayed in Figs. 18-21, respectively.

¹ R. Woo, personal communication, Radio Frequency and Microwave Subsystems Section, Jet Propulsion Laboratory, Pasadena, California, August 1992.

The measured means and errors of the Allan deviations are given in Table 2 for each time interval (outlier passes were removed as well as three passes where the data were insufficient to estimate Allan deviations at 1000 sec). The pre-flight Allan deviation measurements of the Galileo USO for these time intervals are also presented in Table 2.² The pre-flight measurements were performed at the JPL Hydrogen Maser Test Facility on May 1, 1980. Figure 22 is a plot of the measured flight Allan deviation averages with the complete set of pre-flight measurements superimposed.

For 1 and 10 sec, the flight-measured Allan deviations significantly exceed the pre-flight Allan deviations. This was expected since it is known that wideband thermal system noise (white phase noise) dominates at these time intervals due to the low signal-to-noise ratios resulting from using Galileo's low-gain antennas. The flight-measured Allan deviations agree with estimates derived from the measured signal levels, system noise temperatures and receiver bandwidths. The Allan deviations at 1 sec (Fig. 18) and 10 sec (Fig. 19) also correlate with spacecraft range (compare with Fig. 23). If the HGA ever becomes available, and if there is an opportunity to turn off the telemetry modulation to increase the signal strength, there will be better visibility of the true USO performance at the 1-sec and 10-sec time intervals.

The flight-measured Allan deviations at 100 and 1000 sec agree well with the corresponding pre-flight values in Table 2 and Fig. 22. This implies that the flight data are dominated by the true behavior of the USO at these time intervals. The hump at 70 sec in the pre-flight values in Fig. 22 has been attributed to the thermal oscillation of the inner oven current driven by noise.³ The consistent behavior of the flight data between 100 and 1000 sec with the pre-flight measurements at these time intervals implies that this effect is visible in the flight data. In the absence of this thermal cycling, one would then expect the Allan deviation behavior to be flat over these time scales, and thus be consistent with flicker frequency noise. The expected noise due to unmodeled media fluctuations lies well below this level at these time intervals.

The Allan deviation behavior for the flight data measurements can be broken down into several regions. The

² A. Kirk, "Frequency Stability Measurements of Galileo Project High Stability Crystal Oscillators," JPL 331-TRAK-800527 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 1, 1980.

³ G. E. Wood, personal communication, Telecommunications Systems Section, Jet Propulsion Laboratory, Pasadena, California, February 1993.

Allan deviation in the 1- to 10-sec region falls roughly as τ^{-1} , characteristic of white phase noise dominating at these time intervals. The USO behavior here is masked by this noise at these time scales. The behavior of the region from 100 to 1000 sec can be interpreted as being consistent with the known inner-oven thermal cycling effect being superimposed on a flicker frequency noise floor (using the pre-flight measurement information). Between 1000 sec to about 40,000 sec, one can assume a continuation of the flicker noise floor upon inspection of the pre-flight measurements and error bars. The 17-mHz rms scatter of the estimated flight frequencies about an aging model (see Section VII.F) is consistent with the random walk inferred by extrapolating the pre-flight Allan deviation measurements as $\tau^{1/2}$ to a weekly time interval, but could include some unknown mismodeling.

E. Signal Levels

The USO passes have relatively low signal levels as exhibited in the plot of spacecraft signal levels (AGC) of Fig. 24. The observed values of the Allan deviations at small time intervals are consistent with those derived for thermal noise theory using measurements of system noise temperature and the measured signal levels presented in this plot. The AGC correlates well with spacecraft range (Fig. 23).

After correcting the received signal levels displayed in Fig. 24 for receiver station antenna gain, space loss, spacecraft LGA antenna gain, and telemetry state carrier suppression, the resulting spacecraft-transmitted 2.3-GHz power levels at the RFS/antenna subsystem interface of the spacecraft were computed. The 42.2-dBm mean value of the LGA-2 passes was in good agreement with the expected power level, while the 40.0-dBm mean value for the LGA-1 passes was about 2 dB lower than the expected power level.⁴ The resulting rms scatters of about 2 dB were consistent with the known calibration uncertainties at the ground stations.

F. Analysis of Spacecraft-Transmitted Frequency Measurements

The frequency transmitted by the spacecraft was estimated for all 82 passes. Figure 25 displays the USO-referenced spacecraft-transmitted frequencies as estimated by STBLTY. Each point on the plot is the USO frequency

estimated at the first time tag for that pass and then corrected to the center of weight of all of the residuals over that pass. The assigned uncertainties of the estimated frequencies run about 3 mHz and are dominated by the uncertainty inferred from not performing an ionospheric correction. The plot is annotated with the times the USO was powered on and off. The time axis is in days since January 1, 1989. The USO was initially powered on in flight on December 5, 1989 (DOY 339). There was one instance of cycling the USO off (91-217) and then back on (91-228) in this data set.

A preliminary pass of a few minutes' duration was conducted shortly after initial turn-on on December 5, 1989 (and after the inter-oven current was allowed to stabilize), in order to verify operation of the USO. Here the frequency was observed to be increasing at a very high rate and exhibited dynamic behavior due to early impurity migration and/or stress relief resulting in the very poor stability expected shortly after turn-on. The first valid USO pass was conducted on December 7, 1989, several hours after initial turn-on.

Changes in the USO frequency with time are referred to as resonator aging. In general, the principal causes of aging are stress relief in the mounting structure of the crystal unit, mass transfer to or from the resonator's surface due to adsorption or desorption of contamination, changes in the oscillator circuitry, and possibly changes in the quartz [5]. Aging effects seen in this data set likely include surface liberation of impurities, impurity migration across and within the crystal, and linear aging (diffusion). The significant positive logarithmic increase in frequency shortly after initial turn-on (see Fig. 25) is probably due to the liberation of contamination on the crystal resonator surfaces which were acquired during the long period of inoperation prior to launch.⁵ The USO was powered off for much of the time that the spacecraft was dormant on the ground prior to launch (1986-1989). The linear region is expected to be reached after the USO has been turned on for a sufficiently long enough period of time such that the only significant aging mechanism is diffusion. The curvature observed for the 6 passes conducted after the USO was cycled off and back on in August 1991 could be attributed to stress relief and migration of impurities.

During the first USO-on cycle (December 1989 to August 1991), 76 passes of USO data were acquired. Each pass consisted of about 2 hr of 1/sec sampled Doppler

⁴ A. Makovsky, *Galileo Orbiter Functional Requirements*, module *Galileo Orbiter Telecommunications Functional Requirement Document*, JPL Document 625-205, 3-300B (internal document), Jet Propulsion Laboratory, Pasadena, California, May 9, 1989.

⁵ R. L. Sydnor, personal communication, Communications Systems Research Section, Jet Propulsion Laboratory, Pasadena, California, September 3, 1992.

Allan deviation in the 1- to 10-sec region falls roughly as τ^{-1} , characteristic of white phase noise dominating at these time intervals. The USO behavior here is masked by this noise at these time scales. The behavior of the region from 100 to 1000 sec can be interpreted as being consistent with the known inner-oven thermal cycling effect being superimposed on a flicker frequency noise floor (using the pre-flight measurement information). Between 1000 sec to about 40,000 sec, one can assume a continuation of the flicker noise floor upon inspection of the pre-flight measurements and error bars. The 17-mHz rms scatter of the estimated flight frequencies about an aging model (see Section VII.F) is consistent with the random walk inferred by extrapolating the pre-flight Allan deviation measurements as $\tau^{1/2}$ to a weekly time interval, but could include some unknown mismodeling.

E. Signal Levels

The USO passes have relatively low signal levels as exhibited in the plot of spacecraft signal levels (AGC) of Fig. 24. The observed values of the Allan deviations at small time intervals are consistent with those derived for thermal noise theory using measurements of system noise temperature and the measured signal levels presented in this plot. The AGC correlates well with spacecraft range (Fig. 23).

After correcting the received signal levels displayed in Fig. 24 for receiver station antenna gain, space loss, spacecraft LGA antenna gain, and telemetry state carrier suppression, the resulting spacecraft-transmitted 2.3-GHz power levels at the RFS/antenna subsystem interface of the spacecraft were computed. The 42.2-dBm mean value of the LGA-2 passes was in good agreement with the expected power level, while the 40.0-dBm mean value for the LGA-1 passes was about 2 dB lower than the expected power level.⁴ The resulting rms scatters of about 2 dB were consistent with the known calibration uncertainties at the ground stations.

F. Analysis of Spacecraft-Transmitted Frequency Measurements

The frequency transmitted by the spacecraft was estimated for all 82 passes. Figure 25 displays the USO-referenced spacecraft-transmitted frequencies as estimated by STBLTY. Each point on the plot is the USO frequency

estimated at the first time tag for that pass and then corrected to the center of weight of all of the residuals over that pass. The assigned uncertainties of the estimated frequencies run about 3 mHz and are dominated by the uncertainty inferred from not performing an ionospheric correction. The plot is annotated with the times the USO was powered on and off. The time axis is in days since January 1, 1989. The USO was initially powered on in flight on December 5, 1989 (DOY 339). There was one instance of cycling the USO off (91-217) and then back on (91-228) in this data set.

A preliminary pass of a few minutes' duration was conducted shortly after initial turn-on on December 5, 1989 (and after the inter-oven current was allowed to stabilize), in order to verify operation of the USO. Here the frequency was observed to be increasing at a very high rate and exhibited dynamic behavior due to early impurity migration and/or stress relief resulting in the very poor stability expected shortly after turn-on. The first valid USO pass was conducted on December 7, 1989, several hours after initial turn-on.

Changes in the USO frequency with time are referred to as resonator aging. In general, the principal causes of aging are stress relief in the mounting structure of the crystal unit, mass transfer to or from the resonator's surface due to adsorption or desorption of contamination, changes in the oscillator circuitry, and possibly changes in the quartz [5]. Aging effects seen in this data set likely include surface liberation of impurities, impurity migration across and within the crystal, and linear aging (diffusion). The significant positive logarithmic increase in frequency shortly after initial turn-on (see Fig. 25) is probably due to the liberation of contamination on the crystal resonator surfaces which were acquired during the long period of inoperation prior to launch.⁵ The USO was powered off for much of the time that the spacecraft was dormant on the ground prior to launch (1986-1989). The linear region is expected to be reached after the USO has been turned on for a sufficiently long enough period of time such that the only significant aging mechanism is diffusion. The curvature observed for the 6 passes conducted after the USO was cycled off and back on in August 1991 could be attributed to stress relief and migration of impurities.

During the first USO-on cycle (December 1989 to August 1991), 76 passes of USO data were acquired. Each pass consisted of about 2 hr of 1/sec sampled Doppler

⁴ A. Makovsky, *Galileo Orbiter Functional Requirements*, module *Galileo Orbiter Telecommunications Functional Requirement Document*, JPL Document 625-205, 3-300B (internal document), Jet Propulsion Laboratory, Pasadena, California, May 9, 1989.

⁵ R. L. Sydnor, personal communication, Communications Systems Research Section, Jet Propulsion Laboratory, Pasadena, California, September 3, 1992.

known as retrace is defined as the nonrepeatability of the frequency versus temperature characteristic at a fixed temperature upon on-off cycling of the oscillator [5]. An example of retrace is the 12-Hz jump between passes on 91-154 and 91-247 (see Fig. 25). Between these passes, the USO and its oven were turned off for an 11-day period in August 1991. Several mechanisms which can cause retrace include strain changes, changes in the quartz, oscillator circuitry changes, contamination redistribution in the crystal enclosure and apparent hysteresis [5].

H. Assessment of Environmental Effects

The frequency of the USO can change due to variations in environmental parameters which include temperature, pressure, acceleration, magnetic field, and radiation. The crystal frequency is also dependent on the power level. Discussions are presented in [5-7] on the effects of the various phenomena on the behavior of crystal oscillator devices.

The Galileo USO was designed to minimize the effects of magnetic fields. A worst-case estimate of the magnetic field of the spacecraft in the environment of the USO found that the resulting fluctuations in frequency were expected to be negligible for the 10- to 1000-sec time intervals.⁸

Since the USO is oven-controlled so as to maintain a constant temperature, and the crystal temperature is designed to operate at an optimum point on the f versus T curve, noise due to temperature fluctuations is expected to be insignificant.

The majority of the USO passes were conducted during quiescent periods on the spacecraft when there was no

⁸ R. Postal, "A Concern of USO Stability as a Function of Magnetic Field," JPL IOM 3362-87,019 (internal document), Jet Propulsion Laboratory, Pasadena, California, June 3, 1987.

scheduled dynamic activity. The few known exceptions of dynamic activity involved no net thrusts on the spacecraft. Acceleration effects on the spacecraft USO are considered to be negligible.

It is assumed that in the deep space environment, changes due to atmospheric pressure and humidity are virtually nonexistent. Pre-launch testing showed that the Galileo USO exhibited spikes of less than 1 mHz during changes in pressure.⁹

Due to the high level of shielding, no charged particles are expected to hit the crystal during cruise. High-energy particles are expected to be stopped by the lead shielding, producing photons which could possibly hit the crystal.¹⁰ The level of radiation reaching the crystal during the cruise phase is expected to cause negligible shifts in frequency. The USO frequency could, however, shift about 1 Hz during passage through the Jovian radiation belts.¹¹

VIII. Conclusion

The Galileo USO appears to be healthy and functioning in a reasonable manner based on the analysis presented in this article. The evaluation of the Radio Science instrument will continue for the duration of the Galileo radio propagation investigations.

⁹ A. Gussner, "Summary of Galileo Ultra Stable Oscillator (USO) Testing," JPL IOM 3364-80-080 (internal document), Jet Propulsion Laboratory, Pasadena, California, August 20, 1980.

¹⁰ G. E. Wood, personal communication, Telecommunications Systems Section, Jet Propulsion Laboratory, Pasadena, California, February 1993.

¹¹ G. E. Wood, "Radiation Testing of Ultra Stable Oscillator S/N 004," JPL IOM 3396-76-095 (internal document), Jet Propulsion Laboratory, Pasadena, California, August 20, 1976.

Acknowledgments

P. Eshe, R. Herrera and T. Horton of the Radio Science Support Team provided valuable support. The comments and discussions provided by R. Sydnor and G. Wood, and R. Filler of LABCOM, New Jersey, were greatly appreciated. We acknowledge assistance by the Galileo Flight Team and the Galileo Navigation Team. The Deep Space Network generated the data. The authors thank W. Mayo of the Galileo Telecom Team for providing USO oven current and spacecraft temperatures.

References

- [1] T. P. Krisher, D. D. Morabito, and J. D. Anderson, "The Galileo Solar Redshift Experiment," *Physical Review Letters*, vol. 70, pp. 2213–2216, April 12, 1993.
- [2] H. T. Howard, V. R. Eshleman, D. P. Hinson, A. J. Kliore, G. F. Lindal, R. Woo, M. K. Bird, H. Volland, P. Edenhofer, M. Pätzold, and H. Porsche, "Galileo Radio Science Investigations," *Space Science Reviews*, vol. 60, pp. 565–590, 1992.
- [3] J. D. Anderson, J. W. Armstrong, J. K. Campbell, F. B. Estabrook, T. P. Krisher, and E. L. Lau, "Gravitation and Celestial Mechanics Investigations with Galileo," *Space Science Reviews*, vol. 60, pp. 591–610, 1992.
- [4] V. M. Pollmeier and P. H. Kallemeyn, "Galileo Orbit Determination from Launch through the First Earth Flyby," *Proceedings of the 47th Ann. Meeting of the Inst. of Nav.*, Williamsburg, Virginia, pp. 9–16, June 10–12, 1991.
- [5] J. R. Vig, *Introduction to Quartz Frequency Standards*, Research and Development Technical Report SLCET-TR-92-1 (Rev. 1), Army Research Laboratory, Electronics and Power Sources Directorate, Fort Monmouth, New Jersey, October 1992.
- [6] D. P. Howe, *Frequency Domain Stability Measurement*, National Bureau of Standards, Technical Note 679, U.S. Department of Commerce, PB-252-171, Washington, D.C., March 1976.
- [7] B. Parzen, *Design of Crystal and Other Harmonic Oscillators*, New York: John Wiley and Sons, 1983.

Table 1. Galileo USO pass summary.

Year	DOY	Start hr:min:sec	End hr:min:sec	DSS id.	AGC, dBm	(Spacecraft-Transmitted Frequency - 2294997000), Hz
89	341	21:34:22	23:27:59	14	-149.5	690.321
89	350	00:02:40	01:57:59	14	-142.4	696.287
89	360	22:35:42	00:00:59	14	-145.5	699.954
90	2	17:09:45	18:59:59	14	-149.0	701.425
90	9	17:37:22	19:26:08	14	-149.6	702.669
90	15	17:34:22	19:28:11	14	-153.1	703.597
90	19	16:10:46	17:57:50	14	-156.1	704.134
90	28	16:00:09	18:00:00	14	-154.3	705.128
90	32	15:33:51	17:29:59	14	-155.8	705.508
90	37	10:50:57	12:12:36	63	-155.0	705.895
90	44	03:00:00	05:19:59	43	-151.6	706.383
90	46	00:19:29	01:04:57	43	-154.3	706.523
90	49	19:44:31	21:22:29	43	-153.9	706.855
90	56	02:10:12	03:26:48	43	-157.4	707.295
90	58	23:08:28	01:01:01	43	-158.5	707.457
90	61	00:17:51	01:55:53	43	-162.2	707.586
90	68	02:12:12	03:42:02	43	-159.5	707.987
90	76	02:19:06	03:28:03	43	-165.0	708.376
90	78	00:41:32	02:27:47	43	-166.5	708.452
90	83	19:13:51	20:56:11	43	-167.5	708.708
90	89	00:09:38	02:03:16	43	-166.8	708.888
90	97	22:18:28	23:55:22	43	-165.3	709.205
90	104	22:14:52	23:56:39	43	-167.6	709.422
90	110	00:47:05	02:26:45	43	-168.0	709.570
90	113	20:46:51	22:01:00	43	-169.0	709.671
90	121	23:21:02	00:57:56	43	-166.5	709.875
90	128	09:26:40	11:04:42	63	-168.3	710.014
90	136	23:22:03	01:16:30	43	-167.0	710.111
90	139	19:55:00	21:48:51	43	-166.9	710.166
90	149	19:17:45	21:02:08	43	-165.0	710.337
90	155	19:18:09	19:45:01	43	-165.7	710.395
90	162	18:14:46	19:43:47	14	-166.6	710.457
90	172	15:18:15	17:04:29	14	-164.8	710.566
90	176	16:19:35	18:05:46	14	-165.7	710.587
90	183	18:20:15	20:07:25	43	-165.6	710.634
90	193	14:48:37	16:34:51	14	-162.8	710.704
90	197	18:19:45	20:02:33	43	-165.0	710.720
90	206	21:17:32	23:00:05	43	-165.3	710.737
90	213	21:12:49	22:03:33	43	-165.6	710.769
90	221	20:43:23	22:29:37	43	-163.7	710.778
90	228	18:14:37	19:59:44	43	-162.8	710.775
90	233	15:17:38	17:03:53	14	-168.1	710.785
90	243	19:44:31	21:24:40	43	-165.5	710.764
90	250	01:41:02	02:45:10	63	-162.7	710.757
90	253	19:13:26	20:55:34	43	-161.1	710.735

Table 1. (cont'd).

Year	DOY	Start hr:min:sec	End hr:min:sec	DSS id.	AGC, dBm	(Spacecraft-Transmitted Frequency - 2294997000), Hz
90	262	19:26:55	21:11:16	43	-160.5	710.731
90	268	00:30:00	03:40:00	63	-162.5	710.706
90	274	18:56:25	20:43:34	43	-160.1	710.658
90	281	17:12:34	18:54:36	43	-156.9	710.612
90	303	01:44:22	03:23:19	63	-153.0	710.495
90	312	01:16:55	02:57:56	63	-148.5	710.432
90	321	03:18:28	04:58:11	63	-145.1	710.326
90	330	01:13:20	02:57:56	63	-145.0	710.271
90	345	02:14:52	03:42:02	42	-138.4	710.169
90	350	08:18:28	10:12:02	61	-147.7	710.115
90	360	17:09:14	18:57:41	43	-150.3	709.991
91	4	07:32:13	07:49:48	63	-149.4	709.922
91	6	18:01:47	19:58:45	43	-149.1	709.885
91	14	16:55:11	18:46:30	43	-152.5	709.792
91	16	17:04:22	18:59:59	43	-151.3	709.769
91	19	16:05:14	17:59:59	43	-152.5	709.743
91	21	18:04:52	19:58:57	43	-153.1	709.709
91	26	16:07:05	17:50:51	43	-154.8	709.657
91	29	06:07:42	07:58:54	63	-156.1	709.638
91	33	15:58:12	17:43:56	43	-158.1	709.577
91	36	06:07:11	07:58:50	63	-162.6	709.548
91	39	06:05:08	08:01:13	63	-160.9	709.507
91	43	15:05:08	16:59:59	43	-162.2	709.459
91	49	21:21:14	23:13:13	43	-160.6	709.370
91	55	15:35:23	17:30:30	43	-160.7	709.300
91	64	14:05:14	16:02:30	43	-158.6	709.173
91	73	16:23:17	18:13:59	43	-160.2	709.068
91	81	15:08:37	17:02:30	43	-155.4	708.929
91	109	19:02:38	20:45:14	43	-155.4	708.563
91	141	00:02:40	01:59:52	63	-157.6	708.161
91	154	23:04:06	00:59:41	61	-164.6	707.997
91	247	16:39:14	18:23:50	63	-167.5	696.304
91	259	21:28:12	23:11:46	14	-164.8	696.556
91	275	00:59:23	02:43:34	43	-165.9	696.678
91	292	15:15:11	17:00:01	63	-168.2	696.728
91	318	03:54:25	05:29:56	43	-168.1	696.772
91	334	20:41:17	22:10:45	14	-165.9	696.769

Table 2. Summary of Galileo USO Allan deviations.

Time interval, sec	Measured flight $\times 10^{12}$	Measured pre-launch $\times 10^{12}$
1	29.4 ± 1.1	0.56
10	3.93 ± 0.17	0.56
100	0.90 ± 0.03	1.1
1000	0.71 ± 0.03	0.68

Note: Uncertainties are errors in the mean.

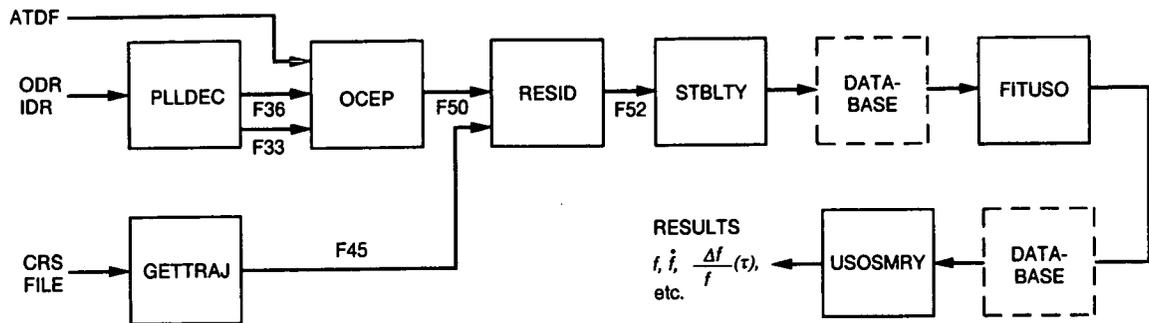


Fig. 1. STBLTY program-set block diagram.

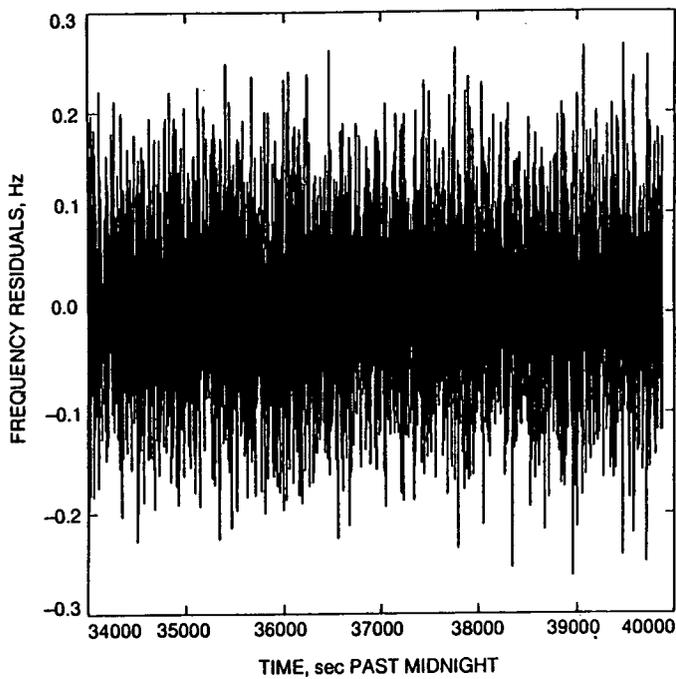


Fig. 2. Frequency residuals of sampled 1/sec Doppler for the USO pass of May 8, 1990.

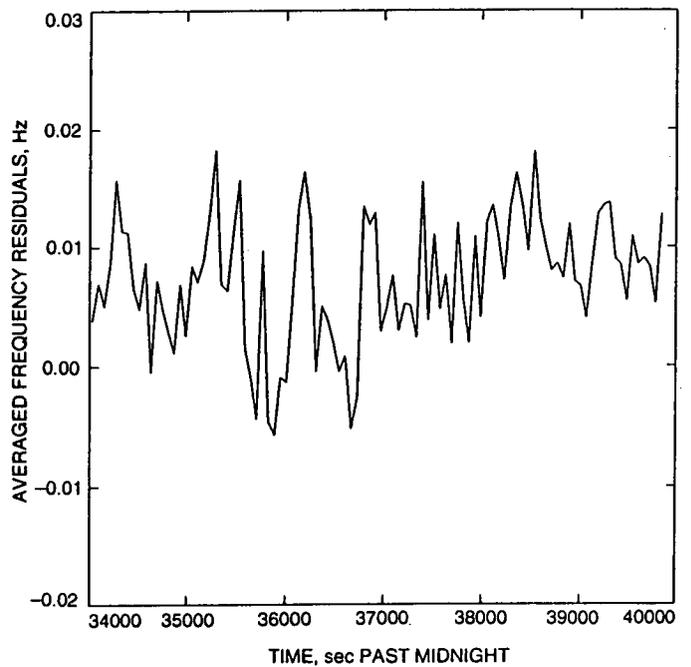


Fig. 3. Frequency residuals averaged every 60 sec for the USO pass of May 8, 1990.

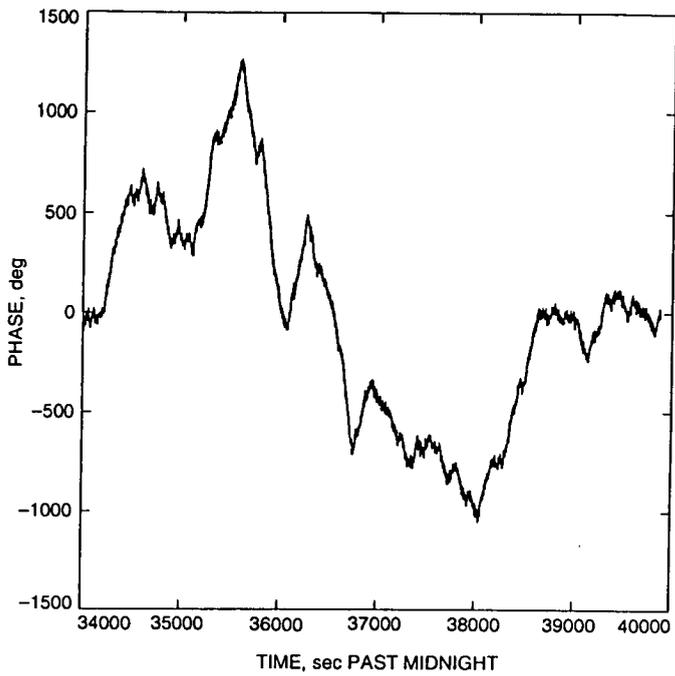


Fig. 4. Phase reconstructed from frequency residuals for the USO pass of May 8, 1990.

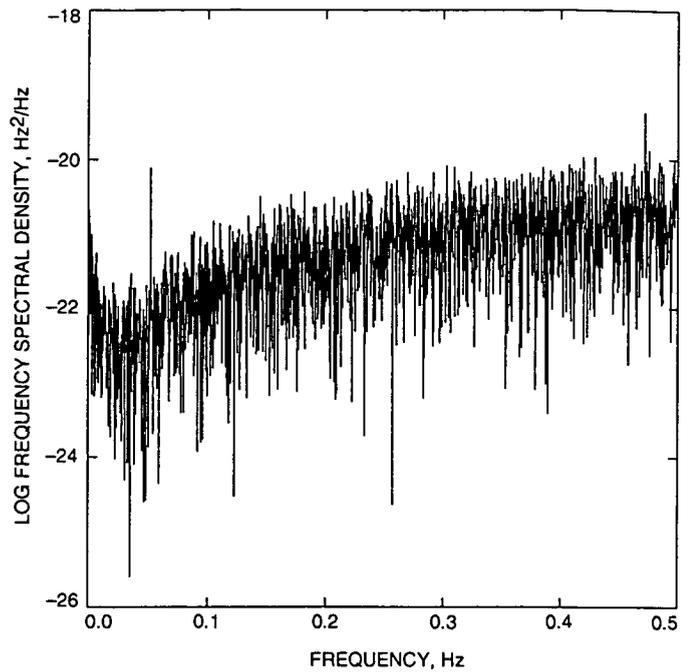


Fig. 6. Log of frequency spectral density of frequency residuals for the USO pass of May 8, 1990.

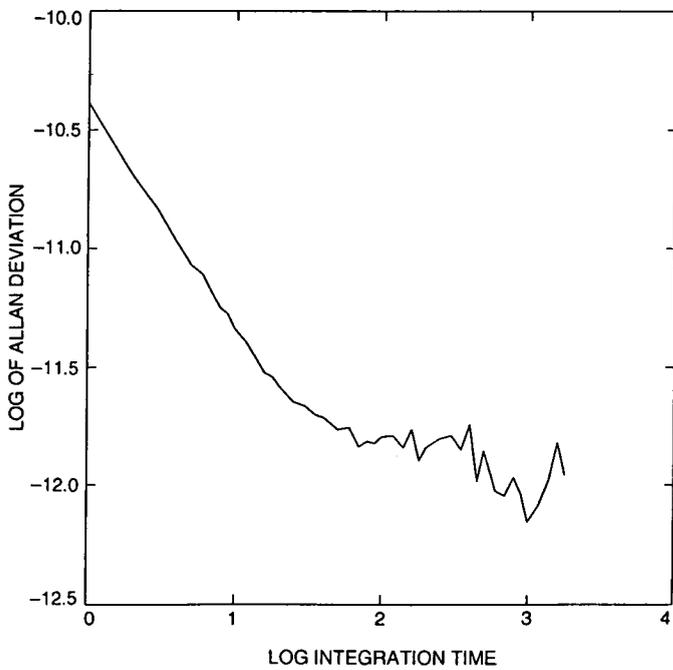


Fig. 5. Log of Allan deviation of frequency residuals for the USO pass of May 8, 1990.

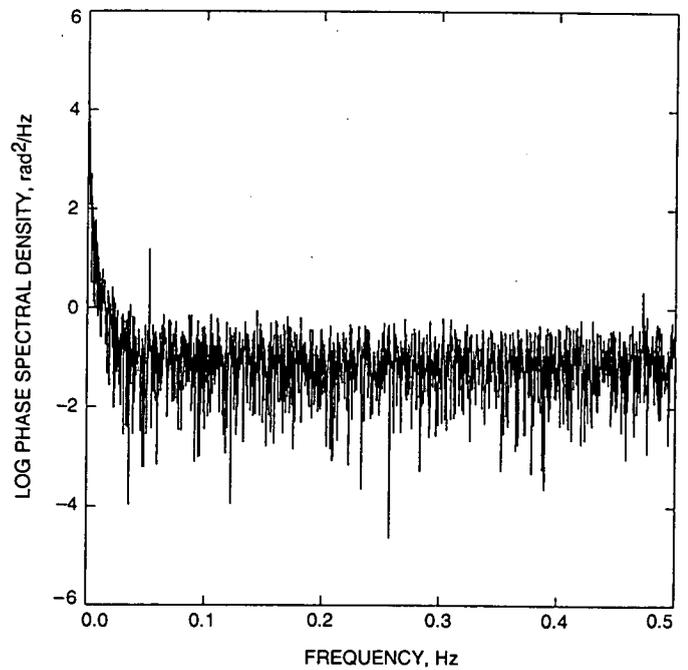


Fig. 7. Log of phase spectral density of frequency residuals for the USO pass of May 8, 1990.

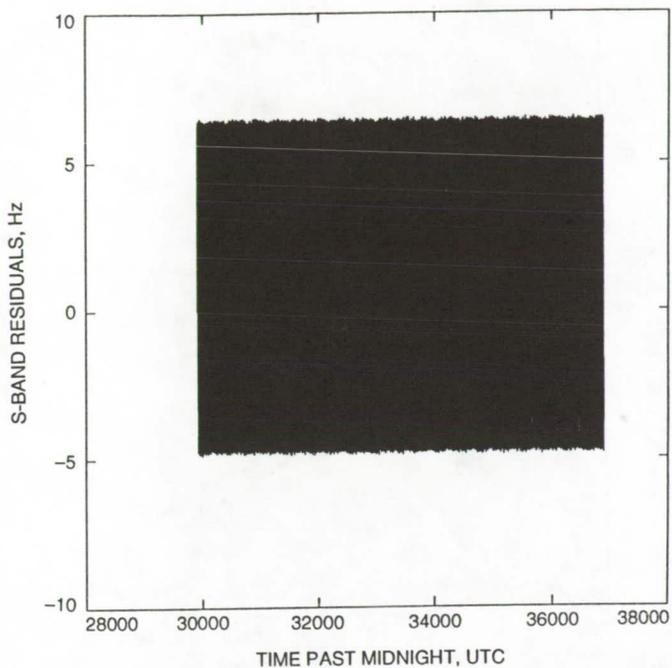


Fig. 8. Frequency residuals of sampled 1/sec Doppler for USO pass of December 16, 1990, where LGA-2 was the spacecraft antenna.

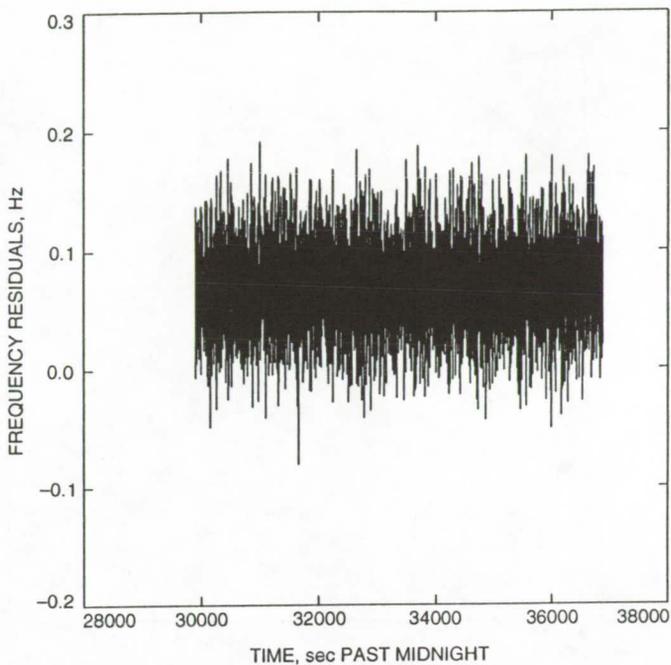


Fig. 10. Frequency residuals after removing the sinusoid fit from residuals displayed in Fig. 8 for USO pass of December 16, 1990, where LGA-2 was the spacecraft antenna.

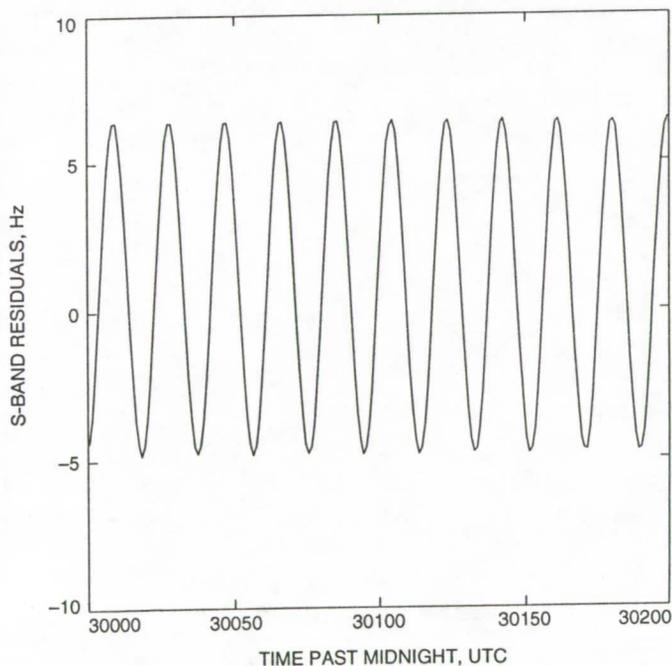


Fig. 9. Selected 200-sec period of frequency residuals of sampled 1/sec Doppler for USO pass of December 16, 1990, where LGA-2 was the spacecraft antenna.

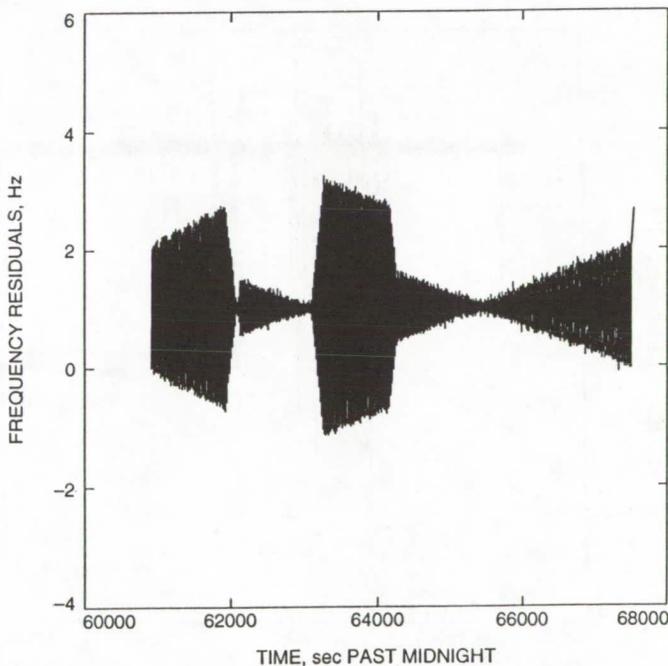


Fig. 11. Frequency residuals of sampled 1/sec Doppler for the USO pass of January 14, 1991, where LGA-2 was the signal source and dynamic motion occurred on board the spacecraft.

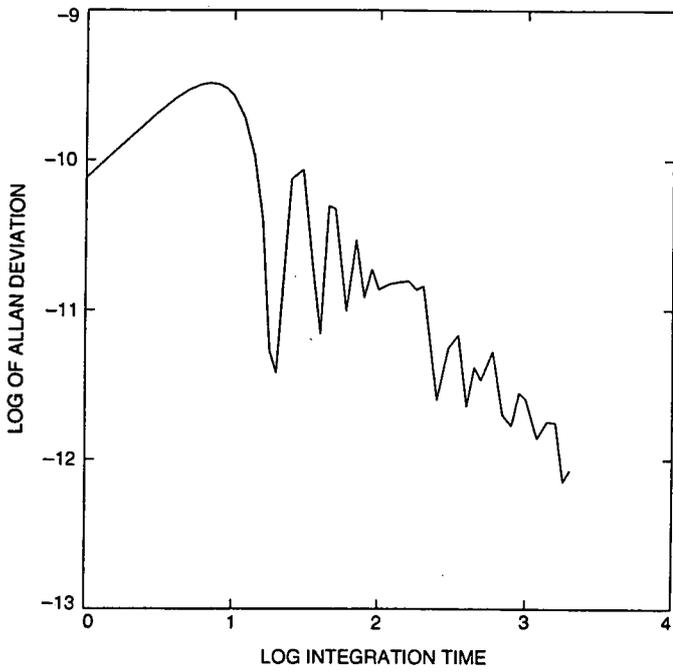


Fig. 12. Log of Allan deviation of frequency residuals of Fig. 11.

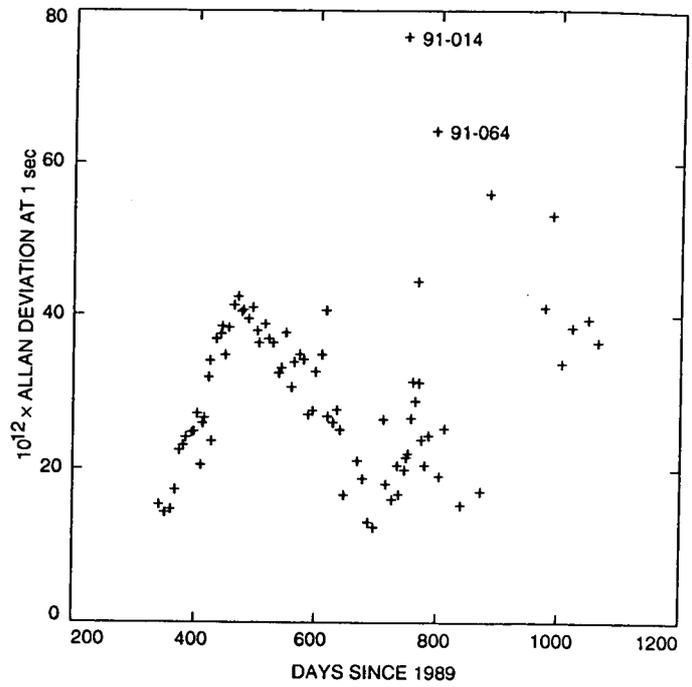


Fig. 14. Allan deviations at 1 sec for the 82 USO passes.

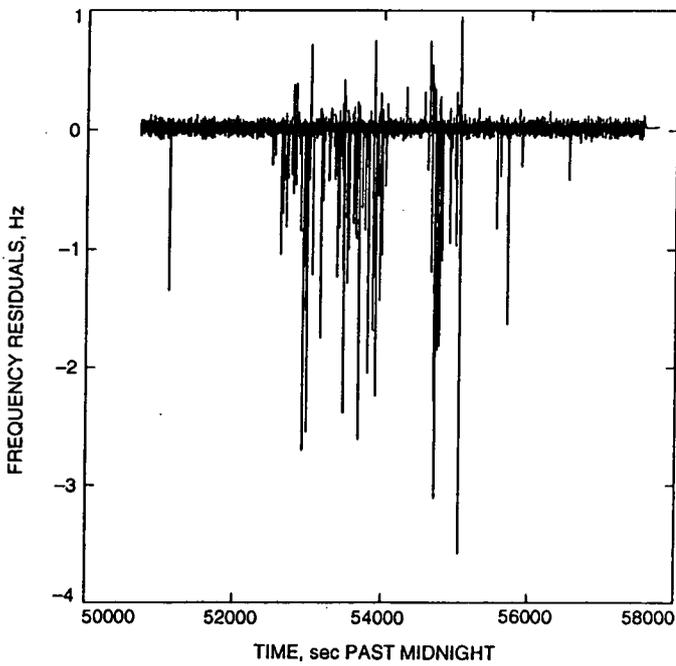


Fig. 13. Frequency residuals of the USO pass of March 5, 1991, where solar activity was known to have occurred.

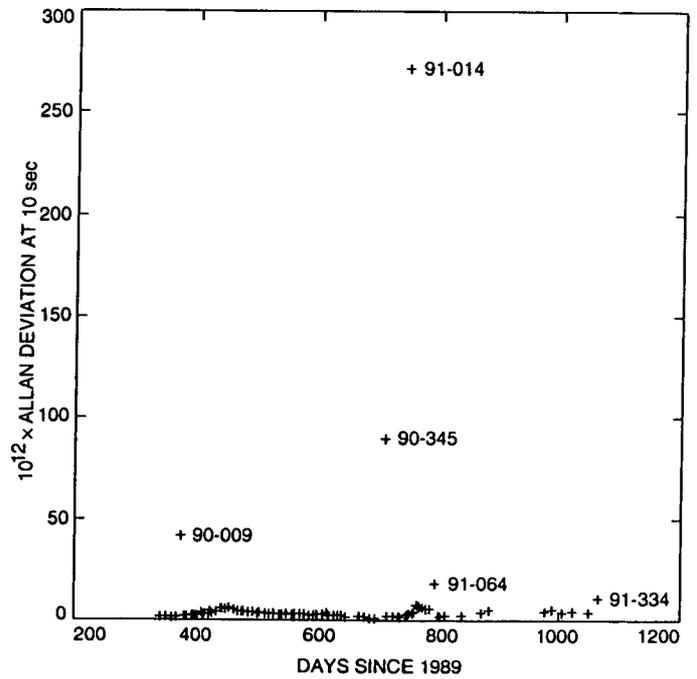


Fig. 15. Allan deviations at 10 sec for the 82 USO passes.

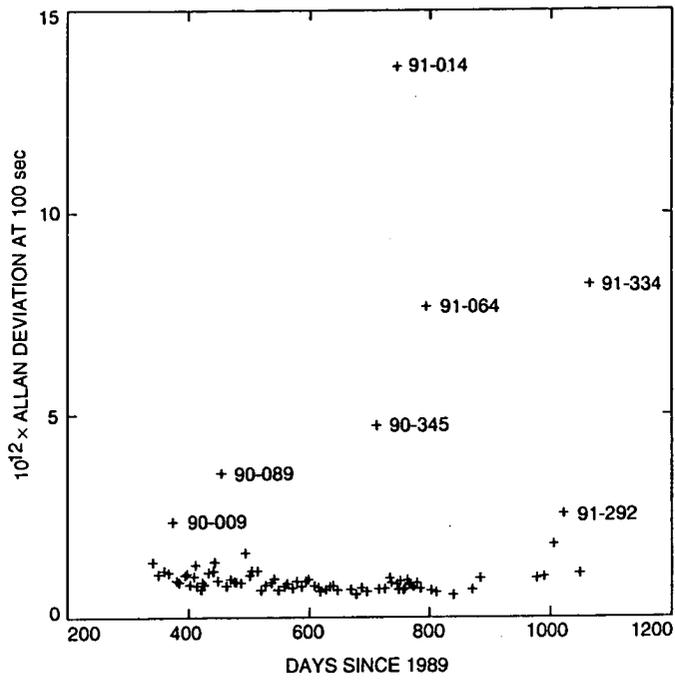


Fig. 16. Allan deviations at 100 sec for the 82 USO passes.

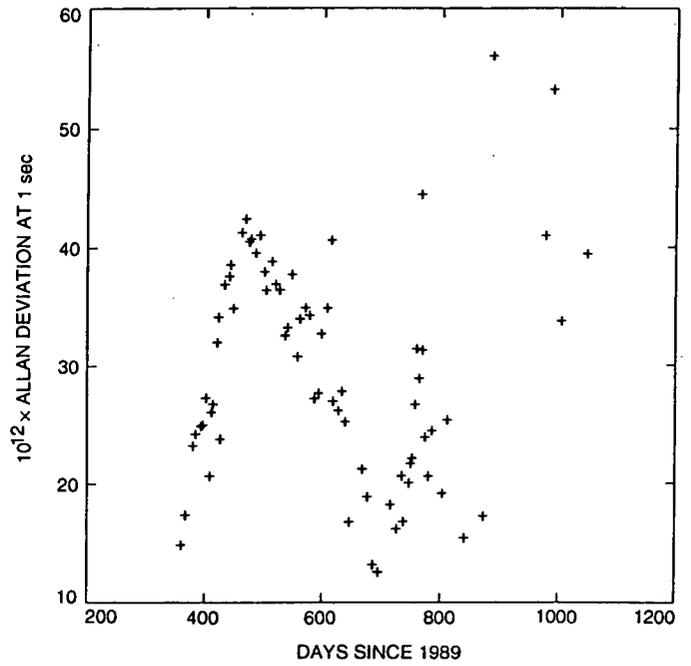


Fig. 18. Allan deviations at 1 sec for 73 USO passes (outliers identified in Figs. 14-17 have been removed).

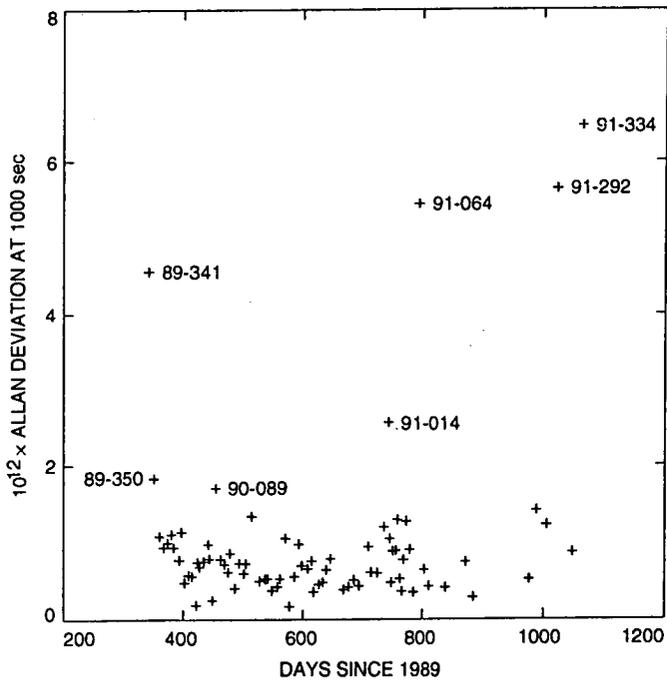


Fig. 17. Allan deviations at 1000 sec for the 79 USO passes.

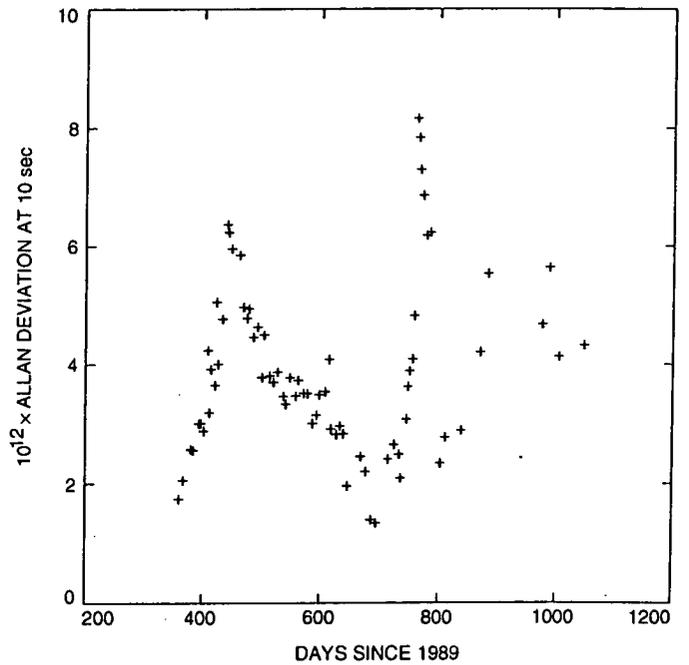


Fig. 19. Allan deviations at 10 sec for 73 USO passes (outliers identified in Figs. 14-17 have been removed).

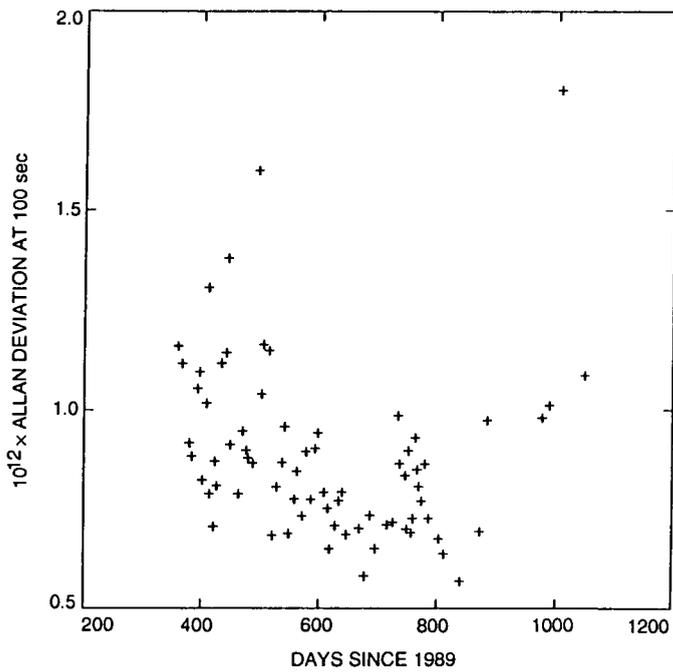


Fig. 20. Allan deviations at 100 sec for 73 USO passes (outliers identified in Figs. 14-17 have been removed).

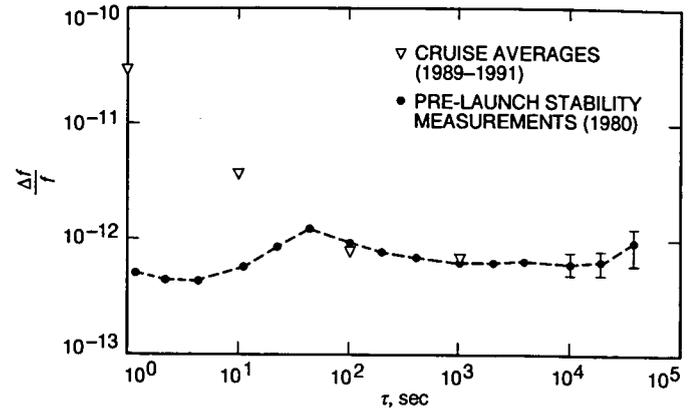


Fig. 22. In-flight USO-pass Allan-deviation measurement averages for 1, 10, 100 and 1000 sec superimposed with pre-flight Allan deviation measurements.

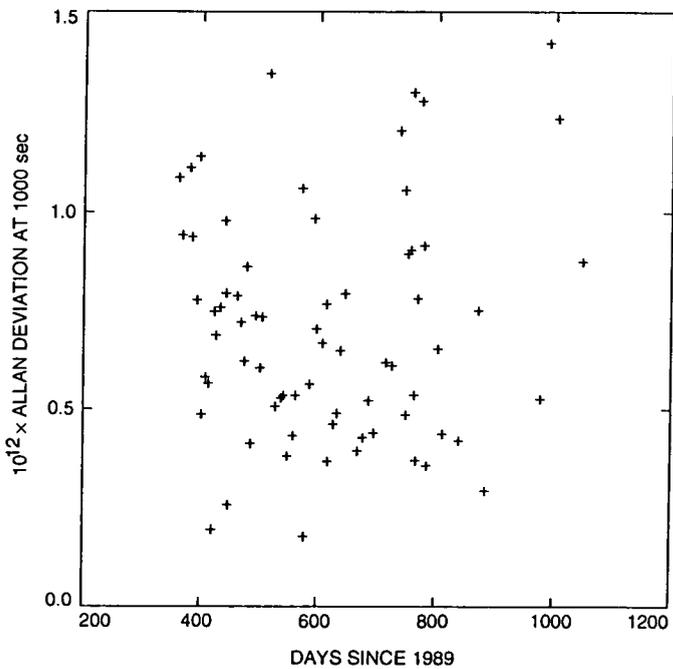


Fig. 21. Allan deviations at 1000 sec for 70 USO passes (outliers identified in Figs. 14-17 have been removed).

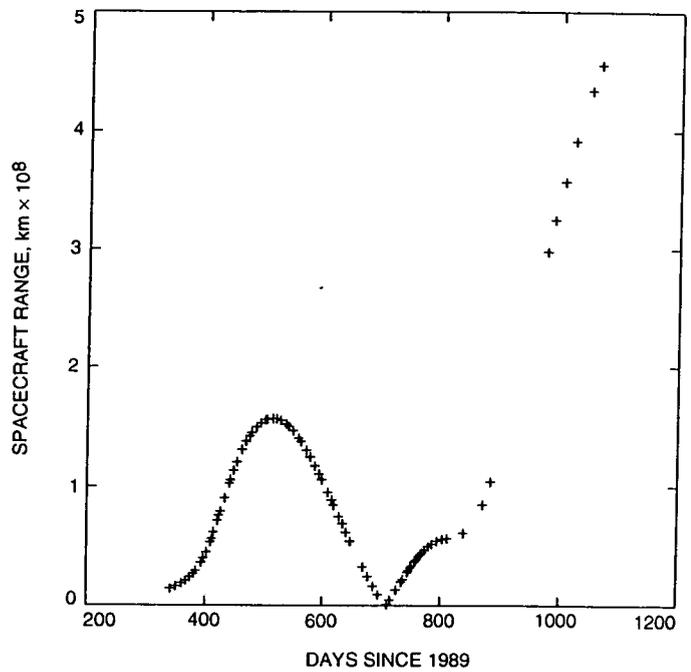


Fig. 23. Spacecraft range for each USO pass. Launch occurred at day 291; the dip at 707 days after 1989.0 was the Earth 1 flyby.

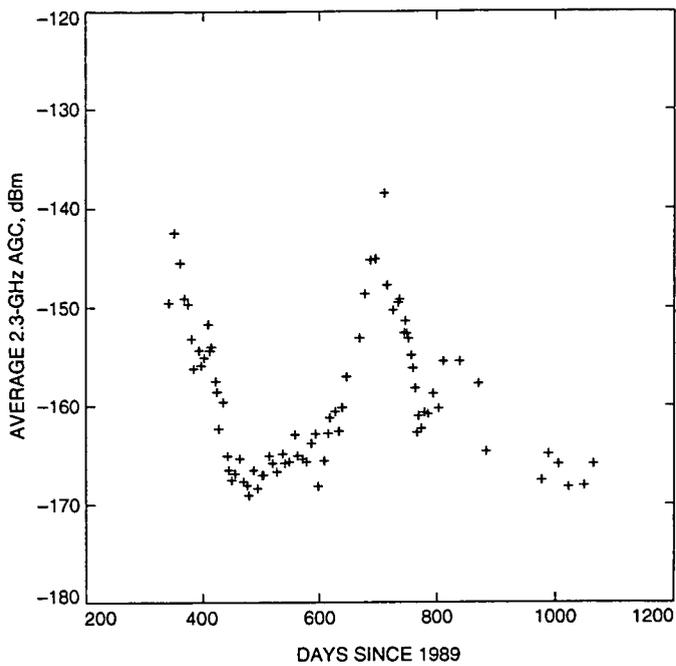


Fig. 24. Signal strength (AGC level) measured for each USO pass.

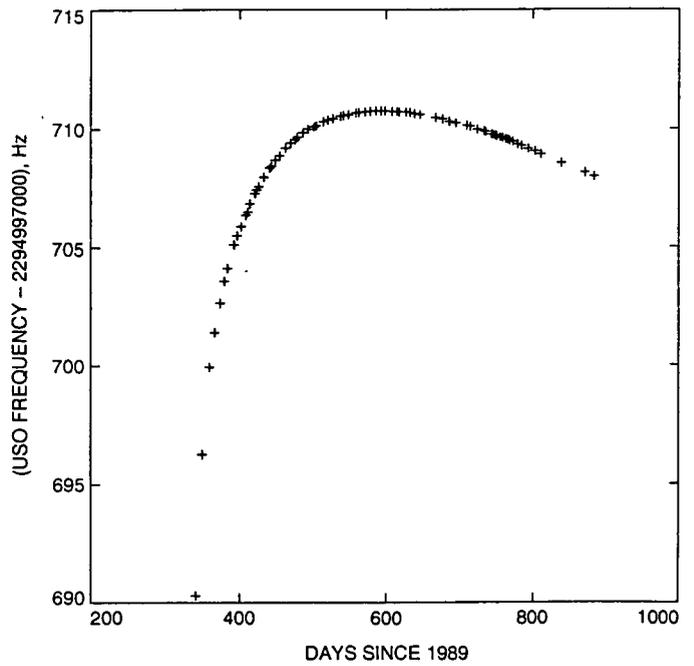


Fig. 26. Estimated spacecraft transmitted frequencies for 76 USO passes during the first USO on-off cycle after launch.

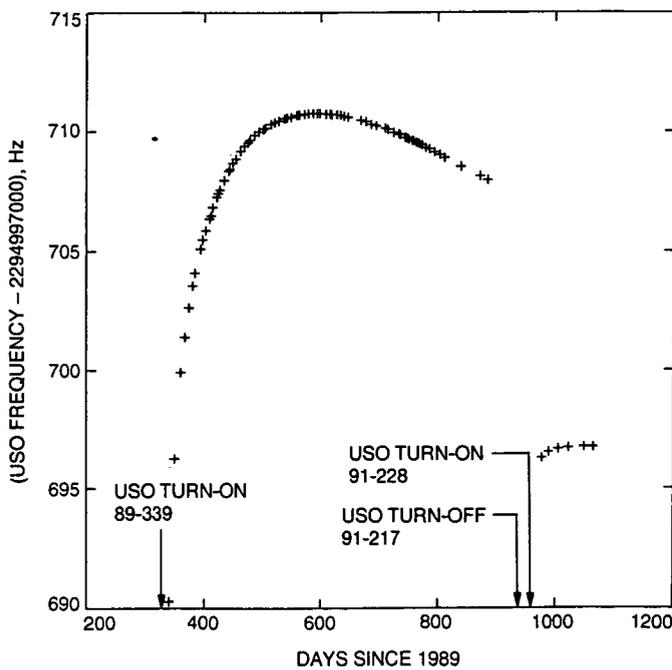


Fig. 25. Estimated spacecraft transmitted frequencies for all 82 USO passes as determined by STBLTY.

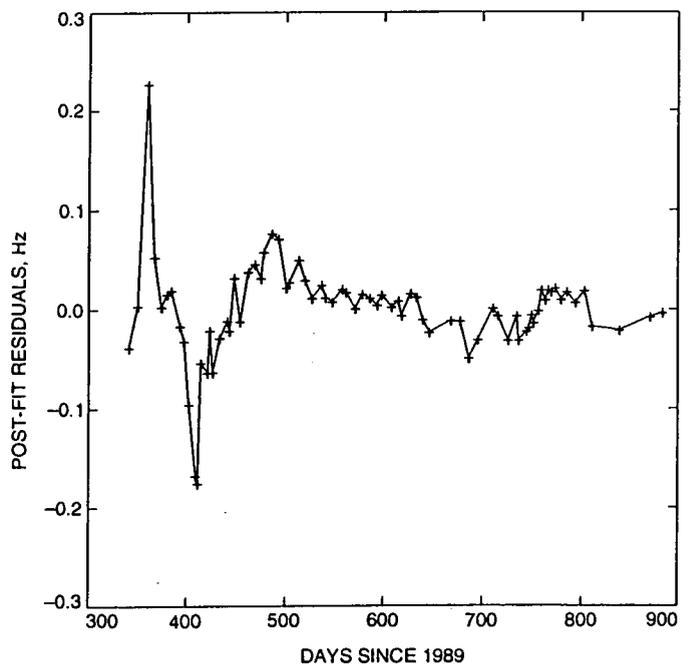


Fig. 27. Post-fit residuals of estimated spacecraft transmitted frequencies of Fig. 26 after fitting and removing an aging model.

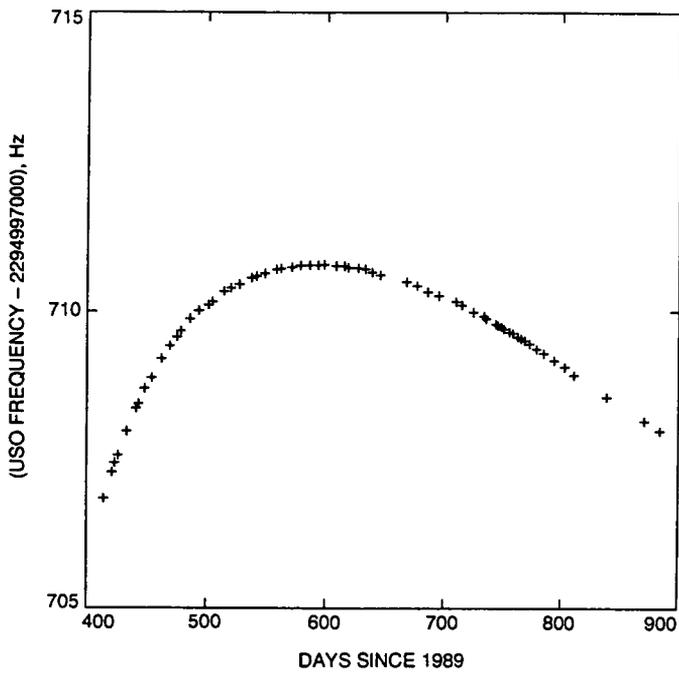


Fig. 28. Estimated spacecraft transmitted frequencies for 64 USO passes during the first USO on-off cycle after launch (the first 12 passes of Fig. 26 have been removed).

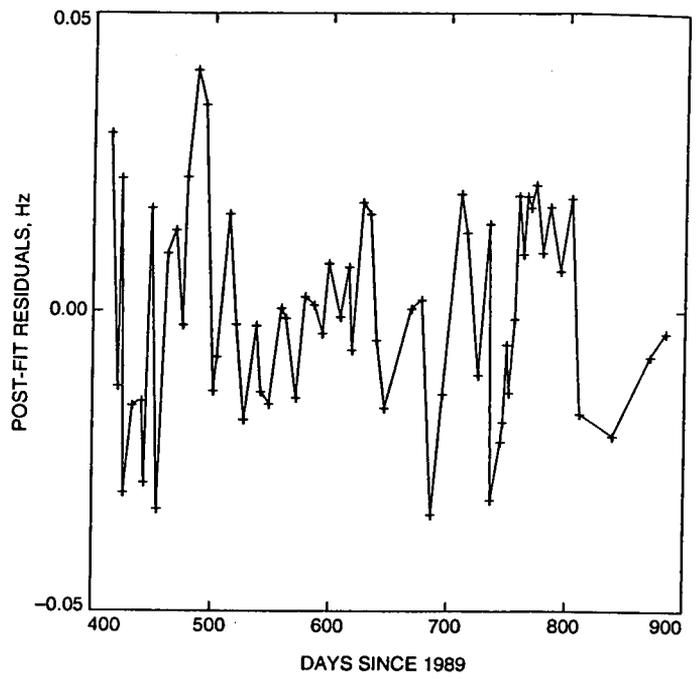


Fig. 29. Post-fit residuals of estimated spacecraft transmitted frequencies of Fig. 28 after fitting and removing an aging model.

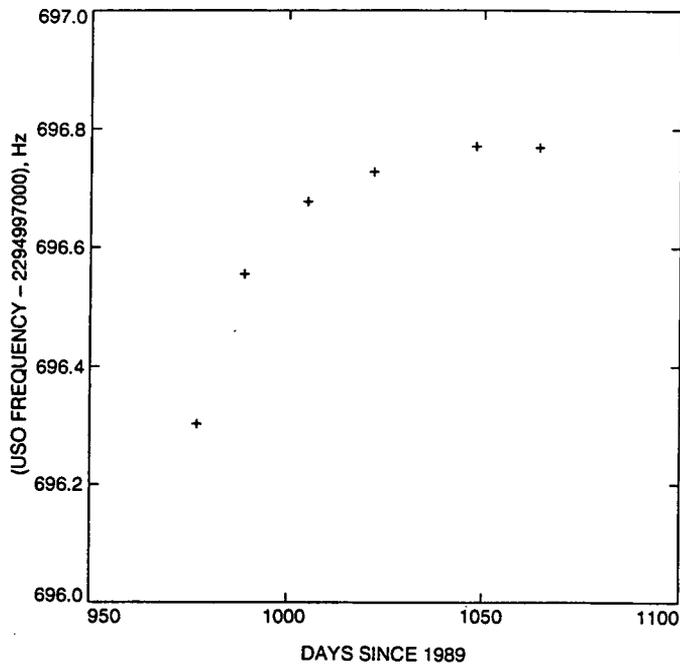


Fig. 30. Estimated spacecraft transmitted frequencies of 6 USO passes conducted after the first USO on-off cycle.