NASA-CR-194388

The Telecommunications and Data Acquisition Progress Report 42-114

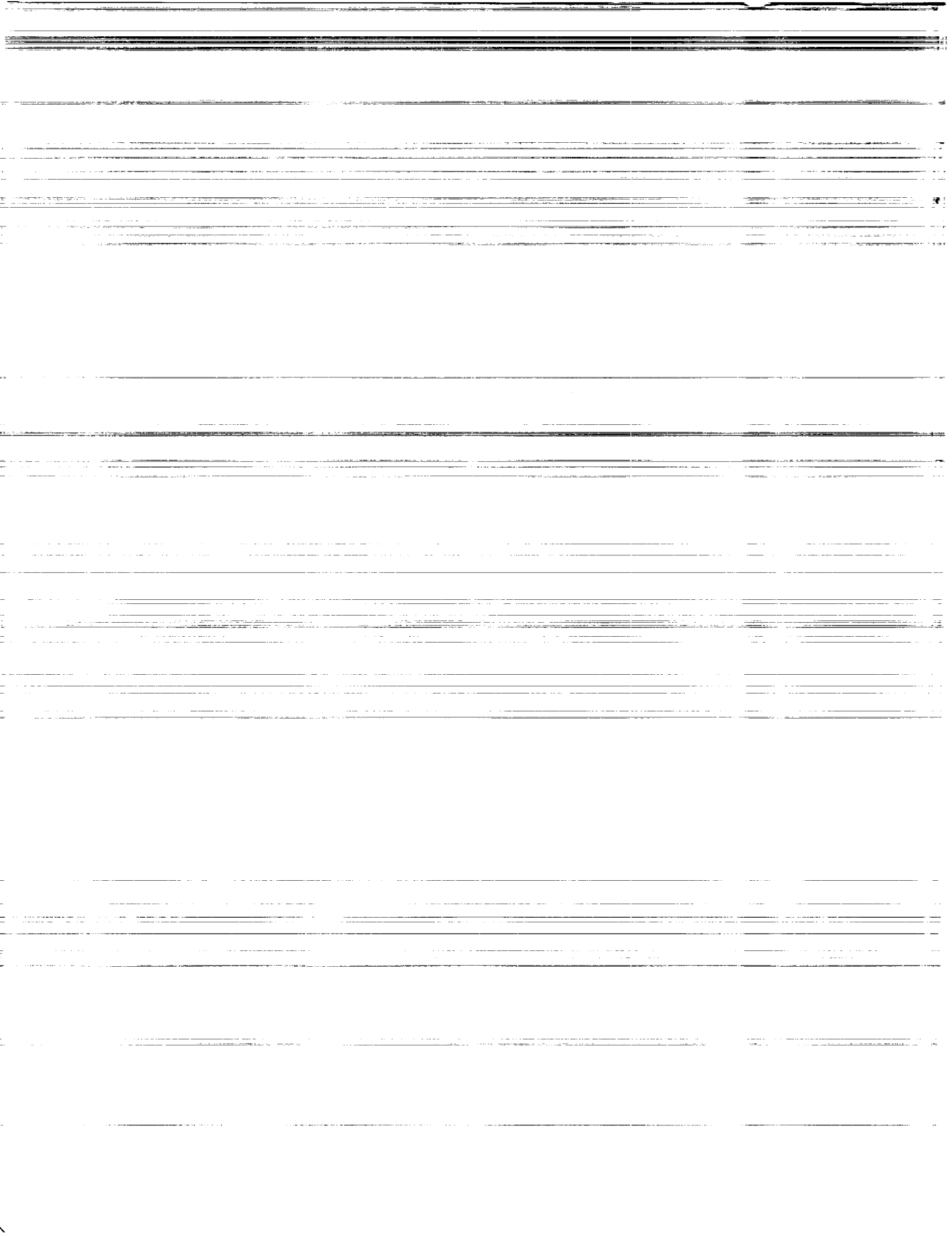
April–June 1993

In Memory of
E. C. Posner
Editor

August 15, 1993

NASA
National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California



The Telecommunications and Data Acquisition Progress Report 42-114

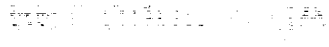
April–June 1993

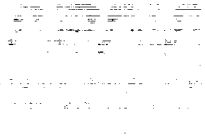
In Memory of
E. C. Posner
Editor

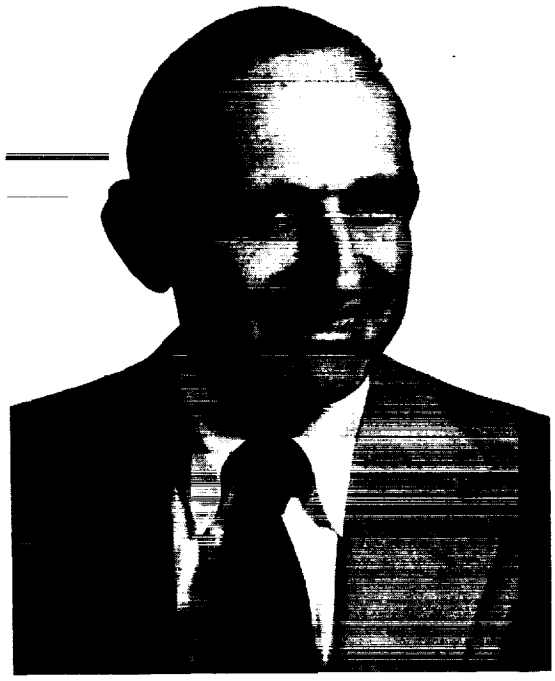
August 15, 1993

NASA
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California







Dr. Edward C. Posner
August 10, 1933, to June 15, 1993

Ed Posner was born in New York City on August 10, 1933, and graduated first in his class from Stuyvesant High School in 1950. Two years later, he received his B.A. in physics, and then his M.S. and Ph.D. (in 1957) from the University of Chicago. While a graduate student, he married Sylvia Cohen. They have a daughter, Joyce, and a son, Stephen. After four years of teaching mathematics at the University of Wisconsin and Harvey Mudd College, he joined JPL's Communications Systems Research Section in 1961. During the next 32 years, he worked extensively in information theory, communications, and neural networks. He was Supervisor of the Information Processing Group (1962–1967), Deputy Manager of the Communications Systems Research Section (1967–1973), a member of the TDA Planning Office (1973–1978), and Manager of the TDA Planning Office (1978–1983) before being appointed the TDA's Chief Technologist in 1981. In 1982, he became editor of the *TDA Progress Report*.

In addition to his JPL activities, Ed held a joint appointment on Caltech's faculty from 1970 until his death. In this capacity, he supervised numerous Ph.D. students, in research topics covering a variety of communications subjects that generally differed from his JPL communications interests. He founded and, for more than 20 years, organized the Caltech Electrical Engineering Systems Group seminar. He was a founder of Caltech's and JPL's program in neural networks. He sponsored many Summer Undergraduate Research Fellowship (SURF) students and cofounded the SURF satellite program. He also served on the freshman admissions committee.

Ed was a fellow of the Institute of Electrical and Electronics Engineers and a member of the Board of Governors of its Information Theory Society and of the Neural Networks Committee. He served as chairman of the Neural Information Processing Systems Foundation. He was a member of the committees of professional mathematics and engineering societies, program chairman or general chairman of several conferences, and a member of the editorial boards of four journals.

Dedication

This issue is dedicated to the memory of Dr. Edward C. Posner, who died in a tragic traffic accident while bicycling to work on June 15, 1993. He will be greatly missed.

Dr. Posner was the editor of this publication from December 1982 to June 1993. During his tenure, he succeeded in continually improving the technical quality and integrity of this quarterly report. It is widely circulated to 97 organizations in 19 countries.

The editing of this issue had been substantially completed by Dr. Posner. I am honored to have the privilege to complete this issue for him.

To express some of the feelings of those of us who had come to know Ed Posner, two passages follow from Dr. Lew Allen, Director of the Jet Propulsion Laboratory from 1982 to 1990, and Dr. Eberhardt Rechtin, a founder of the Deep Space Network.

Joseph H. Yuen

It was my great pleasure to work closely with Dr. Edward C. Posner in several respects. His contributions to the Deep Space Network have been recounted elsewhere and they, of course, continued during my time as Director. Ed was known nationally and internationally as a leading expert in communications with very low-level signals in a noisy environment. He applied this knowledge to several important problems involving national security. His contributions to these areas were recognized by the customer as very innovative and valuable.

Ed was an excellent teacher. He held a joint appointment with the Electrical Engineering Department at Caltech. He enjoyed working with students and was especially active in creating summer research opportunities for undergraduate Caltech students. His wit and wry view of many matters caused him to be a joy to students and fellow workers. He encouraged and mentored young scientists and engineers at JPL in a way that will be particularly missed.

Ed and I worked together on a few civil applications of communication technologies; here also, his deep understanding of underlying principles was very important. And, I also knew him as a friend. JPL will miss the services of Ed, and I and many others will miss him as a friend.

Lew Allen

The Deep Space Network is a very human organization, for all of its stations, accomplishments, "firsts," and fame. It has a soul. Ed Posner was at the center of that soul—a renaissance man, a teacher, and a brilliant technologist. He understood the importance of unmatched excellence under stress, of performance so high as to be taken for granted, and of deep dedication beyond self. But, without relief, such things make for somber surroundings and, worse yet, somber people who take things too seriously and who fail hard. Ed was a master of the gentle jibe that reminded us that being in the DSN could be—must be—fun. He had the technical and intellectual credentials to get away with it and the tact not to jest too deep.

The *TDA Progress Report* has been an essential part of the DSN from its inception. The report is an expression of how and why the DSN performs as it does—with high standards, integrity, and respect for those it serves. Excellent journals come from excellent organizations, of course. But, it is also true that excellent organizations result from the discipline, responsibility, and pride that comes from writing of one's work in a first-rate publication. Our Progress Report editor is responsible for making and keeping it so, for "keeping the flame," and revealing the soul of the DSN.

Ed Posner understood all that. Being an editor is a tough job. There's not much glory in it. He volunteered for it, among all his other careers. He seldom talked about his feelings, but I suspect he deeply loved the Deep Space Network and its people. As editor, he told their story. He was a lively, gentle, understanding, and brilliant part of their soul. We will miss him more than he could have imagined.

Eb Rechtin

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

OSS Tasks:
- Radio Science
- High-Resolution Microwave Survey
- Goldstone Solar System Radar
- Radio Astronomy

Discretionary Funded Task

Contents

OSC TASKS
DSN Advanced Systems
TRACKING AND GROUND-BASED NAVIGATION

COMMUNICATIONS, SPACECRAFT–GROUND

DSN Systems Implementation
CAPABILITIES FOR EXISTING PROJECTS

NETWORK UPGRADE AND SUSTAINING

DSN Operations

TDA PROGRAM MANAGEMENT AND ANALYSIS

Ground Communications Implementation and Operations

Galileo Optical Experiment

FLIGHT–GROUND ADVANCED ENGINEERING

Long-Range Program Planning

OSS TASKS
High-Resolution Microwave Survey

Radio Science

Radio Metric Errors Due to Mismatch and Offset Between a DSN Antenna Beam and the Beam of a Troposphere Calibration Instrument

R. P. Linfield and J. Z. Wilcox
Tracking Systems and Applications Section

Two components of the error of a troposphere calibration measurement have been quantified by theoretical calculations. The first component is a beam mismatch error, which occurs when the calibration instrument senses a conical volume different from the cylindrical volume sampled by a DSN antenna. The second component is a beam offset error, which occurs if the calibration instrument is not mounted on the axis of the DSN antenna. These two error sources were calculated for both delay (e.g., VLBI) and delay rate (e.g., Doppler) measurements. The beam mismatch error for both delay and delay rate drops rapidly as the beamwidth of the troposphere calibration instrument (e.g., a water vapor radiometer or an infrared Fourier transform spectrometer) is reduced. At a 10-deg elevation angle, the instantaneous beam mismatch error is 1.0 mm for a 6-deg beamwidth and 0.09 mm for a 0.5-deg beam (these are the full angular widths of a circular beam with uniform gain out to a sharp cutoff). Time averaging for 60–100 sec will reduce these errors by factors of 1.2–2.2. At a 20-deg elevation angle, the lower limit for current Doppler observations, the beam-mismatch delay rate error is an Allan standard deviation over 100 sec of 1.1×10^{-14} with a 4-deg beam and 1.3×10^{-15} for a 0.5-deg beam.

A 50-m beam offset would result in a fairly modest (compared to other expected error sources) delay error ($\leq$0.3 mm for 60-sec integrations at any elevation angle $\geq$6 deg). However, the same offset would cause a large error in delay rate measurements (e.g., an Allan standard deviation of 1.2×10^{-14} over 100 sec at a 20-deg elevation angle), which would dominate over other known error sources if the beamwidth is 2 deg or smaller. An on-axis location is essential for accurate troposphere calibration of delay rate measurements. A half-power beamwidth (for a beam with a tapered gain profile) of 1.2 deg or smaller is desired for calibration of all types of radio metrics. A water-vapor radiometer calibration beam of this size with very low sidelobes would require a clear aperture antenna with a diameter of at least 1.5 m if the primary water vapor sensing channel were in the 20–22 GHz range.

I. Introduction

A. Background

Spacecraft tracking with the DSN involves radio metric (delay or delay rate) measurements over the spacecraft–ground path, sometimes with similar measurements of natural radio sources for purposes of calibration. Any nongeometric delay or delay rate will (if not perfectly calibrated) degrade the accuracy of the tracking measurement. With the increase in the radio frequency used for spacecraft tracking from 2.3 GHz (past missions) to 8.4 GHz (current missions) to 32 GHz (future missions), the effect of charged particles along the spacecraft–ground path is becoming much less significant ($\tau_{ion} \propto \nu^{-2}$, where τ_{ion} is the delay due to charged particles and ν is the radio frequency of the link; the same proportionality holds for delay rates). The effect of the neutral (and nondispersive) troposphere is consequently becoming more dominant. For 8.4-GHz single-frequency observations, the contributions from charged particles and from the troposphere are comparable. For 32-GHz single-frequency tracking (future spacecraft missions) or for dual-frequency measurements (current very long baseline interferometry [VLBI] tracking), the troposphere will dominate over other media effects. The radio metric data themselves can be used to solve for the tropospheric delay. However, spatial and temporal irregularities in the troposphere limit the accuracy with which this process can be done.

A VLBI tracking demonstration on a natural source [1] showed that the tropospheric delay in the direction toward a spacecraft could be estimated to an accuracy of approximately 1 cm, giving an angular measurement accuracy of 1 nanoradian (nrad). *Troposphere calibration is essential in order to improve the accuracy of radio metric tracking beyond this capability.* For a goal of an order-of-magnitude improvement in VLBI tracking accuracy to 100 picoradians (prad), the contribution of the troposphere must be reduced to 1 mm, through a combination of calibration and parameter estimation.

Doppler (delay rate) spacecraft tracking measurements are used both for navigation and for gravitational wave searches. The troposphere calibration goals for future Doppler navigation using Ka-band (31–34 GHz) have not yet been completely specified. However, they are approximately 1×10^{-15} in an Allan standard deviation over 3000–10,000 sec.[1] For the Cassini Gravitational Wave Search experiment (the first one to use a Ka-band link), the *goals* for troposphere calibration are Allan standard deviations of 4×10^{-16} at 1000 and 10,000 sec and 1.3×10^{-15} at 100 sec. The experiment *requirements* are a factor of 4 less stringent.[2] The goals for this experiment were chosen so that the contribution to the delay rate from the uncalibrated troposphere would not dominate the total error at any time scale.

In this article, the terms "delay" and "delay rate" refer to characteristics of the troposphere above a single DSN antenna. This is different than one common use of these terms in DSN spacecraft navigation, where "delay" and "delay rate" refer to the difference (or its time derivative) in arrival times at two DSN antennas (which are often widely spaced). For such station-differenced radio metric data types, the effects calculated here apply to both stations. For widely spaced stations (e.g., at different DSN sites), the delay and delay rate effects in this article should be added in quadrature.

B. Beam Mismatch and Offset

Several types of remote-sensing troposphere calibration systems have been constructed or proposed. These include water vapor radiometers (WVR's) [2], infrared Fourier transform spectrometers (FTS's) [3], and tracking of global positioning satellites (GPS's) [4]. For all these systems, the delay is measured (or inferred from thermal emission) over some volume of troposphere. This volume is approximately conical for techniques that measure emission (WVR, FTS) and cylindrical for techniques that measure delay or absorption (GPS). In both cases, the measured volume of troposphere will not match the volume sampled by a DSN antenna, which is cylindrical with a diameter equal to the DSN antenna diameter, axis centered on the antenna, and extending in the direction of the spacecraft or natural source that is being observed. This cylindrical shape results from the circular shape of the DSN aperture and the fact that the observed radio signal arises from a source on the sky that has a very small ($\ll$1-arcsec) angular extent. The volume mismatch can be broken down into two components: a mismatch in the beam shape and/or size (e.g., conical versus cylindrical) and a horizontal offset in the axes of the two instruments (resulting from the calibration instrument being mounted on the edge of the DSN antenna or off the antenna entirely). There are inhomogeneities in tropospheric refractivity, predominately due to variable water vapor density. As a result, both a beam mismatch and a beam offset would be expected to

[1] S. Thurman, personal communication, Navigation Systems Section, Jet Propulsion Laboratory, Pasadena, California, February 3, 1993.

[2] S. Dolinsky, "Cassini Ka-Band Radio Science Study," (internal document), Jet Propulsion Laboratory, Pasadena, California, August 1992.

cause significant errors. The purpose of this article is to quantify these errors.

Other error sources in the calibration process (e.g., the brightness temperature-to-path delay conversion for water-vapor radiometer measurements) will be important and are likely to be larger than the beam-mismatch and beam-offset errors in some cases. However, these error sources have been neglected in this article, so that the effect of the beam-mismatch and beam-offset errors can be more fully studied. In any case, these other (e.g., non-beam-related) error sources could potentially be reduced to very small levels by advanced instruments and data-processing techniques.

II. Definitions and Calculation Methods

A. Basic Definitions

The wet tropospheric delay $\tau_{DSN}(\vec{x}_{DSN}, t)$ sampled at time t by a DSN antenna of diameter d_{DSN} whose vertex is at position $\vec{x}_{DSN}$ and which is pointed at elevation angle θ_{DSN} and azimuth ϕ_{DSN} is

$$\tau_{DSN}(\vec{x}_{DSN}, t) = \frac{4}{\pi d_{DSN}^2 \sin\theta_{DSN}} \int_0^{d_{DSN}/2} R\,dR \int_0^{2\pi} d\omega \times \int_{-R\cos\omega\cos\theta_{DSN}}^{h} dz\, \chi(\vec{x}_{DSN} + \vec{r}(\theta_{DSN}, \phi_{DSN}, R, \omega, z), t) \quad (1)$$

where $\chi(\vec{x}, t)$ is the refractivity ($\chi \equiv n - 1$, where n is the index of refraction) associated with water vapor at position $\vec{x}$ and time t; $\vec{r}(\theta_{DSN}, \phi_{DSN}, R, \omega, z)$ is the position, relative to the antenna vertex, of a vector in direction $(\theta_{DSN}, \phi_{DSN})$ from polar coordinates (R, ω) on the antenna to height z (the antenna vertex is defined as lying at $z = 0$); and h is the height of the wet troposphere slab, above which χ is assumed to be negligible. The DSN aperture has been approximated by a flat disk. In Eq. (1), the delay has been integrated across the DSN aperture. The actual delay measured by a DSN antenna is derived from the phase of a spatially integrated complex signal. However, because the signal is coherent across the DSN aperture (i.e., the phase variations are $<<1$ rad), the phase of the integrated signal is equal to the integral of the phase across the aperture.

The wet troposphere delay $\tau_{cal}(\vec{x}_{cal}, t)$ measured at time t by a calibration instrument with a conical beam at location $\vec{x}_{cal}$ is

$$\tau_{cal}(\vec{x}_{cal}, t) = \frac{1}{\Omega} \int_0^{\psi/2} \sin u\, du \int_0^{2\pi} \frac{d\omega}{\sin\theta(u,\omega)} \times \int_0^{h} dz\, \chi(\vec{x}_{cal} + \vec{r}(\theta, \phi, z), t) \quad (2)$$

The calibration-beam cross section is assumed to be conical, with uniform gain out to a full angular width ψ. The correspondence between the full width of such a beam and the half-power beamwidth (HPBW) of a Gaussian beam profile is discussed in Section III.A. The beam solid angle Ω is $\Omega = 2\pi[1 - \cos(\psi/2)]$ and $\vec{r}(\theta, \phi, z)$ is the position at height z, relative to the calibration instrument, of a vector along elevation angle θ and azimuth ϕ.

In order for the difference between $\tau_{DSN}(\vec{x}_{DSN}, t)$ and $\tau_{cal}(\vec{x}_{cal}, t)$ to have zero mean, the axis of the calibration beam must be pointed at an elevation angle θ_{cal} that is higher than θ_{DSN}. There is a nonlinear gradient of air mass A with elevation angle θ across the calibration beam, so that the mean air mass corresponds to an elevation angle lower than θ_{cal}. [The air mass $A(\theta)$ is the tropospheric path length at elevation angle θ, normalized to the zenith path length, for a homogeneous troposphere. For the plane-parallel, zero-bending approximation used in these calculations, $A(\theta) = 1/\sin\theta$]. This effect has been previously noted [5]. For the circular calibration-beam cross section used here, θ_{cal} was calculated numerically. For the cases in this article, the largest value of $\theta_{cal} - \theta_{DSN}$ was 0.23 deg (for $\theta_{DSN} = 10$ deg and $\psi = 6$ deg). For a more typical case, the difference in elevation angles was much smaller (e.g., for $\theta_{DSN} = 20$ deg, $\psi = 1$ deg, $\theta_{cal} - \theta_{DSN} = 0.003$ deg).

The absolute value of the average instantaneous difference $\Delta\tau_{inst}$ between the tropospheric delay sampled by the DSN antenna and that measured by the calibration instrument at time t is

$$\Delta\tau_{inst}(t) = \sqrt{\left\langle [\tau_{DSN}(\vec{x}_{DSN}, t) - \tau_{cal}(\vec{x}_{DSN} + \vec{\rho}, t)]^2 \right\rangle} \quad (3)$$

where $\vec{\rho}$ is the vector separation in the horizontal plane between the axis of the DSN antenna and the calibration instrument. When delays are averaged over a time Δt (as in a VLBI measurement), the delay difference $\Delta\tau_{avg}$ is

$$\Delta\tau_{avg}(t_0, \Delta t)$$

$$= \frac{1}{\Delta t}\sqrt{\left\langle \left[\int_0^{\Delta t} \tau_{DSN}(\vec{x}_{DSN}, t_0 + t)\, dt - \int_0^{\Delta t} \tau_{cal}(\vec{x}_{DSN} + \vec{\rho}, t_0 + t)\, dt \right]^2 \right\rangle} \quad (4)$$

where t_0 is the epoch at which the integration begins. For all cases in this article where Eqs. (3) and (4) were evaluated, $\theta_{cal} - \theta_{DSN}$ was chosen so as to give a zero mean for the delay difference. Therefore, Eqs. (3) and (4) yielded the standard deviation of the delay difference.

The Allan standard deviation $\sigma_y(\Delta t)$ [6] of the delay difference $\tau_{DSN}(\vec{x}_{DSN}, t) - \tau_{cal}(\vec{x}_{DSN} + \vec{\rho}, t)$ over a time interval Δt is

$$\sigma_y(\Delta t)$$

$$= \sqrt{\frac{\left\langle [\Delta\tau_{inst}(t + 2\Delta t) - 2\Delta\tau_{inst}(t + \Delta t) + \Delta\tau_{inst}(t)]^2 \right\rangle}{2(\Delta t)^2}} \quad (5)$$

$\sigma_y(\Delta t)$ will be used as a measure of the delay rate error due to beam mismatch and offset.

B. Method of Calculation

In order to evaluate Eqs. (3)–(5) for given values of d_{DSN}, θ_{DSN}, ϕ_{DSN}, and ψ, the expressions were all expanded to integrals of χ, using Eqs. (1) and (2). Products of integrals were converted to multidimensional integrals, and the order of integration and expectation value was reversed. This resulted in multidimensional integrals of $\langle\chi(\vec{x}_1)\chi(\vec{x}_2)\rangle$, where $\vec{x}_1$ and $\vec{x}_2$ are two different vector locations above the DSN antenna and/or the calibration instrument. These integrals were then converted to integrals of a structure function $D_\chi(\vec{x}_1, \vec{x}_1 - \vec{x}_2)$. $\left(D_\chi(\vec{x}_1, \vec{x}_1 - \vec{x}_2) \equiv \left\langle [\chi(\vec{x}_1) - \chi(\vec{x}_2)]^2 \right\rangle\right)$. This procedure has been described elsewhere [7,5].

The model of [7] was used to evaluate these integrals. This model assumes that time variations in χ are due to spatial variations which move past the observer at a constant vector wind velocity $\vec{v}_w$. It assumes that the wet troposphere has constant average properties up to a height h, with zero refractivity above. D_χ has the analytical form

$$D_\chi(\vec{x}_1, \vec{x}_1 - \vec{x}_2) = \frac{C_\chi^2 R^{2/3}}{1 + (R/L_{sat})^{2/3}} \quad (6)$$

where $R = \|\vec{x}_1 - \vec{x}_2\|$ and L_{sat} is a saturation length needed to make D_χ converge at large distances. For the calculations reported here, the values of the parameters adopted were $C_\chi = 1.1 \times 10^{-7}\text{m}^{-1/3}$, $L_{sat} = 3000$ km, $v_w = 8$ m/sec, and the slab height $h = 2$ km. These values give the best match to observed 200-sec delay rates and observed daily zenith delay scatter at DSN sites [1]. Connected-element radio interferometric observations over a 20 km-baseline (DSS 13–DSS 15) at Goldstone provide one test of the range over which C_χ can vary.[3] A series of 10 such observations over a period of 2–3 years was analyzed using the model of [7] to fit the observed scatter in residual delays as a function of time separation, angular separation, and elevation angle. One parameter adjustment of C_χ resulted in a χ^2 agreement of approximately 1 per degree of freedom in all cases, with the best-fit value of C_χ always within a factor of 2 of the value used in this article.

The multidimensional integrals of D_χ were evaluated numerically, using Simpson's Rule. The integration step sizes were reduced until convergence was obtained.

C. Validity of Tropospheric Model

Evidence of the validity of this tropospheric model (Kolmogorov turbulence, frozen flow) comes from several arguments. The model prediction for the dependence of delay rate (at a 200-sec time scale) upon elevation angle agrees with the observed shape from 3 years of DSN VLBI data [7]. The phase structure function from non-DSN VLBI observations [8] shows the same power law dependence (with a curvature at integration times near 200 sec) as that predicted by the model. The delay structure function as measured by a solar hygrometer at JPL[4] shows approximately the same power law dependence at short and long time scales as for the model. In addition to the connected-element radio interferometric data (on natural sources) cited above, additional observations of the Ulysses spacecraft with the same interferometer have been performed.[5] When C_χ is adjusted so as to fit point–point delay scatter of adjacent time points, the estimated difference in the zenith delays between the two antennas

[3] C. Edwards, personal communication, Tracking Systems and Applications Section, Jet Propulsion Laboratory, Pasadena, California, May 28, 1993.

[4] R. Treuhaft, personal communication, Tracking Systems and Applications Section, Jet Propulsion Laboratory, Pasadena, California, March 1993.

[5] C. Edwards, op. cit.

(DSS 13 and DSS 15) was reproduced within a χ^2 of 1 per degree of freedom.

The tropospheric slab approximation (constant structure function from the surface up to a cutoff altitude) was tested numerically. As described in Section III.A, $\Delta\tau_{inst}$ was calculated for two $(\theta_{DSN}, \phi_{DSN}, \psi)$ cases with a more realistic structure function, one in which the refractivity variations decrease exponentially with height. In these two cases, $\Delta\tau_{inst}$ differed from the slab model value by 3 and 11 percents.

III. Results: Delay Errors

A. Beam Mismatch Effects

This subsection presents the errors which occur when the calibration instrument is mounted on the axis of the DSN antenna. The errors are due to the mismatch between the conical beam of the calibration instrument and the cylindrical beam (a good approximation for the tropospheric volume sampled when observing a point source) of the DSN antenna. For the case where the conical calibration beam, projected to a typical atmospheric height, subtends a much larger diameter than the DSN beam, this error is sometimes referred to as beam smearing. The instantaneous and time-averaged errors are given by Eqs. (3) and (4), respectively, with $\vec{\rho} = 0$ in both cases.

1. Zero Integration Time (Instantaneous Beam Mismatch Error). The simplest approximation used was a pencil DSN beam (i.e., a cylinder of diameter zero), a conical calibration beam, and zero integration time. Results are shown in Fig. 1. Note that $\theta_{cal} - \theta_{DSN} > 0$ in order to remove the systematic offset between τ_{cal} and τ_{DSN}.

Seven of the values shown in Fig. 1 were recalculated, using a 34-m-diam DSN antenna beam. The physical mismatch between such a beam and a conical calibration beam was smaller than for a pencil DSN beam. When the maximum calibration-beam diameter $(\psi h / \sin\theta_{DSN})$ was not substantially larger than the DSN beam diameter (h is the wet troposphere slab height), there was a significant reduction in the delay calibration error $\Delta\tau$. (For these cases, the DSN cylindrical beam is a significantly better match to the conical calibration beam than is a pencil beam.) However, these were the cases where the delay error was small (<0.1 mm) with a pencil beam calculation. When a pencil beam calculation yielded a delay error >0.1 mm, the delay error with a cylindrical DSN beam was only slightly smaller than with a pencil DSN beam.

For instantaneous ($\Delta t = 0$) results, no transport of refractivity irregularities occurs. The results represent purely spatial irregularities and do not depend on the wind velocity or direction.

2. Nonzero Integration Time. For four cases, the beam mismatch error was calculated for integration times of 10 and 100 sec, with a 34-m-diam DSN beam. For nonzero integration times, some of the refractivity irregularities move through both beams, reducing the beam mismatch error. A wind velocity of 8 m/sec was assumed, with an azimuth 45 deg different from the DSN pointing direction (the dependence of the time-averaged error upon wind direction was not tested, due to limited computer time). The results are shown in Fig. 2. Delay integration for 100 sec reduces the beam mismatch error by factors of 1.2–2.2. The smaller factors correspond to cases where the instantaneous beam mismatch error was largest. For these cases, the DSN beam cross-sectional area at most heights is much smaller than that of the calibration beam. Therefore, most of the refractivity irregularities in the calibration beam which are transported by the wind are never seen by the DSN beam, and time averaging has little effect.

For an overall line-of-sight calibration accuracy goal of 1 mm, it is desirable that any one error component contribute no more than 20 percent to a 1-mm^2 total variance. This yields a maximum allowed beam-mismatch error of $\sqrt{0.20}$ mm = 0.45 mm. For a 100-sec integration time, a 1-deg full-width calibration beam is acceptable at all elevation angles, and a 2-deg full-width beam is acceptable for elevation angles larger than approximately 7–8 deg.

3. Test of Wet Troposphere Slab Approximation. As a test of the validity of the wet troposphere slab model, the instantaneous delay error with a pencil DSN beam was recalculated using a modified structure function for two sets of (θ_{DSN}, ψ) parameters. This structure function was adopted from previous work[6] and had the form

$$D_\chi(\vec{x}_1, \vec{x}_1 - \vec{x}_2) = \frac{C_\chi^2 R^{2/3}}{1 + (R/L_{sat})^{2/3}} e^{-(z_1+z_2)/z_0} \qquad (7)$$

where $R = ||\vec{x}_1 - \vec{x}_2||$, z_1 and z_2 are the heights of the two points $\vec{x}_1$ and $\vec{x}_2$, L_{sat} is a saturation scale length, and z_0 is a wet troposphere scale height. This structure function accounts for the falloff of water vapor density with height.

[6] J. J. Wiseman, "Altitude Dependence and Saturation Modifications to the Treuhaft-Lanyi Model of Tropospheric Delay Fluctuations," JPL Interoffice Memorandum 335.3-87-203 (internal document), Jet Propulsion Laboratory, Pasadena, California, December 21, 1987.

Values of $C_\chi = 1.34 \times 10^{-7} m^{-1/3}$, L_{sat} = 3000 km, and z_0 = 2 km were used.[7] The upper height of integration was 10 km. The results (for zero integration time and a pencil, or zero-diameter, DSN beam) were 0.15 mm for θ_{DSN} = 10 deg and ψ = 1 deg, and 0.066 mm for θ_{DSN} = 30 deg and ψ = 2 deg. These results differed from the corresponding values in a slab model by 11 and 3 percents, respectively. This level of agreement suggests that the use of a wet troposphere slab approximation was a reasonable one for the purpose of calculating beam mismatch errors.

4. Application to Beams With Tapered Gain Profiles. The beam mismatch results presented in Figs. 1 and 2 are for a uniform calibration beam gain out to a sharp cutoff at $\psi/2$. Actual beam profiles will be tapered, with a Gaussian profile often used for analytical calculations. A Gaussian profile was not used for the calculations in this article because it would have required an excessive amount of computer time.

An approximate correspondence between Gaussian HPBW and the full width ψ of a uniform gain beam profile used in these calculations can be made. In Fig. 1, the dependence of the beam mismatch error $\Delta\tau_{inst}$ upon beamwidth ψ is $\Delta\tau_{inst} \propto \psi^{exp}$, where $exp \approx 1$ at low elevation angles (θ_{DSN} = 6 or 10 deg) where the error is of most concern (i.e., largest). For a circular beam with a uniform gain out to a sharp cutoff, the mean value $\bar{\xi}_{uniform}$ for the angular separation from the axis is $\bar{\xi}_{uniform} = \psi/3$. For a circular beam with a Gaussian profile, the mean value $\bar{\xi}_{Gaussian}$ for this separation is $\bar{\xi}_{Gaussian} = \mathrm{HPBW}\sqrt{\pi}/4\sqrt{\ln 2} \approx 0.532$ HPBW. Assume that $\Delta\tau_{inst} \propto \psi^{exp} \Longrightarrow \Delta\tau_{inst} \propto \bar{\xi}^{exp}$. For $\Delta\tau_{inst}(uniform) = \Delta\tau_{inst}(Gaussian), \bar{\xi}_{uniform} = \bar{\xi}_{Gaussian}$, which yields HPBW = 0.63ψ. Therefore, beam mismatch errors in Figs. 1 and 2 will apply approximately to HPBW = 0.63ψ. A roughly similar relation will hold for beam-mismatch delay rates (Tables 1–2, discussed in Section IV.B). For both delay and delay rate, the correspondence factor will differ very little from 0.63 at higher elevation angles (e.g., the factor is 0.64 if $\Delta\tau_{inst} \propto \psi^{0.8}$, as observed at θ_{DSN} = 60 deg for delay errors).

5. Tests of Numerical Computations. The accuracy of the numerical integrations used to generate the beam mismatch results was tested in two ways. The first test used a quadratic structure function.

$$D_\chi(R) = C_\chi^2 R^2 \qquad (8)$$

Using this structure function, the integrals could be evaluated analytically in many cases [8] (e.g., with pencil DSN or calibration beams, cylindrical DSN beams, or fan calibration beams: a nonzero angular extent in only one dimension). The analytical and numerical beam mismatch results for the quadratic structure function agreed to seven significant figures.

The second test consisted of independent numerical integrations, using the structure function of Eq. (6). These calculations used a code written for the formalism in [5], corrected for a numerical error [9]. A diverging beam with a square cross section and uniform gain was used in these calculations. For a square beam, $\bar{\xi}_{uniform}(square\ beam) = 0.383\psi = 1.15\bar{\xi}_{uniform}(circular\ beam)$. Given the assumed relationship between $\Delta\tau_{inst}$ and $\bar{\xi}$ discussed above, one would expect $\Delta\tau_{inst}(square\ beam) \approx 1.15\Delta\tau_{inst}$ $(circular\ beam)$ for the same value of ψ. For the seven (θ_{DSN}, ψ) cases that were checked, the ratio $\Delta\tau_{inst}(square\ beam) : \Delta\tau_{inst}\ (circular\ beam)$ was in the range 1.06–1.27, with a mean value of 1.15.

B. Beam Offset Effects

1. Results for Nonzero Integration Time. For a nonzero DSN-calibration beam offset [i.e., $\vec{\rho} \neq 0$ in Eqs. (3) and (4)], the number of parameters (beamwidth, elevation angle, DSN-calibration offset distance and azimuth, and wind velocity and azimuth) becomes large. A simplified calculation was necessary in order to explore a reasonable volume of parameter space with the available computer resources. A pencil beam (zero diameter) was used for both the DSN antenna beam and the calibration beam. For this case, $\theta_{cal} = \theta_{DSN}$ gives a delay difference with zero mean.

For a given vector separation $\vec{\rho}$, the offset error depends on two azimuths: that of $\vec{\rho}$ and that of $\vec{v}_w$ (in both cases, it is the azimuth relative to the DSN pointing direction which is important). Using pencil beams, the beam offset error was calculated for all combinations of these two azimuths (with grid spacings of 30 deg). Figures 3–5 show the results for separations ($||\vec{\rho}||$) of 17, 50, and 100 m, and integration times of 30, 60, 120, and 240 sec. Figure 3 gives results for an elevation angle of 60 deg, Fig. 4 for an elevation angle of 20 deg, and Fig. 5 for an elevation angle of 6 deg. The lowest elevation angle currently used by the DSN for spacecraft tracking is 6 deg. The error bars in Figs. 3–5 represent the standard deviation for the 144 azimuth pairs for each separation distance ρ and integration time.

[7] Ibid.

[8] G. Lanyi, personal communication, Tracking Systems and Applications Section, Jet Propulsion Laboratory, Pasadena, California, 1992.

For a 60-sec integration time and a maximum allowed error of 0.45 mm (i.e., a 20-percent contribution to a 1-mm^2 total variance), a separation of 50 m is acceptable at all elevation angles, and a separation of 100 m is acceptable at elevation angles larger than 15 deg.

2. Tests of Numerical Results. In order to check the validity of the use of pencil beams for these calculations, two cases were recalculated using more realistic beams (i.e., a 34-m-diam cylindrical DSN beam and a conical calibration beam). Both cases used $\theta_{DSN} = 10$ deg, $\psi = 1$ deg, an integration time of 100 sec, and a wind velocity of 8 m/sec in azimuth 45 deg. Case 1 had a 17-m DSN-calibration instrument separation in azimuth 90 deg, and case 2 had a 50-m separation in azimuth 0 deg. All azimuths are relative to the DSN pointing direction. The results agreed with the quadrature sum of the beam mismatch (i.e., the zero separation error) and the pencil-beam offset error to within 2 percent in both cases.

The pencil beam results for a limited number of cases (16 combinations of elevation angle, separation distance and azimuth, wind azimuth, and integration time) were recalculated, using a different analytical derivation and computer code. In all 16 cases, the time-averaged beam offset error calculations agreed to <2 percent.

IV. Results: Delay Rate Errors

A. Beam Mismatch Effects

The beam mismatch errors for delay rates [expressed as an Allan standard deviation: Eqs. (3) and (5) with $\vec{\rho} = 0$] were calculated for a cylindrical 34-m-diam DSN beam and a conical calibration beam as a function of elevation angle θ_{DSN}, calibration beam full width ψ, and time interval Δt. Results are given in Table 1 ($\theta_{DSN} = 60$ deg) and Table 2 ($\theta_{DSN} = 20$ deg). The values of $\psi = 1$ deg and $\psi = 2$ deg from Table 2 are plotted in Fig. 6. An elevation angle of 20 deg is the lowest currently used for spacecraft Doppler tracking for navigation purposes, and is the lowest planned for use during the Cassini Gravitational Wave experiment. Column 6 in Tables 1 and 2 gives the Allan standard deviation of the uncalibrated wet troposphere for purposes of comparison. (It was calculated for a 34-m-diam cylindrical beam pointed in a constant direction at that elevation angle). The results in Tables 1 and 2 are averages over wind azimuth with a grid of 30 deg. The variation with wind direction was <5 percent for all cases except for the 10-sec values at 20-deg elevation, where the 1σ variation ranged from 5 percent at $\psi = 0.5$ deg to 38 percent at $\psi = 4$ deg. In addition, the 1σ variation for $\Delta t = 32$ sec, $\psi = 4$ deg, and $\theta_{DSN} = 20$ deg was 9 percent.

For short time scales, other known error sources will dominate over the beam mismatch error for small beamwidths. Projected state-of-the-art frequency-standard capability for the year 2000[9] is an Allan standard deviation of 1×10^{-15} for $\Delta t = 1$–100 sec from a superconducting cavity maser oscillator and $1 \times 10^{-14}/\sqrt{\Delta t}$ for $\Delta t = 100$–10,000 sec from a trapped-ion frequency standard. If water vapor radiometers are used for troposphere calibration (as seems likely), thermal receiver noise will be significant on short time scales. A WVR with a 200-MHz bandwidth and a 300-K total on-sky system temperature, operated in a total power mode, will give an Allan standard deviation of $7.4 \times 10^{-13}(\Delta t)^{-1} t_{integ}^{-0.5}$, where t_{integ} is the WVR integration time (in seconds) used for calibration. In Table 3, the quadrature sums of the Allan standard deviations for beam mismatch error, frequency standard noise, and WVR receiver noise (using the above parameters for the WVR and frequency standard) are given for an elevation angle of 30 deg and beamwidths of 0.5, 1, 2, and 4 deg. The assumed WVR integration time was $t_{integ} = 10$ sec for $\Delta t < 100$ sec, and $t_{integ} = 100$ sec for $\Delta t \geq 100$ sec. The improvement with decreasing beamwidth is substantial down to ψ of approximately 1 deg for timescales up to $\Delta t = 1000$ sec. The improvement is weak for $\psi < 1$ deg and $\Delta t > 1000$ sec. If future frequency standards do not achieve the performance goals given above, delay-rate troposphere calibration will be limited by the frequency standards over a larger range of parameter space (e.g., for time scales shorter than 1000 sec and beamwidths wider than $\psi = 1$ deg).

B. Beam Offset Effects

The delay-rate beam offset errors were calculated [Eqs. (3) and (5) with $\vec{\rho} \neq 0$] for offsets of 17, 50, and 100 m, using full beamwidths (ψ) of 0.5, 1, 2, and 4 deg. The resulting Allan standard deviations were averaged over all wind directions (with a 30-deg grid) and over two DSN-calibration separation azimuths: 0 and 90 deg (as before, all azimuths are relative to the DSN pointing direction). Results (mean and standard deviation) are given in Tables 4 ($\theta_{DSN} = 60$ deg) and 5 ($\theta_{DSN} = 20$ deg). The 17-, 50-, and 100-m values from Table 5 are plotted in Fig. 7. Note that, for a 100-m offset, the beam offset error exceeds the uncalibrated troposphere for time scales less than approximately the crossing time by the wind (100 m/8 m/sec = 12.5 sec). For longer time scales, the offset error decreases, due to the transport of refractivity irregularities. Tables 4 and 5 list results for only one beamwidth: $\psi = 1$ deg. Calculations for other beamwidths indicate a

[9] G. J. Dick, personal communication, Communications Systems Research Section, Jet Propulsion Laboratory, Pasadena, California, January 27, 1993.

weak dependence (<5 percent) upon beamwidth for separations $\rho \geq 50$ m. For $\rho = 17$ m, the mean Allan standard deviations varied by approximately ±10 percent with beamwidth at a 60-deg elevation angle. At $\rho = 17$ m and $\theta_{DSN} = 20$ deg, the mean Allan standard deviations for $\psi = 0.5$, 2, and 4 deg are approximately 0.85, 1.5, and 2.5 times that at $\psi = 1$ deg.

A DSN-calibration instrument offset has a severe impact on the accuracy of delay rate calibration. For a 50-m separation, which is probably near the minimum for a location off the antenna, the goals of the Cassini Gravitational Wave Search experiment cannot be met at time scales less than 3000–4000 sec, and the requirements cannot be met at time scales less than 500 sec. The use of multiple troposphere calibration instruments could potentially reduce the error. Due to time variation in wind azimuth, at least three instruments arrayed around the perimeter of a DSN antenna would be needed in order to achieve a significant reduction in the offset error.

V. Discussion

The consequences of a beam mismatch and offset between a DSN antenna and a calibration instrument (e.g., a WVR) are much more serious for delay rate measurements than for time-averaged delay measurements. This situation reflects the fact that differentiating a noisy signal (the wet tropospheric refractivity in this case) increases its amplitude, whereas integrating such a signal decreases its amplitude. For delay measurements with ≥1-min integration, HPBW ≤1.3 deg and an offset ≤50 m are needed for an overall calibration goal of 1 mm. A narrower beamwidth and an on-axis location would be desirable, in case other error sources should prove to be small enough to achieve a calibration accuracy of <1 mm.

For delay rate measurements, the performance cost of an offset location for a calibration instrument is severe. A calibration instrument that could measure line-of-sight delay perfectly, but which had a 50-m offset from the axis of a DSN antenna, would calibrate only 96 percent of the total 1000-sec troposphere Allan standard deviation, 80 percent for 100-sec intervals, and could do no useful calibration for 10-sec time scales (these values are for a 20-deg elevation angle; the fractional calibration errors are approximately a factor of 2 smaller at 60 deg). *An on-axis location for a calibration instrument is essential in order to perform accurate tropospheric delay rate measurements.* Specifically, it is required in order to meet the Cassini Gravitational Wave Search *goals* at time scales less than 3000–4000 sec, or the *requirements* at times scales less than 500 sec. Multiple calibration instruments mounted around the perimeter of the DSN antenna could potentially reduce the offset error, but this has not been quantified. Given the high cost of water-vapor-sensing instruments (e.g., WVR's), this is not an attractive option.

A calculation of the power spectral density error resulting from an offset location reached conclusions similar to the results shown in Tables 4 and 5.[10] That power spectral density calculation was done in a very different manner from the calculations reported here (although it used the same wet troposphere fluctuation model [7]).

The beam mismatch error (i.e., that for zero on-axis offset) is important for time scales of 1000 sec and smaller. This error drops rapidly as the beamwidth is reduced, down to HPBW = 0.7 deg. The improvement for HPBW <0.7 deg is minor, due to contributions from other error sources (frequency standard noise and WVR thermal receiver noise). For HPBW = 0.7 deg, the beam mismatch error will not interfere with the Cassini Gravitational Wave Search *requirements* at any time scale. It will prevent the *goals* from being met at time scales shorter than 200 sec.

[10] P. Richter, personal communication, Telecommunications Systems Section, Jet Propulsion Laboratory, Pasadena, California, 1993.

Acknowledgments

The authors thank G. J. Dick for helpful discussions on the development and calibration of frequency standards. R. N. Treuhaft and G. M. Resch made helpful comments, based on readings of earlier versions of this article.

References

[1] R. N. Treuhaft and S. T. Lowe, "A Measurement of Planetary Relativistic Deflection," *The Astronomical Journal*, vol. 102, no. 5, pp. 1879–1888, November 1991.

[2] G. Elgered, J. L. Davis, T. A. Herring, and I. I. Shapiro, "Geodesy by Radio Interferometry: Water Vapor Radiometry for Estimation of the Wet Delay," *Journal of Geophysical Research*, vol. 96, no. B4, pp. 6541–6555, April 10, 1991.

[3] W. L. Smith, H. E. Revercomb, H. B. Howell, H.-L. Huang, R. O. Knuteson, E. W. Koenig, D. D. LaPorte, S. Silverman, L. A. Sromovsky, and H. M. Woolf, "GHIS—The GOES High-Resolution Interferometer Sounder," *Journal of Applied Meteorology*, vol. 29, pp. 1189–1204, December 1990.

[4] D. M. Tralli, S. M. Lichten, and T. A. Herring, "Comparison of Kalman Filter Estimates of Zenith Atmospheric Path Delays Using the Global Positioning System and Very Long Baseline Interferometry," *Radio Science*, vol. 27, no. 6, pp. 999–1007, November–December 1992.

[5] J. Z. Wilcox, "The Effect of Tropospheric Fluctuations on the Accuracy of Water Vapor Radiometry," *The Telecommunications and Data Acquisition Progress Report 42-110*, vol. April–June 1992, Jet Propulsion Laboratory, Pasadena, California, pp. 33–51, August 15, 1992.

[6] D. W. Allan, "Statistics of Atomic Frequency Standards," *Proceedings of the IEEE*, vol. 54, no. 2, pp. 221–230, February 1966.

[7] R. N. Treuhaft and G. E. Lanyi, "The Effect of the Dynamic Wet Troposphere on Radio Interferometric Measurements," *Radio Science*, vol. 22, no. 2, pp. 251–265, March–April 1987.

[8] A. E. E. Rogers, "Coherence Limits in VLBI Observations at 3-millimeter Wavelength," *Radio Science*, vol. 19, no. 6, pp. 1552–1560, November–December 1984.

[9] J. Z. Wilcox, "Errata: The Effect of Tropospheric Fluctuations on the Accuracy of Water Vapor Radiometry," *The Telecommunications and Data Acquisition Progress Report 42-114*, vol. April–June 1993, Jet Propulsion Laboratory, Pasadena, California, p. 346, August 15, 1993.

Table 1. Beam mismatch Allan standard deviation between a conical calibration beam and a coaxial 34-m-diam DSN beam for a 60-deg elevation angle.

Time interval (Δt), sec	Full beamwidth (ψ), deg				Uncalibrated troposphere, Allan standard deviation
	0.5	1	2	4	
10	1.0×10^{-14}	7.8×10^{-15}	1.1×10^{-14}	2.6×10^{-14}	6.4×10^{-14}
32	3.2×10^{-15}	2.4×10^{-15}	3.5×10^{-15}	8.0×10^{-15}	5.5×10^{-14}
100	1.0×10^{-15}	7.7×10^{-16}	1.1×10^{-15}	2.6×10^{-15}	4.3×10^{-14}
320	3.2×10^{-16}	2.4×10^{-16}	3.5×10^{-16}	8.0×10^{-16}	2.9×10^{-14}
1000	1.0×10^{-16}	7.7×10^{-17}	1.1×10^{-16}	2.6×10^{-16}	1.6×10^{-14}
3200	3.2×10^{-17}	2.4×10^{-17}	3.5×10^{-17}	8.0×10^{-17}	8.0×10^{-15}
10,000	1.0×10^{-17}	7.7×10^{-18}	1.1×10^{-17}	2.6×10^{-17}	3.8×10^{-15}

Table 2. Beam mismatch Allan standard deviation between a conical calibration beam and a coaxial 34-m-diam DSN beam for a 20-deg elevation angle.

Time interval (Δt), sec	Full beamwidth (ψ), deg				Uncalibrated troposphere, Allan standard deviation
	0.5	1	2	4	
10	1.3×10^{-14}	2.2×10^{-14}	4.3×10^{-14}	5.9×10^{-14}	8.3×10^{-14}
32	4.3×10^{-15}	7.4×10^{-15}	1.7×10^{-14}	3.2×10^{-14}	7.5×10^{-14}
100	1.3×10^{-15}	2.4×10^{-15}	5.6×10^{-15}	1.1×10^{-14}	6.3×10^{-14}
320	4.2×10^{-16}	7.4×10^{-16}	1.7×10^{-15}	3.5×10^{-15}	5.1×10^{-14}
1000	1.3×10^{-16}	2.4×10^{-16}	5.6×10^{-16}	1.1×10^{-15}	3.5×10^{-14}
3200	4.1×10^{-17}	7.3×10^{-17}	1.7×10^{-16}	3.5×10^{-16}	1.9×10^{-14}
10,000	1.3×10^{-17}	2.4×10^{-17}	5.6×10^{-17}	1.1×10^{-16}	9.3×10^{-15}

Table 3. Total Allan standard deviation due to beam mismatch, frequency standard noise, and WVR receiver noise for a 30-deg elevation angle and an on-axis location.

Time interval (Δt), sec	Full beamwidth (ψ), deg			
	0.5	1	2	4
10	2.6×10^{-14}	2.7×10^{-14}	3.7×10^{-14}	5.4×10^{-14}
32	8.0×10^{-15}	8.4×10^{-15}	1.2×10^{-14}	2.0×10^{-14}
100	1.7×10^{-15}	1.9×10^{-15}	3.2×10^{-15}	6.3×10^{-15}
320	7.0×10^{-16}	7.4×10^{-16}	1.1×10^{-15}	2.0×10^{-15}
1000	3.5×10^{-16}	3.6×10^{-16}	4.4×10^{-16}	7.0×10^{-16}
3200	1.8×10^{-16}	1.9×10^{-16}	2.0×10^{-16}	2.6×10^{-16}
10,000	1.0×10^{-16}	1.0×10^{-16}	1.0×10^{-16}	1.2×10^{-16}

Table 4. Allan standard deviation for an offset between a conical calibration beam and a 34-m-diam DSN beam for a 60-deg elevation angle and a 1-deg full beamwidth.

Time interval (Δt), sec	DSN-calibration separation, m			Uncalibrated troposphere, Allan standard deviation
	17	50	100	
10	$3.0 \pm 0.4 \times 10^{-14}$	$7.1 \pm 1.0 \times 10^{-14}$	$9.2 \pm 1.3 \times 10^{-14}$	6.4×10^{-14}
32	$1.1 \pm 0.1 \times 10^{-14}$	$2.8 \pm 0.2 \times 10^{-14}$	$4.6 \pm 0.5 \times 10^{-14}$	5.5×10^{-14}
100	$3.8 \pm 0.3 \times 10^{-15}$	$1.0 \pm 0.1 \times 10^{-14}$	$1.7 \pm 0.1 \times 10^{-14}$	4.3×10^{-14}
320	$1.2 \pm 0.1 \times 10^{-15}$	$3.2 \pm 0.2 \times 10^{-15}$	$5.6 \pm 0.2 \times 10^{-15}$	2.9×10^{-14}
1000	$4.0 \pm 0.2 \times 10^{-16}$	$1.0 \pm 0.1 \times 10^{-15}$	$1.8 \pm 0.1 \times 10^{-15}$	1.6×10^{-14}
3200	$1.2 \pm 0.1 \times 10^{-16}$	$3.2 \pm 0.2 \times 10^{-16}$	$5.6 \pm 0.2 \times 10^{-16}$	8.0×10^{-15}
10,000	$4.0 \pm 0.2 \times 10^{-17}$	$1.0 \pm 0.1 \times 10^{-16}$	$1.8 \pm 0.1 \times 10^{-16}$	3.8×10^{-15}

Table 5. Allan standard deviation for an offset between a conical calibration beam and a 34-m-diam DSN beam for a 20-deg elevation angle and a 1-deg full beamwidth.

Time interval (Δt), sec	DSN-calibration separation, m			Uncalibrated troposphere, Allan standard deviation
	17	50	100	
10	$3.4 \pm 1.8 \times 10^{-14}$	$7.1 \pm 4.1 \times 10^{-14}$	$9.4 \pm 4.4 \times 10^{-14}$	8.3×10^{-14}
32	$1.3 \pm 0.6 \times 10^{-14}$	$3.6 \pm 1.5 \times 10^{-14}$	$5.2 \pm 2.3 \times 10^{-14}$	7.5×10^{-14}
100	$4.8 \pm 2.1 \times 10^{-15}$	$1.2 \pm 0.5 \times 10^{-14}$	$2.0 \pm 0.8 \times 10^{-14}$	6.3×10^{-14}
320	$1.6 \pm 0.6 \times 10^{-15}$	$4.0 \pm 1.6 \times 10^{-15}$	$7.2 \pm 2.6 \times 10^{-15}$	5.1×10^{-14}
1000	$5.2 \pm 2.0 \times 10^{-16}$	$1.3 \pm 0.5 \times 10^{-15}$	$2.4 \pm 0.8 \times 10^{-15}$	3.5×10^{-14}
3200	$1.6 \pm 0.6 \times 10^{-16}$	$4.0 \pm 1.6 \times 10^{-16}$	$7.2 \pm 2.6 \times 10^{-16}$	1.9×10^{-14}
10,000	$5.2 \pm 2.0 \times 10^{-17}$	$1.3 \pm 0.5 \times 10^{-16}$	$2.4 \pm 0.8 \times 10^{-16}$	9.3×10^{-15}

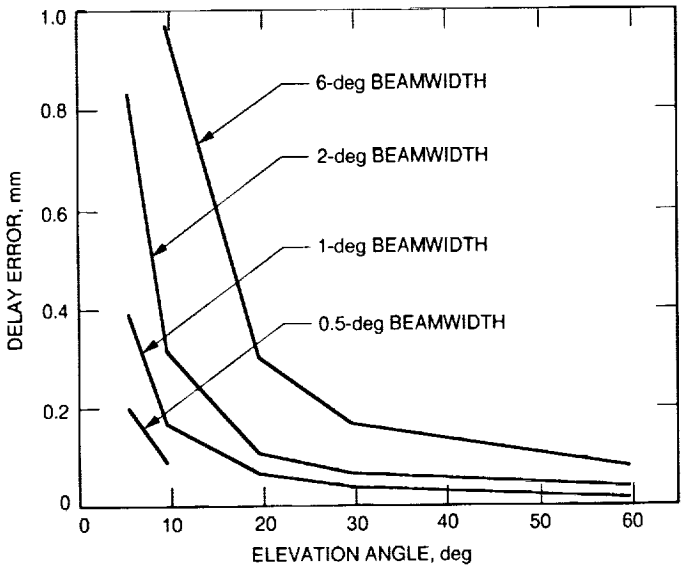


Fig. 1. Instantaneous delay calibration error due to beam mismatch as a function of elevation angle (θ) and full beamwidth (ψ) (circular beamwidth with a uniform gain) for a pencil DSN beam.

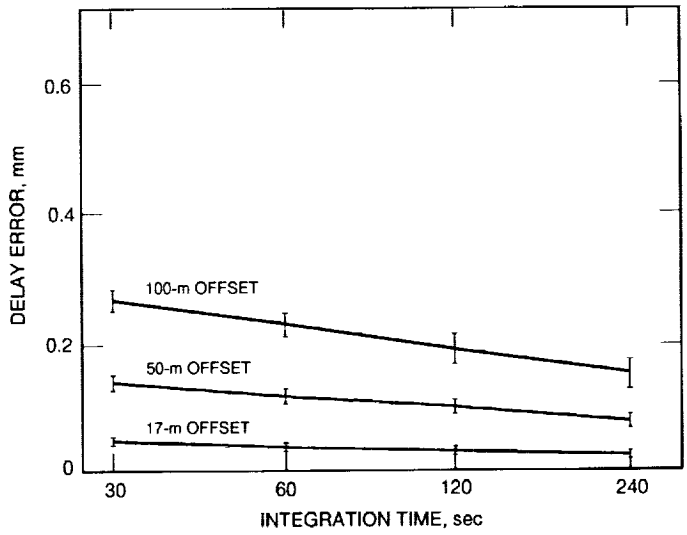


Fig. 3. Time-averaged delay calibration error due to beam offset as a function of time for offsets (between the axis of the DSN antenna and the calibration instrument) of 17, 50, and 100 m. The elevation angle is 60 deg. A pencil DSN beam was used in the calculations.

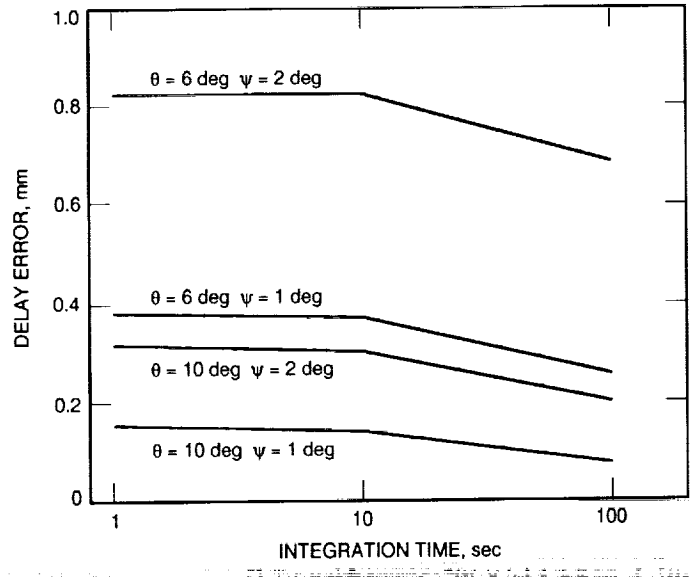


Fig. 2. Time-averaged delay calibration error due to beam mismatch for four combinations of elevation angle (θ) and full beamwidth (ψ) (circular beam with a uniform gain). A 34-m diam DSN beam was used in the calculations.

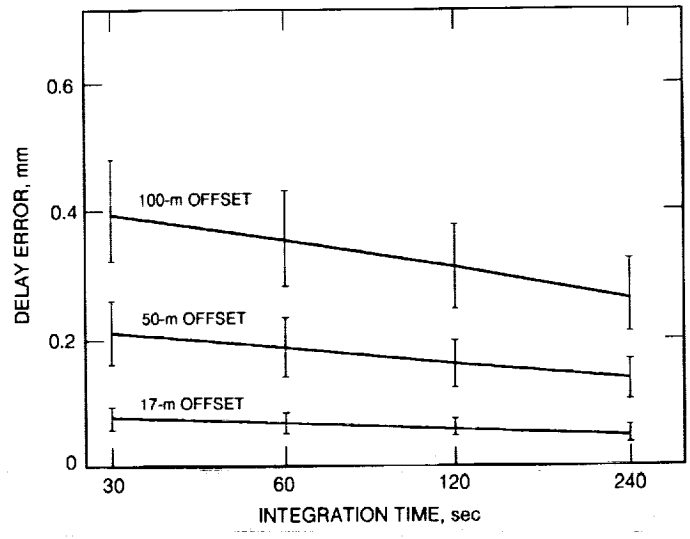


Fig. 4. Time-averaged delay calibration error due to beam offset as a function of time for offsets (between the axis of the DSN antenna and the calibration instrument) of 17, 50, and 100 m. The elevation angle is 20 deg. A pencil DSN beam was used in the calculations.

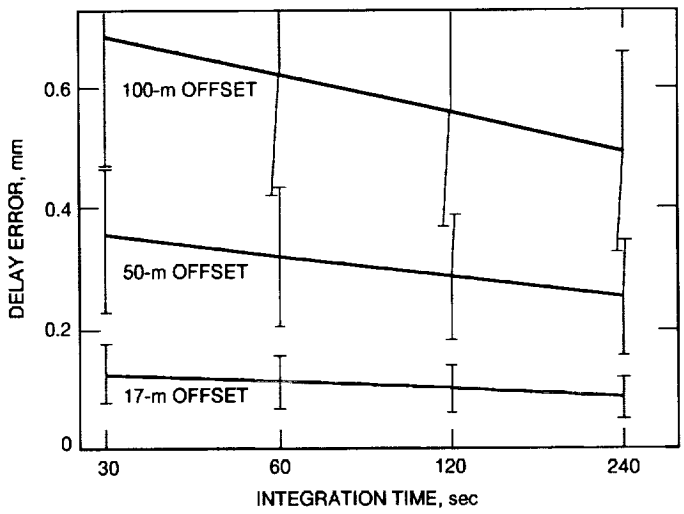


Fig. 5. Time-averaged delay calibration error due to beam offset as a function of time for offsets (between the axis of the DSN antenna and the calibration instrument) of 17, 50, and 100 m. The elevation angle is 6 deg. A pencil DSN beam was used in the calculations.

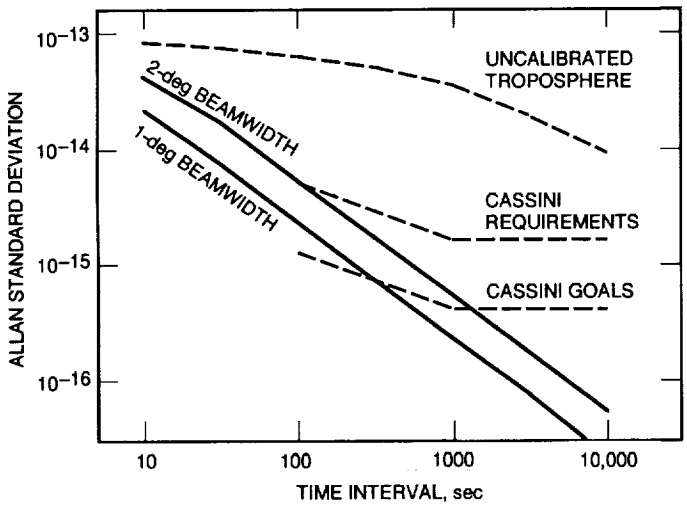


Fig. 6. Delay rate calibration error (expressed as an Allan standard deviation) due to beam mismatch as a function of time interval for full beamwidths ψ (circular beam with a uniform gain) of ψ = 2 deg and ψ = 1 deg. The elevation angle is 20 deg. The Allan standard deviation of the uncalibrated troposphere is plotted for comparison, as are the goals and requirements of the Cassini Gravitational Wave Search experiment. A 34-m-diameter DSN beam was used in the calculations.

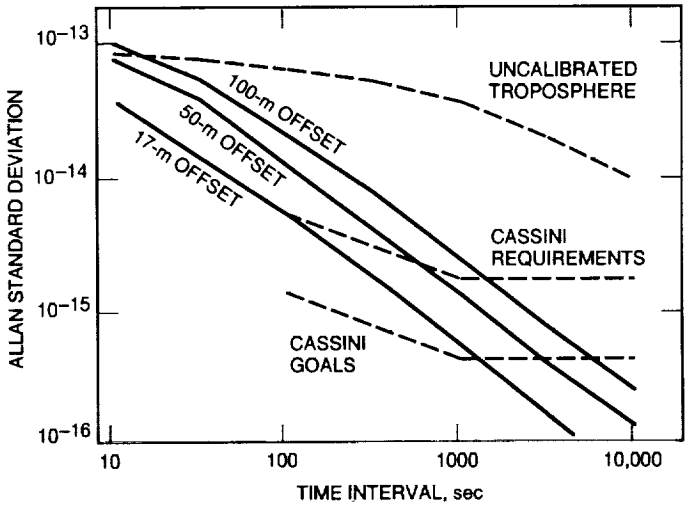


Fig. 7. Delay rate calibration error (expressed as an Allan standard deviation) due to beam offset as a function of time interval for offsets (between the axis of the DSN antenna and the calibration instrument) of 17, 50, and 100 m. The elevation angle is 20 deg. The Allan standard deviation of the uncalibrated troposphere is plotted for comparison, as are the goals and requirements of the Cassini Gravitational Wave Search experiment. A 34-m-diameter DSN beam was used in the calculations.

Improved Treatment of Global Positioning System Force Parameters in Precise Orbit Determination Applications

Y. Vigue, S. M. Lichten, R. J. Muellerschoen, G. Blewitt, and M. B. Heflin
Tracking Systems and Applications Section

Data collected from a worldwide 1992 experiment have been processed at JPL to determine precise orbits for the satellites of the Global Positioning System (GPS). A filtering technique has been tested to improve modeling of solar-radiation pressure force parameters for GPS satellites. The new approach improves orbit quality for eclipsing satellites by a factor of two, with typical results in the 25- to 50-cm range. The resultant GPS-based estimates for geocentric coordinates of the tracking sites, which include the three DSN sites, are accurate to 2 to 8 cm, roughly equivalent to 3 to 10 nrad of angular measure.

I. Introduction

The satellites of the Global Positioning System (GPS) are maintained by the U.S. Department of Defense for navigational purposes. These satellites are distributed in six evenly spaced orbit planes, at an orbit altitude of 20,000 km and with an orbit period of approximately 12 hr. GPS measurements collected from globally distributed ground receivers are also being used to estimate Earth rotation, polar motion, and station coordinates. Estimation of the geocenter, or Earth's center of mass, has geophysical and scientific implications. GPS estimates of the geocenter can be used for precise reference-frame calibration and alignment and to eliminate a bias-type error in DSN site coordinates. Over time intervals of weeks to months, the GPS measurements can be used to precisely monitor variations in tracking site coordinates due to crustal motion and continental drift. The data used in this analysis are taken from the International Global Positioning System Geodynamics Service 1992 (IGS'92) campaign. IGS'92 consisted of approximately 30 globally distributed Rogue receiver sites tracking the 18 GPS satellites active during this time. The focus of this analysis was to assess the effects of mismodeling satellite force parameters due to solar radiation on the estimation of other parameters, such as GPS orbits and geocentric DSN coordinates.

II. Estimation Strategy

A unique strength of GPS measurements is that the satellites are sensitive to the geocenter, yet relatively insensitive to errors in the gravity field because of their high orbit altitude and the relatively short data arcs (30 hr) needed for the solution. Mismodeling of satellite force parameters, however, can have a significant effect on satellite orbits, especially in orbit prediction [1]. The goal of this study was to improve satellite modeling in order to achieve centimeter-level accuracy for global geocentric coordinates. To achieve this level of accuracy, it is essential

to include corrections for numerous potential errors, including Earth's rotation and orientation, atmospheric distortion of the radio signals from the satellites, gravitational and nongravitational forces acting on the satellites, and various geophysical effects.

The data used in this analysis were taken from GPS week 660, which ran from August 30 through September 5, 1992. In general, the data contain carrier phase and pseudo-range measurements from 18 available GPS satellites tracked by approximately 30 globally distributed JPL Rogue receivers. These tracking sites are shown in Fig. 1 and listed in Table 1. The data were processed using the GPS Inferred Positioning SYstem/Orbit Analysis and Software Integration System, (GIPSY/OASIS II) software [2,3].[1] All nonfiducial station locations were estimated, as well as Earth orientation parameters, GPS carrier phase biases, and random-walk zenith troposphere delays for each tracking site; all transmitter/receiver clocks but one were treated as white noise parameters. X and Y polar motion, pole rate, and UT1–UTC rate are estimated as constant parameters (reset every 24 hr).

One of the most important recent innovations is a new approach to modeling the effects of solar radiation pressure on the satellite orbits. The solar radiation environment of the GPS satellites is generally constant except when a satellite's orbit is in eclipse season. When this occurs, the satellites pass through Earth's shadow, changing the amount of solar radiation the satellite receives. In general, three body-fixed solar-radiation pressure parameters are estimated for all GPS satellite orbits, regardless of whether or not those satellites are in eclipsing orbits. For this analysis however, the GPS orbits were estimated with five solar pressure parameters, which are shown in Table 2. Two solar radiation pressure parameters were estimated as constant: GYC and GX/GZ, where X, Y, and Z represent spacecraft body-fixed coordinates and GYC is the "Y-bias" parameter [4]. GX/GZ represents a single combined scale factor for the ROCK4[2] solar radiation parameter [4], while GX and GZ are independent scale factors for the X and Z directions. The three remaining solar pressure parameters are estimated as stochastic corrections to the constant solar pressure parameters, modeled as a first-order Gauss–Markov process [5]. This technique has enabled the achievement of few-centimeter geocentric coordinate accuracy [6].

III. Results and Discussion

The GPS constellation consists of satellites in Earth orbit configured in six evenly spaced orbit planes. At times, satellites in certain orbit planes experience what will be referred to in this article as eclipsing, or shadowing. Figure 2 shows a satellite in an eclipsing orbit plane. GPS satellites eclipsing during GPS week 660 were PRN02, PRN14, PRN16, PRN20, PRN21, and PRN23.

The results presented here illustrate recent improvements in GPS orbit accuracy. The three-dimensional root-sum-square (*3DRSS*) orbit repeatability for each GPS satellite (j) is defined as [7]

$$3DRSS(j) = \sqrt{\frac{1}{T}\sum_{t=1}^{T} d_{jkt}^2} \qquad (1)$$

where j is the satellite PRN number, T is the number of epochs within both overlap segments, t is the time index, k indicates which overlap segment, and d_{jkt}^2 is the three-dimensional distance between corresponding points on the two overlapping segments. In other words, the orbit quality of a single day is quantified as the rms difference between the ephemerides computed over the corresponding 3 hr of orbit overlap at both ends of that day. This concept is illustrated in Fig. 3.[3]

Figure 4 shows the *3DRSS* orbit repeatability for all GPS satellites active during GPS week 660. The improvement in the orbits with stochastic solar pressure parameters is approximately 25 percent overall for all the GPS satellites (averaged), with a 44 percent improvement in the eclipsing orbits and only an 8 percent improvement in the noneclipsing orbits. This demonstrates how mismodeling of satellite force parameters due to solar radiation can have a significant effect on GPS orbit accuracy, especially for an eclipsing satellite. In an eclipsing orbit, a satellite does not experience constant solar radiation forces throughout the entire orbit arc.

The orbit repeatability for each satellite using both strategies is shown in Table 3. These values represent

[1] S. C. Wu, Y. Bar-Sever, S. Bassiri, W. I. Bertiger, G. A. Hajj, S. M. Lichten, R. P. Malla, B. K. Trinkle, and J. T. Wu, *TOPEX/POSEIDON Project: Global Positioning System (GPS) Precison Orbit Determination (POD) Software Design*, JPL D-7275 (internal document), Jet Propulsion Laboratory, Pasadena, California, March 1990.

[2] ROCK4 is a solar radiation pressure model provided by Rockwell Corp. for GPS satellites.

[3] J. F. Zumberge, "Regular Analysis of Global Positioning System Data from a Globally-Distributed Network of Rogue Receivers," Interoffice Memorandum 335.4-92-012 (internal document), Jet Propulsion Laboratory, Pasadena, California.

the average orbit repeatability over the 7 days in GPS week 660, with the eclipsing satellites indicated. This table shows the improvement in orbit accuracy due to estimating stochastic corrections to the GPS solar radiation parameters (srp). The table also shows how the eclipsing satellites benefit more from improved estimation strategy.

IV. Concluding Remarks

This analysis shows that GPS orbit accuracy can be improved by estimating stochastic corrections to the GPS dynamic parameters. This new approach improves orbit quality for eclipsing satellites from 85 to 47 cm. This level of orbit accuracy is in agreement with results given by Zumberge et al. [7], where routine processing of GPS data shows orbit accuracy in the range of 25 to 50 cm. A direct result of the improvements in orbit accuracy can be seen in the improvement of the geocentric station coordinate accuracy [6]. The goal of the analysis described in [6] was to achieve centimeter-level accuracy for global geocentric coordinates. Those GPS results were obtained with 3 months of GPS measurements and compared to satellite laser ranging (SLR) solutions from many years of repeated observations. It was demonstrated that the geocenter estimates from GPS are accurate to better than 2 cm in the X and Y components and to approximately 8 cm in the Z (where Z is parallel to the axis of rotation). This capability has important benefits for DSN tracking and for geophysical research, such as geocentric crustal motion studies, and for understanding the magnitude and time scale of geocenter variations and their origin. Precise tracking of interplanetary spacecraft and Earth orbiters requires that DSN geocentric station coordinates be determined to high accuracy. The accuracy for geocentric DSN coordinates determined from the GPS data is approximately 3 to 10 nrad in angular measure.

Figure 5 shows a history of the improvements in the GPS determination of the geocenter. Most of the recent improvements can be attributed to the changes in the technique used for modeling solar radiation pressure, changes that have been described in this article. These new results enable the tracking sites to be precisely specified in a reference frame whose origin is at the geocenter, and will enable precise alignment of different reference frames used for Earth-based tracking, interplanetary navigation, and geophysical measurements. Some future work in this area will examine systematic error sources, further improve the orbit modeling, study fixing phase ambiguities, and incorporate data from low Earth orbiters, such as TOPEX/POSEIDON, to look for time-varying signatures of the geocenter.

Acknowledgments

The authors thank Jim Zumberge and the JPL Fiducial Laboratory for an International Natural Science Network (FLINN) data processing center for special software used in this analysis.

References

[1] Y. Vigue, *Thermal Imbalance Effects on a GPS Satellite*, Technical Memorandum 90-1, University of Texas Center for Space Research, Austin, Texas, 1990.

[2] S. Lichten and J. Border, "Strategies for High Precision Global Positioning System Orbit Determination," *Journal of Geophysical Research*, vol. 92, pp. 12751–12762, November 10, 1987.

[3] O. Sovers and J. Border, *Observation Model and Parameter Partials for the JPL Geodetic GPS Modeling Software "GPSOMC,"* JPL Publication 87-21, Rev. 2, Jet Propulsion Laboratory, Pasadena, California, 1990.

[4] H. F. Fliegel and T. E. Gallini, "Radiation Pressure Models for Block II GPS Satellites," Interoffice Memorandum, Aerospace Corporation, El Segundo, California, 1989.

[5] G. Bierman, *Factorization Methods for Discrete Sequential Estimation*, vol. 128, New York: Academic Press, 1977.

[6] Y. Vigue, G. Blewitt, S. M. Lichten, M. B. Heflin, and R. J. Muellerschoen, "Recent High-Accuracy GPS Estimates of the Geocenter," *EOS Transactions, American Geophysical Union*, vol. 73, no. 43, p. 135, October 1992.

[7] J. F. Zumberge, D. Jefferson, G. Blewitt, F. Webb, and M. Heflin, "Processing IGS Data at JPL," paper presented at the American Geophysical Union, Fall Meeting, San Francisco, California, 1992.

Table 1. GPS Rogue receiver sites.

ALBH	Albert Head, B.C., Canada	NALL	Ny Ålesund, Norway
ALGO	Algonquin, Ontario, Canada	ONSA	Onsala, Sweden
CANB	Canberra, Australia	PAMA	Pamatai, Tahiti
FAIR	Fairbanks, Alaska (USA)	PENT	Penticton, B.C., Canada
GOLD	Goldstone, California (USA)	PGC1	Victoria, Canada
HART	Hartebeesthoek, South Africa	PINY	Pinyon, California (USA)
HARV	Harvest Platform, California (USA)	QUIN	Quincy, California (USA)
HERS	Herstmonceux, Great Britain	RCM2	Richmond, Florida (USA)
HONE	Honefoss, Norway	SANT	Santiago, Chile
JPLM	Pasadena, California (USA)	SCRI	La Jolla, California (USA)
KOKB	Kokee, Hawaii (USA)	STJO	St. Johns, Canada
KOSG	Kootwijk, The Netherlands	TAIW	Taiwan
MADR	Madrid, Spain	TROM	Tromso, Norway
MASP	Maspalomas, Grand Canary Is., Africa	USUD	Usuda, Japan
MATE	Matera, Italy	WETB	Wettzell, Germany
MCMU	McMurdo Station, Ross Is., Antarctica	YAR1	Yarragadee, Australia
METS	Metsahovi, Finland	YELL	Yellowknife, Canada

Table 2. Estimated solar radiation parameters.

Parameter	Model	A priori sigma
GX/GZ	Constant[a]	100%
GYC	Constant[a]	2 nm/sec^2
GX	First-order G-M[b]	10%
GY	First-order G-M[b]	0.1 nm/sec^2
GZ	First-order G-M[b]	10%

[a] Estimated as constant parameter with no process noise.

[b] Gauss-Markov (G-M) with time constant of 4 hr and steady-state sigma of 0.1 nm/sec^2. Typical magnitude of GX and GZ accelerations is 100 nm/sec^2.

Table 3. *3DRSS* orbit repeatability for GPS week 660.

PRN number	With stochastic srp, m	Without stochastic srp, m
2[a]	0.76	0.64
3	0.70	0.28
11	0.38	0.32
12	0.43	0.45
13	0.32	0.56
14[a]	0.48	0.94
15	0.51	0.58
16[a]	0.28	1.02
17	0.52	0.56
18	0.43	0.60
19	0.66	0.44
20[a]	0.72	1.19
21[a]	0.30	0.75
23[a]	0.29	0.54
24	0.45	0.56
25	0.42	0.47
26	0.30	0.83
28	0.48	0.43

[a] Eclipsing satellites.

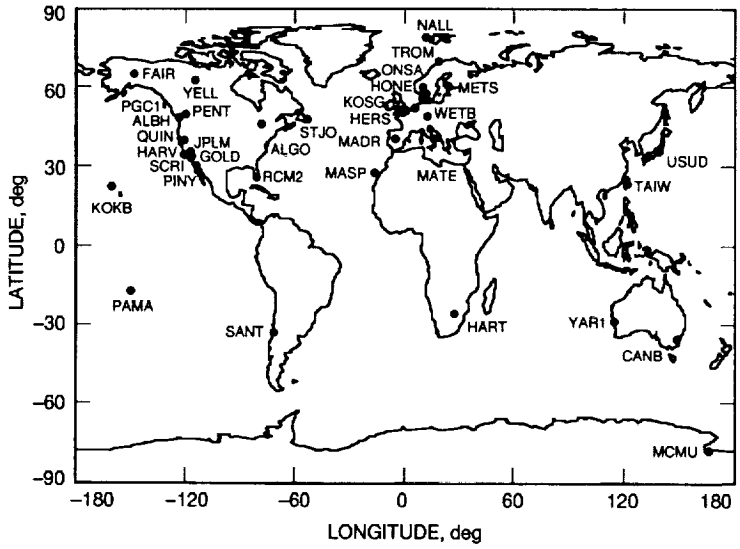


Fig. 1. 1992 GPS tracking sites.

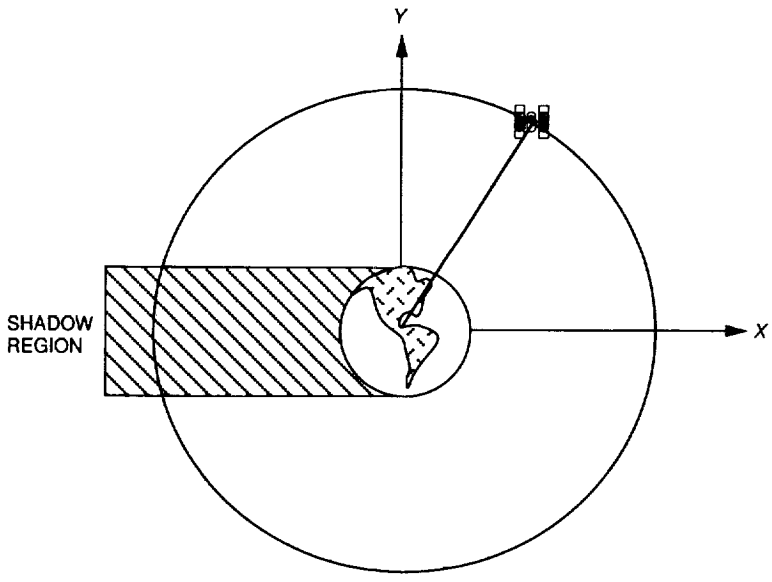


Fig. 2. Satellite in Earth orbit with simple cylindrical shadow model.

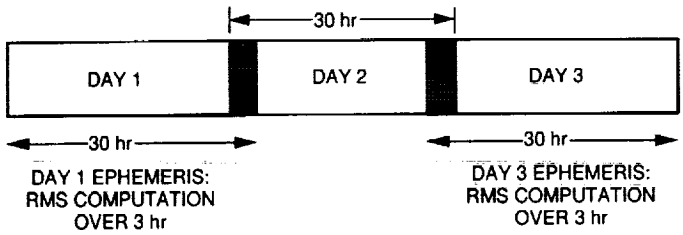


Fig. 3. Assessment of GPS orbit accuracy.

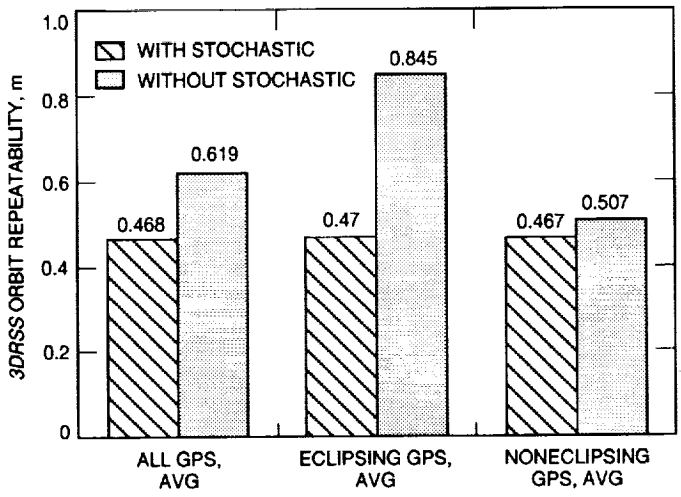


Fig. 4. *3DRSS* GPS orbit repeatability—7 days, GPS week 660.

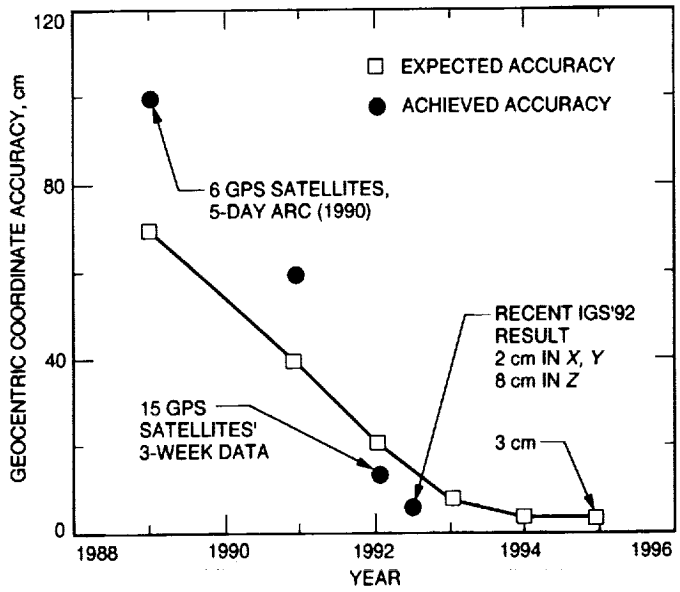


Fig. 5. Recent improvements in GPS geocentric coordinate accuracy.

Use of Global Positioning System Measurements to Determine Geocentric Coordinates and Variations in Earth Orientation

R. P. Malla,[1] S.-C. Wu, and S. M. Lichten
Tracking Systems and Applications Section

Geocentric tracking station coordinates and short-period Earth-orientation variations can be measured with Global Positioning System (GPS) measurements. Unless calibrated, geocentric coordinate errors and changes in Earth orientation can lead to significant deep-space tracking errors. Ground-based GPS estimates of daily and subdaily changes in Earth orientation presently show centimeter-level precision. Comparison between GPS-estimated Earth-rotation variations, which are the differences between Universal Time 1 and Universal Coordinated Time (UT1–UTC), and those calculated from ocean tide models suggests that observed subdaily variations in Earth rotation are dominated by oceanic tidal effects. Preliminary GPS estimates for the geocenter location (from a 3-week experiment) agree with independent satellite laser-ranging estimates to better than 10 cm. Covariance analysis predicts that temporal resolution of GPS estimates for Earth orientation and geocenter improves significantly when data collected from low Earth-orbiting satellites as well as from ground sites are combined. The low Earth GPS tracking data enhance the accuracy and resolution for measuring high-frequency global geodynamical signals over time scales of less than 1 day.

I. Introduction

Measurement of Earth orientation determines the Earth's rotation rate and the position of the pole (rotation axis) with respect to inertial space. Monitoring changes in global geodynamical parameters (GGP)—Earth-orientation parameters and the location of the geocenter, or Earth's center of mass, relative to a defined terrestrial-reference frame—is necessary to correctly model deep-space tracking data used for navigation and trajectory determination. At the present time, the Deep Space Network (DSN) has no direct means of measuring geocentric station coordinates. While DSN baselines (*relative* coordinates) have been measured to better than 10 cm using very long baseline interferometry (VLBI), these coordinates can share much larger common geocentric coordinate bias errors. Global Positioning System (GPS) data, on the other hand, are very sensitive to the geocentric components which cannot be provided by VLBI. For Earth orientation, present-day calibrations for deep-space

[1] Consultant, Sterling Software, Pasadena, California.

tracking are provided from a combination of DSN VLBI quasar-observing sessions and non-DSN VLBI data which are provided on a "best-efforts" basis. Unfortunately, the DSN VLBI measurements require the use of large DSN antennas, which are unavailable for deep-space tracking or telemetry during these sessions. The incorporation of GPS data for deep-space Earth-orientation calibrations is motivated by a need to reduce the burden on DSN antennas for making Earth platform calibrations, thus freeing up blocks of antenna time for spacecraft tracking.

Recent GPS experiments have demonstrated baseline determination across continental plates to the centimeter level [1,2] and regional baseline determination to within a few millimeters' precision [3]. Analyses have indicated that GPS measurements can additionally provide subnanosecond global clock synchronization [4,5], subcentimeter-level media calibrations [6], and subdecimeter orbits of low Earth satellites [7]. In the past, two strategies were devised to determine the offset between the reference-frame origin and the geocenter using GPS observations [8]. Initial Casa Uno (1988) GPS solutions, with data from just seven satellites, agreed with satellite laser ranging (SLR) measurements to within 20–90 cm per geocenter component [9]. The first GPS International Earth Rotation Service (IERS) and Geodynamics (GIG '91) experiment (January–February 1991) provided a somewhat superior and more global data set. Preliminary GIG '91 GPS solutions for the geocenter agree with SLR to within 10–15 cm [15]. The improved accuracy was mainly due to the more uniformly distributed ground network equipped with better receivers and the greater number of GPS satellites in the constellation. As discussed in this article, the authors expect that GPS estimates for the geocenter will continue to improve to the few-centimeter level.

As the accuracy of GGP estimates improves, the capability to model their variation with time is enhanced. Particularly for Earth-orientation parameters, it is desirable to monitor subdaily variations with relatively short satellite-tracking data arcs in order to minimize systematic orbit-modeling errors. Past studies have shown that the addition of data from a low Earth satellite observing GPS to data collected from ground receivers improves geometrical strength for determining GPS orbits, clock-synchronization parameters, and ground baselines [4,10]. This article also investigates the expected enhancement in performance of GGP determination by the introduction of low Earth satellites. Because the orbital period of a low Earth satellite (typically 90–120 min) is much shorter than the GPS orbital period (12 hr) and the Earth-rotation period (24 hr), a GPS flight receiver on the low Earth satellite can track more GPS satellites than a ground receiver does in a shorter period of time. Precise GPS orbits can be obtained in shorter time and this, in turn, helps improve time resolution for estimation of Earth-orientation parameters. The flight GPS data also provide stronger correlation between GPS orbits. This in effect reduces the error in determining nonrotational coordinate parameters (such as the geocenter). For example, the U.S./French Ocean Topography Experiment (TOPEX/POSEIDON) satellite [11] launched on August 10, 1992, is the first low Earth satellite to carry on board a high-precision dual-frequency multichannel GPS receiver. A full (or nearly full) constellation of 21 GPS satellites is expected during the 3-year mission of TOPEX/POSEIDON. The orbital period of TOPEX/POSEIDON is far shorter (112 min) than that of GPS satellites (12 hr). These data should enhance sensitivity to subdaily Earth-rotation variations and could eventually improve tidal models by resolving subtle centimeter-level signatures in the Earth-orientation time series. The anticipated performance of GPS ground data with or without the GPS data from TOPEX/POSEIDON for determination of the GGP, which were modeled as random-walk parameters, has been studied. Future low Earth satellites, such as the Earth Observing System (EOS) platforms, Gravity Probe-B, and Aristoteles, will also carry onboard GPS flight receivers. It is very likely that there will be opportunities to include GPS data from more than one low Earth satellite for the determination of the GGP. To investigate the improved performance due to such a scenario, a covariance analysis using GPS data from two low Earth satellites in orthogonal orbital planes is also presented in this article.

II. GGP Variations

During the GIG '91 experiment, a global network of over 100 stations collected GPS measurements for a period of 3 weeks (January 22–February 13, 1991). Data from 21 Rogue GPS receivers were processed (Fig. 1) [12] using the Jet Propulsion Laboratory's GPS Inferred Positioning System (GIPSY) software. Fifteen GPS satellites were operational during this experiment. In this analysis, the coordinates of Goldstone (California), Kokee (Hawaii), and Kootwijk (The Netherlands) were fixed to align the coordinate system with the SV5 reference frame [13], which is based on a combination of SLR and VLBI observations. The origin and scale of the SV5 reference system is defined by the Center for Space Research (CSR)-8902-SLR [14], and its orientation is consistent with the 1989 International Terrestrial Reference Frame (ITRF '89). The estimated parameters include GPS orbits, nonfiducial ground-station position vectors, GGP (Earth orientation and geocenter location), random-walk zenith tropospheric delays for each station, and white-noise receiver/transmitter clocks

(Table 1). A priori random-walk constraints for the zenith tropospheric delay variation were 1.2 cm/$\sqrt{\text{hr}}$. The GPS solutions for the difference between Universal Time 1 and Universal Coordinated Time (UT1–UTC) and polar motion variations are relative to the *IERS Bulletin-B (B37 and B38)* nominal time series, which contains a smoothed time series from VLBI and SLR measurements separated by several days.

GPS pseudorange and the carrier-phase measurements over a period of time can be combined to solve for the satellite positions and velocities at an epoch. The geocenter is the origin of a satellite-based dynamical system. Any time-dependent origin shift can be estimated from the satellite tracking data as a common coordinate offset of all the ground station locations that realigns the dynamical origin and the origin of the coordinate system. Geocenter offsets with respect to the SV5 reference frame were estimated along with other parameters. The 3-week GIG '91 data were processed with 1-day and 2-day arcs, which gave very similar results. The mean of the daily GPS geocenter solutions differs from the SV5 value by $\Delta X = -8.3$ cm, $\Delta Y = 13.4$ cm, and $\Delta Z = -7.7$ cm [15]. The GPS single-day arc (daily) geocenter estimates had rms repeatabilities of 5 cm in X and Y and 30 cm in Z, with corresponding formal errors of 5 cm and 20 cm, respectively. The weaker result for the daily Z component estimates is due to several factors, such as the uneven distribution of ground sites with fewer sites in the southern hemisphere and polar regions and the incomplete GPS constellation which was available in 1991. During the International GPS Geodynamics Service 1992 campaign, 18 available GPS satellites were observed from 30 globally distributed Rogue receiver sites [16]. The geocenter estimates differ from the 1991 International Terrestrial Reference Frame (ITRF '91) by $\Delta X = 0.0 \pm 1.4$ cm, $\Delta Y = 1.5 \pm 1.3$ cm, and $\Delta Z = -8.2 \pm 3.0$ cm. The Z component still suffers from the lesser number of participating sites in the southern hemisphere. The improvement in results indicates better quality of data and more even distribution of global network sites.

Variability in Earth rotation is measured by estimation of UT1–UTC (typically in milliseconds). Changes in Earth rotation cause changes in UT1, while UTC is a fixed point of reference based on energy levels in the cesium atom. Well-modeled satellite dynamics are essential to detect time variations of UT1–UTC with satellite tracking data. For the GIG '91 experiment, independent GPS orbits were estimated every 24 hr, with new unconstrained GPS orbit parameters introduced at midnight every day. At noon every day (the midpoint of the GPS orbit solution arc), the UT1–UTC parameters were reset with a white process noise update in the filter. This offset between white noise resets for UT1–UTC and GPS orbital states is important to minimize the natural coupling between the orbital nodes and UT1. The results are presented and discussed in Lichten et al. [17]: The UT1–UTC daily estimates agree to about 0.04 msec rms with international radio interferometric surveying (IRIS)-intensive VLBI daily solutions (2-hr VLBI experiments using only two stations) and with the JPL Kalman Earth Orientation Filter daily solutions incorporating IRIS multibaseline VLBI data, the U.S. Naval Observatory's VLBI network (NAVNET) data, SLR polar motion, and NASA DSN VLBI data. The GPS daily estimate formal errors are between 0.01 and 0.02 msec ($\approx$1 cm), and the agreement between the GPS and VLBI estimates is consistent with the combined formal errors from the GPS and VLBI results. The GPS UT1 estimates also agree with the operational SLR estimates for UT1[2] with an rms difference of approximately 0.04 msec. Similarly accurate (1–2 cm) daily polar motion estimates have been reported with the same data set by Herring et al. [18] and by Lindqwister et al. [19].

Stochastic estimates of subhourly UT1–UTC and polar motion (UTPM) variations were also made using a factorized Kalman filter. The UT1–UTC variations were estimated with a random-walk model constrained at 2 msec/$\sqrt{\text{hr}}$. The polar motion variations (also random walk) were constrained at 0.6 cm $\sqrt{\text{hr}}$ ($\approx$3 cm over a day), but the initial overall a priori constraint was several meters, a fairly loose constraint. Note that these estimated variations correspond to corrections to the *IERS Bulletin B* Earth orientation time series. The UTPM solutions were obtained every 6 min with the GPS data. The UT1–UTC formal errors range from a few hundredths of a msec to nearly 0.1 msec near the end of the 24-hr period. Polar motion estimates show formal errors below 2 cm with the 24-hr solution. The formal errors do not reflect systematic effects caused by possible errors in fiducial station coordinates and some types of GPS orbit mismodeling. Blewitt et al. [20] demonstrate GPS–VLBI baseline agreement at the 1- to 2-cm level and a consistency with ITRF at the cm level. Even an error of 5 cm in each of the fiducial station coordinates would lead to less than a 0.01-msec error over 1 day in UT1–UTC changes. Figures 2 and 3 show a sample of smoothed estimates of UTPM fluctuations with 6-min time resolution. Figure 2 compares the UT1–UTC variation with the predictions from ocean tide models developed by Brosche et al. [21] and Herring and Dong [22] at every 6-min interval. Herring and Dong [22] used the approach of empirically fitting the major tidal

[2] Provided by R. Eanes, University of Texas, Austin, Texas, February 1992.

components to subdaily VLBI observations. The variations in UT1–UTC are clearly well correlated with models for UT1 variations from diurnal and semidiurnal oceanic tides [17]. The stochastic estimation of polar motion components (X_p, Y_p) fluctuations with 6-min time resolution are shown in Fig. 3. The observed time variations are being studied for further understanding. The influence of the atmosphere on rapid polar motion variations has been studied by Gross and Lindqwister [23], which shows that the atmosphere can play a major role in polar motion excitation to cause diurnal variations.

These preliminary GIG '91 results show that GPS tracking has the potential for high time-resolution estimation of GGP. It would be desirable, however, to further improve the strength of the ground GPS data to monitor changes in Earth orientation, especially UT1–UTC. In the following section, the use of additional data from low-Earth orbiters is discussed. While centimeter-accurate results have already been achieved for subdaily UT1–UTC variations and for daily polar motion variations, the data from GIG '91 did not have sufficient strength to achieve this level of accuracy on a *daily* basis for the geocenter. The International GPS Geodynamic Service (IGS) '92 campaign results, however, show a 5-cm-level offset between the GPS mean geocenter estimate and the ITRF '91 value [16]. There also remain some questions to be answered regarding the stability of the GPS daily time series for UT1 over periods of more than 1 week. The recent results reported by Freedman et al. [24] on subdaily Earth-rotation determination show highly encouraging results. In the next section, future expected performance of the GPS tracking system for GGP estimation with a full global ground network and a complete (24-satellite) GPS constellation augmented with GPS tracking data from one or more low Earth orbiters is examined.

III. Covariance Analysis

A covariance study was performed to evaluate anticipated speed and accuracy improvement in GGP estimation when the GPS data collected by a precise GPS flight receiver on board a low Earth satellite are combined with those from a global ground network. The TOPEX/POSEIDON satellite will provide such opportunities. A global network of 10 evenly distributed stations (Fig. 4) was used for the study. As in previous covariance studies [8], the fiducial baselines between NASA DSN sites at Goldstone, California; Madrid, Spain; and Canberra, Australia, were fixed as reference baselines in the estimation process. The data noise was assumed to be 20 cm in pseudorange and 0.4 cm in carrier phase at 5-min intervals (Table 2). Post-fit residuals from GIG '91 with Rogue GPS receivers were typically 20–30 cm for pseudorange and 0.3 cm for carrier phase at 6-min intervals (corrected for the ionosphere). Thus, the assumptions in this covariance analysis for data noise are consistent with present-day receiver performance. A full constellation of 24 GPS satellites distributed in 6 orbital planes was assumed for this study [25]. A full or nearly full GPS constellation is expected to be operational by mid-1993. The abundance of the GPS measurements allows rapid simultaneous estimation of the geocenter offset, polar motion, and changes in UT1–UTC, along with GPS orbits, ground stations, and other parameters. To allow for temporal variations, the tropospheric delays were modeled as random-walk parameters with the same constraints as are currently used when processing the real data.

The performance of low Earth satellites in enhancing the GGP estimation can be described in terms of three different cases. Case one includes only GPS measurement data from the 10 ground stations (Fig. 4). Case two includes GPS data from one low Earth orbiter and data from the ground sites. Case three analyzes the situation when GPS data from two low Earth orbiters are combined with the ground data. Both low Earth satellites were considered to be similar to TOPEX/POSEIDON, but in two orthogonal orbital planes with respect to each other with their ascending nodes separated by 180 deg. The temporal variations of polar motion and UT1–UTC were represented in this analysis by random-walk parameters [26]. The geocenter offset can still be treated as constant since it is unlikely to change significantly over short data spans (a few hours). Since GPS observations are sensitive to the rate of change of UT1–UTC rather than its absolute (inertially determined) offset, a perfect a priori value was assumed prior to the observation; the temporal UT1 variation was then estimated using the GPS tracking data. This assumption implies that VLBI is available to calibrate at least one UT1–UTC value in the GPS time series. Error in such a priori VLBI estimates of UT1–UTC would result in a common GPS node shift without affecting UT1–UTC *variation* determination. The random-walk constraint (1 sec) for polar motion was 2 cm/$\sqrt{\text{day}}$, and for UT1–UTC variations the constraint was 10 msec/$\sqrt{\text{day}}$ (Table 2).

In this covariance study, the effect of mismodeled dynamics on the TOPEX/POSEIDON spacecraft was reduced by adjusting its orbital elements along with a constrained three-dimensional fictitious force treated as process noise [7]. Adjustment of the fictitious force reduces the effects of mismodeled dynamics in the low Earth orbiter. The errors calculated also include systematic error due to mismodeled unadjusted parameters. These include 6 cm in each component of the fiducial baselines;

TOPEX/POSEIDON dynamic errors, which were quantified as 20 percent of the nominal values of solar radiation pressure, atmospheric drag, and albedo; and a gravity error, which was assumed to be 25 percent of the difference between two existing models, Goddard Earth Model (GEM)10 and GEML2 [27,28]. This gravity error is comparable to the error covariance of GEMT2 gravity [29].

IV. Comparison of GGP Estimation Errors

The future anticipated error in GGP as indicated by covariance analyses both with and without GPS measurements from TOPEX/POSEIDON is presented here. The expected errors from all three cases discussed in the previous section were calculated by using data spans of 2, 4, and 8 hr at 5-min intervals. The findings of the covariance analyses are strictly based upon the scenario as specified in Table 2. Although the assumptions made here are kept as close as possible to the real data processing, as in the case of GIG '91, there are still several differences. One of the differences in the covariance study was the scenario of fewer, more evenly distributed tracking sites to represent a routine operation rather than a dedicated experiment. A major difference is also due to the partial (2/3) GPS constellation operating during GIG '91 and somewhat incomplete tracking data from the southern hemisphere. For example, during GIG '91 one of the southern hemisphere sites which would have provided critical common view of GPS satellites was operating for less than 12 hr per day, which significantly degraded the accuracy of the Z component of the geocenter estimates. Because of these differences in the assumptions, a one-to-one comparison with the GIG '91 experiment results will not be meaningful. The main objective here is to demonstrate the extent of anticipated improvement in the determination of GGP variations when the GPS data from the low Earth satellites are introduced in the future, with the full GPS constellation.

Figure 5 presents only the X component of polar motion and geocenter offset; similar behavior was observed in other components as well. The comparison of the GGP estimates between case one, where no low Earth satellite was available, and case two, where data from one low Earth satellite were included, shows that TOPEX/POSEIDON helps to improve the accuracy and convergence speed by a factor of two, especially in the first 2–4 hr. The single dominating error source in the GGP variations estimates for a short data span, as shown in Fig. 5, is the GPS data noise. The introduction of the second low Earth satellite serves the similar purposes of improving the geometry and providing more measurements to combat this dominating effect of the data noise. As a result, further improvement can be expected both in accuracy and speed of convergence by another factor of two (compared to the case of one low Earth satellite) when GPS data from two low Earth satellites are included.

Some improvement in time resolution of GPS solutions could be achieved by simply adding more ground sites, as opposed to incorporating data from a low Earth orbiter. However, the two approaches are not equivalent. After a certain minimal global coverage is reached—for example, with 12 ground sites evenly distributed—additional sites tend to improve the formal estimation error proportionally with the square root of the additional number of measurements. Hence, a very large number of additional ground sites are needed to lead to a significant improvement, particularly in temporal resolution. The data processing also significantly increases in complexity when many ground sites are included. A more fundamental limitation of a ground-based-only approach, however, is the 12-hr period of the GPS satellites. Even with many more ground sites, ground-based carrier-phase data require several hours to determine a solution for the phase biases, due to the relatively slow GPS satellite velocities. The low Earth orbiter, however, provides rapid change in viewing geometries, completing its orbit approximately once every 90 min. This richness in geometrical coverage, rather than simply the additional measurements, provides the significant improvement in time resolution of the GPS data. Finally, the incorporation of Earth orbiters with periods much shorter than the 12-hr GPS orbital period may enable researchers to better separate semidiurnal geodetic signals and tidal effects from GPS orbit error, which may also resonate at 12-hr periods. Thus, there are certainly important advantages to adding low Earth GPS tracking data which cannot be achieved simply by adding additional ground sites. This has not been fully quantified in the analysis, but there are plans to study it in the future with GPS data from the TOPEX/POSEIDON satellite.

It appears that the speed and accuracy in the determination of the GGP can be enhanced by including the GPS data from one or more low earth satellites like TOPEX/POSEIDON. Consequently one should eventually be able to substantially better resolve centimeter- and subcentimeter-level signatures in the Earth-orientation time series over a few hours, which may be expected to appear due to tidal mismodeling or atmospheric effects.

V. Conclusions

Preliminary estimates of Earth-orientation variations and location of the geocenter using global GPS data have been obtained. The average geocenter offset estimate

from the 3-week-long GPS experiment agrees with the SLR determined value to better than 10 cm. The estimated UTPM variations show few-centimeter-level agreement with VLBI time series. A comparison with expected variations from diurnal and semidiurnal ocean tidal models suggests that the observed subdaily variations in UTPM with GPS are dominated by tidal effects. With the 1991 data, the GPS daily Earth-orientation estimates are accurate to the 1- to 2-cm level, and filtered solutions at 6-min intervals for stochastic UT1–UTC and polar motion are precise to the level of 2–4 cm. The preliminary results appear to be somewhat limited by the placement of the ground receivers and the fact that in 1991 only two-thirds of the GPS constellation was operational. However, the geocenter offset results from the IGS '92 campaign [16] have already shown 5-cm-level agreement with ITRF '91, and improved results have been reported for subdaily determination of the Earth's rotation [24]. Further improvement is expected, particularly for short-arc (daily or subdaily) geocenter estimation, as the full GPS constellation and evenly distributed ground network becomes available in the future.

The covariance studies indicate that temporal variations in global geodynamical parameters can be monitored with high precision using high-quality GPS pseudorange and carrier-phase data from a global ground network and a low Earth satellite. The expected errors in polar motion variations and the geocenter offset may be lowered to a few centimeters, and in UT1–UTC variations to a few hundredths of 1 msec, with only 4 hr of GPS data. The temporal resolution is expected to be a factor of two better with ground and low Earth GPS data as compared to ground-based tracking only. Including two low Earth orbiters placed in orthogonal orbital planes further improves the accuracy and time resolution (by an additional factor of two). The incorporation of low Earth orbiters with periods (90–120 min) much shorter than the 12-hr GPS orbital period may enable a better separation of diurnal and semidiurnal geodetic signals and tidal effects from GPS orbit error, which may also resonate at 12-hr periods. These important advantages of adding GPS data from low Earth satellites cannot be achieved simply by including more ground sites. Demonstration of the accuracy and resolution enhancements for subdaily (and even subhourly) Earth-orientation parameters predicted by these covariance analyses is planned when the GPS data from TOPEX/POSEIDON become available in the near future. Ultimately, such data may provide valuable information about tidal and other geophysical models which complement data from other techniques, such as VLBI and laser tracking.

Acknowledgments

The authors thank R. Eanes, who sent the SLR estimates for UT1–UTC to JPL. The oceanic tidal effects on UT1 were calculated by A. Freedman from the models cited in the text.

References

[1] M. Heflin, W. Bertiger, G. Blewitt, A. Freedman, K. Hurst, S. Lichten, U. Lindqwister, Y. Vigue, F. Webb, T. Yunck, and J. Zumberge, "Global Geodesy Using GPS Without Sites," *Geophysical Research Letters*, vol. 19, no. 2, pp. 131–134, January 24, 1992.

[2] S. M. Lichten and W. I. Bertiger, "Demonstration of Sub-Meter GPS Orbit Determination and 1.5 Parts in 10^8 Three-Dimensional Baseline Accuracy," *Bulletin Geodesique*, vol. 63, no. 2, pp. 167–189, 1989.

[3] U. J. Lindqwister, J. F. Zumberge, G. Blewitt, and F. H. Webb, "Few Millimeter Precision for Baselines in the California Permanent GPS Geodetic Array," *Geophysical Research Letters*, vol. 18, no. 6, pp. 1135–1138, June 1991.

[4] S. C. Wu, "The Role of Low Earth Orbiter in Intercontinental Time Synchronization Via GPS Satellites," *Proc. 5th Annual Precise Time and Time Interval (PTTI) Application and Planning Meeting*, Washington D.C., pp. 371–388, December 6–8, 1983.

[5] C. Dunn, S. M. Lichten, D. Jefferson, and J. S. Border, "Sub-Nanosec Clock Synchronization and Precision Deep Space Tracking," *Proceedings of Precise Time and Time Interval Conference (PTTI)*, Pasadena, California, pp. 89–102, December 3–5, 1991.

[6] S. M. Lichten, "Precise Estimation of Tropospheric Path Delays with GPS Techniques," *TDA Progress Report 42-100*, vol. October–December 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 1–12, February 15, 1990.

[7] T. P. Yunck, S. C. Wu, J. T. Wu, and C. L. Thornton, "Precise Tracking of Remote Sensing Satellites with the Global Positioning System," *IEEE Trans. Geoscience and Remote Sensing*, vol. 28, no. 1, pp. 108–116, January 1990.

[8] R. P. Malla and S. C. Wu, "GPS Inferred Geocenter Reference Frame for Satellite Positioning and Navigation," *Bulletin Geodesique*, vol. 63, no. 3, pp. 263–279, 1989.

[9] R. P. Malla and S. C. Wu, "The Geocenter Estimation Results Using GPS Measurements," *Proc. 2nd International Symposium on Precise Positioning With Global Positioning System*, GPS-90, Ottawa, Canada, pp. 286–295, September 3–7, 1990.

[10] T. P. Yunck, S. C. Wu, S. M. Lichten, W. I. Bertiger, U. J. Lindqwister, and G. Blewitt, "Toward Centimeter Orbit Determination and Millimeter Geodesy with GPS," *Proc. 5th Int. Symp. on Satellite Positioning*, Las Cruces, New Mexico, pp. 271–281, March 13–17, 1989.

[11] G. H. Born, R. H. Stewart, and C. A. Yamarone, "TOPEX—A Spaceborne Ocean Observing System," in *Monitoring Earth's Ocean, Land, and Atmosphere from Space—Sensors, Systems, and Applications*, edited by A. Schnapf, New York: AIAA, Inc., pp. 464–479, 1985.

[12] J. M. Srinivasan, T. K. Meehan, and L. E. Young, "Code and Codeless Ionospheric GPS Measurements with NASA's Rogue GPS Receivers," *Proceedings ION GPS-89 Conference, Institute of Navigation*, Colorado Springs, Colorado, pp. 451–454, September 27–29, 1989.

[13] M. H. Murray, R. W. King, and P. J. Morgan, "SV5: A Terrestrial Reference Frame for Monitoring Crustal Deformation With the Global Positioning System," *EOS Trans. AGU*, vol. 71, no. 43, p. 1274, October 23, 1990.

[14] B. E. Schutz, B. D. Tapley, R. J. Eanes, and M. M. Watkins, "Earth Rotation From Lageos Laser Ranging," *IERS Technical Note 2*, Observatoire de Paris, Paris, France, pp. 53–57, June 1989.

[15] Y. Vigue, S. M. Lichten, G. Blewitt, M. B. Heflin, and R. P. Malla, "Precise Determination of Earth's Center of Mass Using Measurements From the Global Positioning System," *Geophysical Research Letters*, vol. 19, no. 14, pp. 1487–1490, July 24, 1992.

[16] Y. Vigue, G. Blewitt, S. M. Lichten, M. B. Heflin, and R. J. Muellerschoen, "Recent High-Accuracy GPS Estimates of the Geocenter," *EOS Transactions, American Geophysical Union*, vol. 73, no. 43, p. 135, October 1992.

[17] S. M. Lichten, S. L. Marcus, and J. O. Dickey, "Sub-Daily Resolution of Earth Rotation Variations with Global Positioning System Measurements," *Geophysical Research Letters*, vol. 19, no. 6, pp. 537–540, March 20, 1992.

[18] T. A. Herring, D. Dong, and R. W. King, "Sub-Milliarcsecond Determination of Pole Position Using Global Positioning System Data," *Geophysical Research Letters*, vol. 18, no. 10, pp. 1893–1896, October 1991.

[19] U. J. Lindqwister, A. P. Freedman, and G. Blewitt, "Daily Estimates of the Earth's Pole Position with the Global Positioning System," *Geophysical Research Letters*, vol. 19, no. 9, pp. 845–848, May 4, 1992.

[20] G. Blewitt, M. B. Heflin, F. H. Webb, U. J. Lindqwister, and R. P. Malla, "Global Coordinates With Centimeter Accuracy in the International Terrestrial Reference Frame Using GPS," *Geophysical Research Letters*, vol. 19, no. 9, pp. 853–856, May 1992.

[21] P. Brosche, U. Seiler, J. Sundermann, and J. Wunsch, "Oceanic Tide Effects in Universal Time Detected by VLBI," *Astron. Astrophys*, vol. 245, no. 2, pp. 676–682, May 1991.

[22] T. A. Herring and D. Dong, "Current and Future Accuracy of Earth Orientation Measurements," *Proceedings of the AGUE Conference on Geodetic: Monitoring Global Change* (ON Technical Report No. 137NGS49), pp. 306–324, 1991.

[23] R. S. Gross and U. J. Lindqwister, "Atmospheric Excitation of Polar Motion During the GIG '91 Measurement Campaign," *Geophysical Research Letters*, vol. 19, no. 9, May 4, 1992.

[24] A. P. Freedman, Y. Vigue, R. Ibanez-Meier, R. J. Muellerschoen, S. L. Marcus, S. M. Lichten, J. O. Dickey, and U. J. Lindqwister, "Sub-Daily Earth Rotation With GPS: Recent Results," *EOS Trans. American Geophysical Union*, vol. 73, no. 43, p. 135, October 1992.

[25] G. B. Green, P. D. Massatt, and N. W. Rhodus, "The GPS 21 Primary Satellite Constellation," *Proc. Satellite Div. Int. Tech. Meeting*, Inst. Navigation, pp. 15–26, September 1988.

[26] A. P. Freedman, "Measuring Earth Orientation With the Global Positioning System," *Bulletin Geodesique*, vol. 65, no. 1, pp. 53–65, 1991.

[27] F. J. Lerch, S. M. Klosko, R. E. Laubscher, and C. A. Wagner, "Gravity Model Improvement Using Geos 3 (GEM 9 and 10)," *J. Geophys. Res.*, vol. 84, no. B8, pp. 3897–3916, July 1979.

[28] F. J. Lerch, S. M. Klosko, R. E. Laubscher, and C. A. Wagner, "A Gravity Model for Crustal Dynamics (GEM-L2)," *J. Geophys. Res.*, vol. 90, no. B11, pp. 9301–9311, September 1985.

[29] J. G. Marsh, F. J. Lerch, B. H. Putney, T. L. Felsentreger, B. V. Sanchez, S. M. Klosko, G. B. Patel, J. W. Robbins, R. G. Williamson, T. L. Engelis, W. F. Eddy, N. L. Chandler, D. S. Chinn, S. Kapoor, K. E. Rachlin, L. E. Braatz, and E. C. Pavlis, "The GEM-T2 Gravitational Model," *J. Geophys. Res.*, vol. 95, no. B13, pp. 22043–22071, December 10, 1990.

Table 1. Estimation strategy applied to the GIG '91 data processing.[a]

Parameters	Treatment	A priori σ
GPS orbits	Adjusted	Unconstrained
G_y (Y bias)	Adjusted	2×10^{-12} km/sec^2
G_x and G_z	Adjusted	100 percent (scale factor)
GPS carrier biases	Adjusted (real-valued)	Unconstrained
3 fiducial site locations	Fixed at a priori coordinates	—
18 nonfiducial site locations	Adjusted	Unconstrained
Tropospheric delay (1 per site)	Adjusted (as random walk)	50 cm + 1.2 cm/$\sqrt{\text{hr}}$ (6 cm over 1 day)
Clock biases (stations/transmitters)	Adjusted (as white noise)	Unconstrained
Geocenter offset	Adjusted	Unconstrained
UT1–UTC variations	Adjusted	2 msec/$\sqrt{\text{hr}}$ (10 msec over 1 day)
(X_p, Y_p) pole position	Adjusted	0.6 cm/$\sqrt{\text{hr}}$ (3 cm over 1 day)
Gravity	Fixed at [GEM-T2 (12 x 12)] value	—
Earth's GM	Fixed at IERS value	—

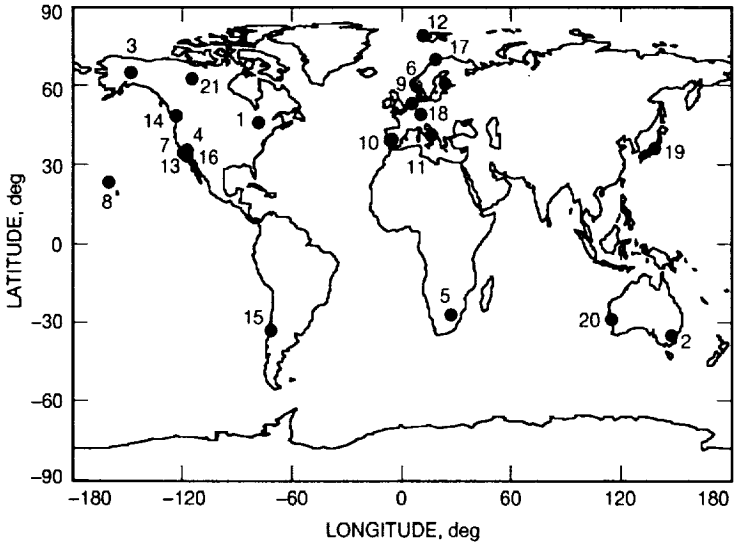
[a] Note that the nominal models for G_x, G_y, and G_z were believed to be accurate to about 10 percent. Therefore, the a priori constraints used are fairly loose.

Data span = 24 hr
Data interval = 6 min
Data noise = 100-cm pseudorange
= 1-cm carrier phase
GPS constellation = 15 satellites

Table 2. Basic estimation models applied to the covariance analysis.[a]

Parameters	Treatment	A priori σ
GPS orbits	Adjusted	Unconstrained
G_y (Y bias)	Adjusted	2×10^{-12} km/sec^2
G_x and G_z	Adjusted	100 percent (scale factor)
GPS carrier biases	Adjusted	Unconstrained
TOPEX/POSEIDON orbits	Adjusted	Unconstrained
Fictitious force on TOPEX/POSEIDON	Adjusted (as process noise)	0.5 μm/sec^2 bias; 0.35 μm/sec^2 batch-to-batch
Gravity on TOPEX/POSEIDON	Considered	25 percent (GEM10–GEML2)
Solar pressure on TOPEX/POSEIDON	Considered	20 percent of nominal model
Albedo TOPEX/POSEIDON	Considered	20 percent of nominal model
Atmospheric Drag TOPEX/POSEIDON	Considered	20 percent of nominal model
3 fiducial site locations	Considered	6 cm each component
7 nonfiducial site locations	Adjusted	20 cm each component
Tropospheric delay (1 per site)	Adjusted (as random walk)	40 cm + 12 cm over 1 day
Clock biases (stations/transmitters)	Adjusted (as white process noise)	Unconstrained
Geocenter offset	Adjusted	20 m each component
UT1–UTC variations	Adjusted	10 msec over 1 day
(X_p, Y_p) pole position	Adjusted	2 cm over 1 day
Earth's GM	Adjusted	2 parts in 10^8

[a] Tracking network = 10 sites (cf. Fig. 4)
Data span = 2–8 hr
Data noise = 20-cm pseudorange
= 0.4-cm carrier phase
Data interval = 5 min
GPS constellation = 24 satellites
Cutoff elevation = 10 deg (ground receivers)
= 0 deg (TOPEX/POSEIDON receiver)
Number of GPS observed = Up to 8 at a time (ground receivers)
= Up to 6 at a time (TOPEX/POSEIDON receiver)



1. ALGONQUIN, CANADA
2. CANBERRA, AUSTRALIA
3. FAIRBANKS, ALASKA
4. GOLDSTONE, CALIFORNIA
5. HARTEBEESTHOEK, SOUTH AFRICA
6. HONEFOSS, NORWAY
7. JPL, PASADENA, CALIFORNIA
8. KOKEE, HAWAII
9. KOOTWIJK, THE NETHERLANDS
10. MADRID, SPAIN
11. MATERA, ITALY
12. NY ÅLESUND, NORWAY
13. PINYON, CALIFORNIA
14. VICTORIA, CANADA
15. SANTIAGO, CHILE
16. LA JOLLA, CALIFORNIA
17. TROMSO, NORWAY
18. WETTZEL, GERMANY
19. USUDA, JAPAN
20. YARRAGADEE, AUSTRALIA
21. YELLOWKNIFE, CANADA

Fig. 1. The 21 GPS Rogue receiver sites of the GIG '91 network.

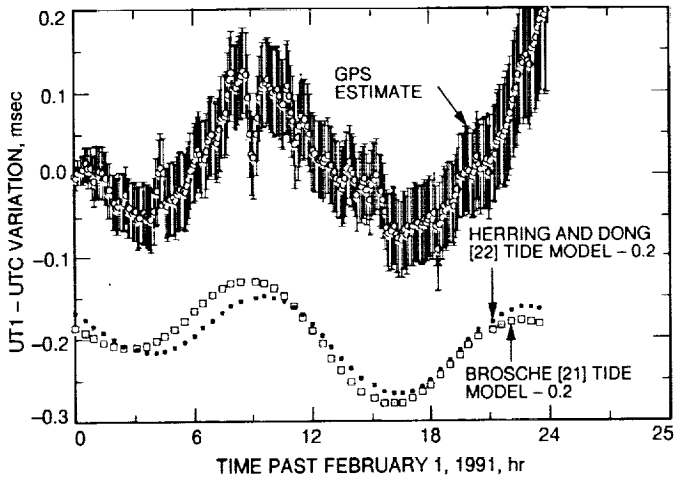


Fig. 2. Variations in UT1–UTC estimated from GPS relative to the *IERS Bulletin B* nominal series. One-sigma GPS formal errors are also plotted. Comparison with oceanic tidal model predictions from Herring and Dong [22] and from Brosche et al. [21] (plotted with an offset of 0.2 msec for clarity in the figure) shows clearly the effect of tides detected in the GPS 6-min estimates.

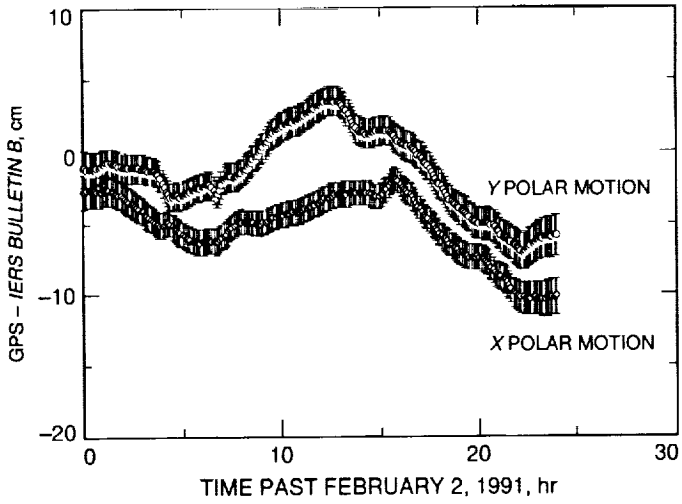


Fig. 3. GPS stochastic estimation of polar-motion fluctuations with 6-min time resolution. GPS formal one-sigma errors are also plotted.

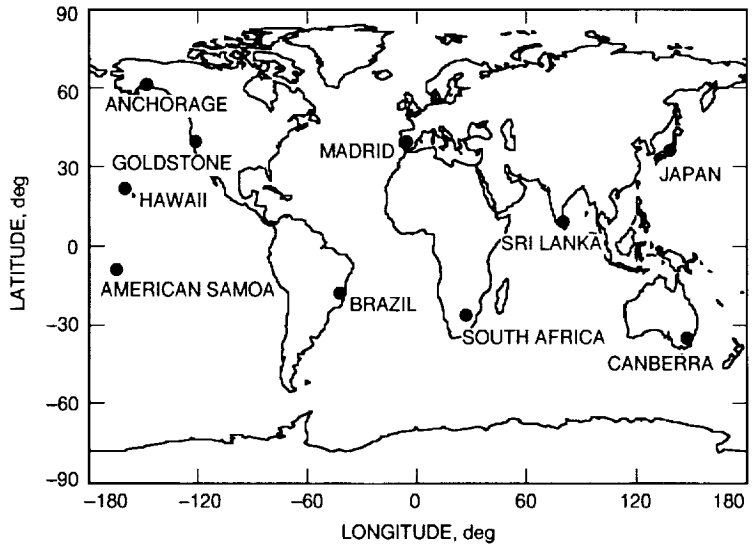


Fig. 4. Global GPS tracking network for covariance analysis. All the sites are assumed to be occupied by GPS Rogue-type receivers tracking GPS satellites simultaneously with high-precision dual-frequency multichannel flight receivers on board one or two low Earth satellites.

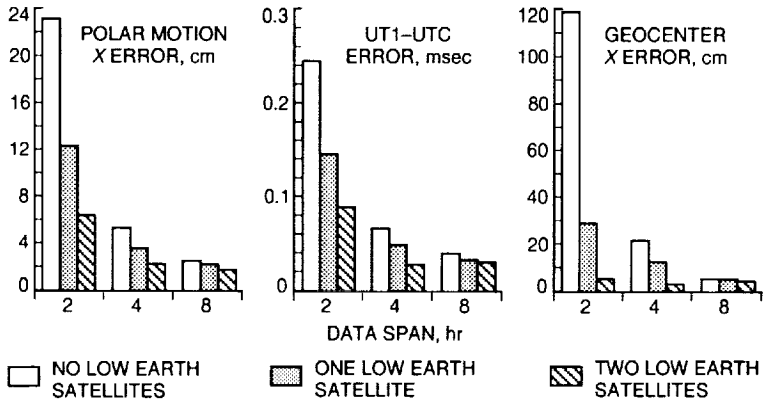


Fig. 5. Anticipated errors in GGP using global network with precision GPS data. The *X* components of polar motion and geocenter offset are presented here; the other components vary similarly.

Failure Modes of Reduced-Order Orbit Determination Filters and Their Remedies

D. J. Scheeres
Navigation Systems Section

This article discusses ways in which failure can occur in reduced-order, orbit-determination filter, error covariance calculations. In the context of this article, reduced-order filters denote nonoptimal filters which include fixed levels of uncertainty in some parameters of the measurement models or in the spacecraft dynamical model which are not explicitly estimated in the filter equations. Failure is defined herein as an increase in the orbit determination covariance with the addition of data or as an unreasonable growth in the covariance with time, i.e., nonasymptotic behavior of the covariance. Some simple, known cases of failure are discussed along with their traditional remedies. In addition, more modern remedies are discussed which are currently under development at the Jet Propulsion Laboratory. The article first describes the known problems of reduced-order filters when they are employed for orbit determination, and their traditional remedies. Then, having defined these, the relevancy and desirability of the more modern remedies are made apparent.

I. Introduction

As the terminology used for reduced-order filters, which are sometimes called consider filters, is not uniform, it is important that the definition of the consider filter used herein be understood from the outset. What is called the consider filter is actually known as the consider option in JPL's orbit determination software system. In this approach, which is commonly used in spacecraft navigation at JPL, the relevant estimated state parameters (spacecraft position and velocity, etc.) are determined using only knowledge of other estimated state parameters; the filter is closed upon itself. It is acknowledged, however, that there are other parameters which may not be known exactly and which have an effect on measurements made of the state and on the dynamics of the state.

For a variety of reasons it is sometimes not desired to estimate these parameters. These reasons may be that there are no adequate models for the evolution of these parameters, that estimating them would require excessive computation time, or that, if estimated, the computed uncertainty in these parameters would be reduced far below

the level warranted by model accuracy. Instead, in order to gauge how uncertainties in these parameters manifest themselves as uncertainties in the state estimates, the estimation error covariance of the state (the computed covariance) is modified to account for these errors in some way. It should be noted that this modified covariance, called the consider covariance, is not fed back into the filter. Thus the filter knows nothing about the consider contribution to the uncertainty in the state estimate. The mathematical definition of these filters is reviewed later in the article.

In this article a consider filter is defined to have failed when the addition of data yields an increase in the consider covariance, and hence less confidence in the estimate, or when the consider covariance begins to grow unreasonably large as time increases. These types of failure are usually alleviated by artificially increasing, or deweighting, the measurement noise variances of one or more data types so that additional data yield a smaller or equivalent consider covariance. Data weights are, ideally, the inverse of the noise variance for the measurements being employed. They also can serve, however, as free parameters with which one may emphasize or deemphasize certain data types. If this does not alleviate the failure, then the data arc may be restarted, essentially discarding old data. Two general modes of failure for consider filters are discussed. In analyzing these modes it becomes evident why and how a filter should be tuned. The ramifications of a consider filter that needs substantial tuning are also discussed. It is hoped that the reader can develop an intuitive feel for the modes of failure of consider filters from this article, and that it will shed light on some of the seemingly nonintuitive behavior that can be exhibited by this type of estimation scheme.

A special effort is made to discuss a more sophisticated sequential orbit-determination filter model being developed at JPL, the Enhanced Filter [1,2]. The use of this filter alleviates many problems associated with more traditional consider filters. At various points in this article mention is made of this approach.

II. Modes of Failure

The modes of failure discussed below correspond to two extreme conditions in which a reduced-order filter may operate. It is hypothesized that generic cases of failure are a combination of these modes. There may exist, of course, possible modes not discussed here. In the first case, process noise is not assumed to be present in the filter model. This is called the batch mode. This case corresponds to a simplified covariance analysis or, in a limiting case, to a situation where the frequency of the data measurements is large enough (or the characteristic time of the process noise is long enough) to render the process noise component essentially constant. In this case, the filter is modeled as a batch filter. The second mode is the converse of the first, namely, it is the case where process noise effects dominate the filter behavior. This is called the sequential mode. In this case, the filter is said to be saturated, and the uncertainty in the filter parameters has reached a quasi-steady state. The filter is modeled as a sequential Kalman–Bucy filter in this case.

III. Batch Mode

To begin, the classical results of batch covariance analysis with consider parameter effects are restated.

A. Batch Consider Covariance Definitions

Assume the state model to be

$$x_k = \Phi(t_k, t_{k-1})x_{k-1} \tag{1}$$

and the measurement model to be

$$z_k = H_{x_k}x_k + \nu_k \tag{2}$$

with $E(\nu_i\nu_j^T) = R_i\delta_{ij}$. The term x_k is the state at time t_k; $\Phi(t_k, t_{k-1})$ is the state transition matrix; z_k is the measurement at time t_k; ν_k is the measurement noise; R_i is the variance of the measurement noise, and δ_{ij} is the usual Kronecker delta function. Then the classical approach to covariance computation, in the absence of process noise, can be stated as [3, p. 27]:

$$P_k = \left[\sum_{j=0}^{N} \Lambda(t_j, t_k)\right]^{-1} \tag{3}$$

where P_k is the state-estimation error covariance matrix at time t_k, the $\Lambda(t_j, t_k)$ are information matrices propagating measurements taken at time t_j to time t_k, and N is the total number of measurements. Essentially, the filter reduces all measurements to the same epoch and then applies a weighted least-squares estimation process. Here the time at $j = 0$ may either be an initial measurement or an a priori information matrix.

Explicitly, the information matrices are defined as

$$\Lambda(t_j, t_k) = \Phi^T(t_j, t_k) H_{x_j}^T R^{-1} H_{x_j} \Phi(t_j, t_k) \qquad (4)$$

where R is the covariance of the measurements at time t_j (often assumed to be diagonal and constant), H_{x_j} is the partial derivative of the observation vector with respect to the state at time t_j, and $\Phi(t_j, t_k)$ is the state transition matrix which describes the linear evolution of state perturbations about the nominal trajectory from time t_k to t_j.

Note that the information matrices $\Lambda(t_j, t_k)$ are at least positive semidefinite, and if the proper types of measurements (including possible a priori information) are taken, the sum of the information matrices will be positive definite for some $j \geq 0$; thus the inverse of the sum will exist if enough measurements are taken. As the information matrices are positive semidefinite, it is obvious that additional measurements will always cause the size of P_k to decrease or remain constant. Thus, to minimize the trace of the P_k matrix, or $tr(P_k)$, a common measure of the size of the covariance matrix, it is useful to increase the number of measurements N over any given data arc. Due to the nature of the dynamics of the state, the size of the information matrix will often vary with $\mid t_j - t_k \mid$ depending on the specific dynamics of the state (governed by the state transition matrices).

A consider parameter is a parameter whose exact value is not known, which has some influence on the measurements or dynamical environments of the state, and is not estimated. If such parameters are present, then the covariance of the state must be augmented to account for the uncertainty of the consider parameters. The resultant covariance matrix is called the consider covariance[1]:

$$P_{c_k} = P_k + S_k \Pi S_k^T \qquad (5)$$

where P_k is the computed state covariance matrix as defined above, Π is the covariance matrix of the consider parameters, and S_k is the sensitivity matrix. The sensitivity matrix can be defined as the partial derivative of the estimated state with respect to the consider parameters. For this discussion, it is useful to view it in a more elemental form:

$$S_k = -P_k \sum_{j=0}^{N} \Phi^T(t_j, t_k) H_{x_j}^T R^{-1} \left[H_{x_j} \theta(t_j, t_k) + H_{c_j} \right] \qquad (6)$$

The repeated quantities are defined as before. The new quantities are defined as follows: H_{c_j} is the partial derivative of the observation vector with respect to the consider parameters at time t_j, and $\theta(t_j, t_k)$ represents the influence that the consider parameters have on the evolution of the state. The term $\theta(t_j, t_k)$ is defined as the particular solution to the nonhomogeneous equation

$$\dot{\theta}(t, t_k) = A(t)\theta(t, t_k) + B(t) \qquad (7)$$

$$\theta(t_k, t_k) = 0 \qquad (8)$$

where θ is a matrix of size $n \times p$ where n is the number of elements in the state and p is the number of consider parameters. $A(t)$ is the partial derivative of the equations of motion with respect to the state and $B(t)$ is the partial derivative of the equations of motion with respect to the consider parameters. The solution of this equation expresses the influence the consider parameters have on the state, and can be stated in terms of the state transition matrix:

$$\theta(t, t_k) = \int_{t_k}^{t} \Phi(t, \tau) B(\tau) d\tau \qquad (9)$$

$$\theta(t_k, t) = -\Phi(t_k, t)\theta(t, t_k) \qquad (10)$$

Thus the sensitivity matrix may be expressed as

$$S_k = M_k + \bar{\theta}_k \qquad (11)$$

where two new quantities are defined. These are

$$M_k = -P_k \sum_{j=0}^{N} \Phi^T(t_j, t_k) H_{x_j}^T R^{-1} H_{c_j} \qquad (12)$$

$$\bar{\theta}_k = \left[\sum_{j=0}^{N} \Lambda(t_j, t_k) \right]^{-1} \sum_{j=0}^{N} \Lambda(t_j, t_k) \theta(t_k, t_j) \qquad (13)$$

The term M_k denotes the effect of the consider parameter uncertainty as the uncertainty acts through the measurement alone; this term is called the measurement

[1] S. C. Wu, W. I. Bertiger, J. S. Border, S. M. Lichten, R. F. Sunseri, B. G. Williams, P. J. Wolff, and J. T. Wu, *OASIS Mathematical Description, Version 1.0*, JPL D-3139, Jet Propulsion Laboratory (internal document), Pasadena, California, pp. 6–11, April 1, 1986.

consider contribution. The effect of this term is usually to add a nearly constant uncertainty to the consider covariance. Its effect may dominate over $\bar{\theta}_k$ in many situations where the consider parameters do not enter strongly into the state dynamics, yet this term will in general not lead to unbounded growth in the consider covariance.

The matrix $\bar{\theta}_k$ represents the effects of the consider parameters on the state dynamics. This term may lead to covariance divergence and some of the counterintuitive orbit determination results which are associated with the use of consider filters. As the information matrices Λ are positive semidefinite, they may be viewed as analogues to masses or to weights, hence $\bar{\theta}_k$ may be viewed as analogous to the center of mass of the accumulated information matrix:

$$\left[\sum_{j=0}^{N} \Lambda(t_j, t_k)\right] \bar{\theta}_k = \sum_{j=0}^{N} \Lambda(t_j, t_k)\theta(t_k, t_j) \qquad (14)$$

Alternatively it may be viewed as an average or mean of the function $\theta(t_k, t_j)$ weighted by the information matrices. This is especially clear if the number of measurements becomes large, leading to

$$\bar{\theta}_k \approx \left[\int_{t_0}^{t_N} \Lambda(\tau, t_k)d\tau\right]^{-1} \int_{t_0}^{t_N} \Lambda(\tau, t_k)\theta(t_k, \tau)d\tau \qquad (15)$$

which is a classical definition of averaging. The term $\bar{\theta}_k$ is called the dynamic consider contribution.

B. General Properties of the Sensitivity Matrix

The sensitivity matrix with its measurement and dynamic contributions has some properties that may be stated without recourse to specific example. Some of these are listed below.

1. Performance Characterization. The performance of the consider covariance may be characterized by $tr(P_{c_k}) = tr(P_k) + tr(S_k \Pi S_k^T)$. If the consider parameter covariance Π is diagonal, as is often the case, then $tr(S_k \Pi S_k^T) = \sum_{i=1}^{p} \Pi_i \|S_{k_i}\|^2$, where $\|-\|$ is the usual Euclidian norm, Π_i is the ith consider parameter variance, and S_{k_i} is the ith column of the sensitivity matrix. As noted previously, $tr(P_k)$ always decreases with additional data. As will be shown, $\|S_k\|$ may increase with additional data, implying that the consider filter has failed. Note that in this formulation (diagonal Π matrix) the consider parameters may be discussed independent of each other.

2. Data Weighting Properties. A classical result associated with the sensitivity matrix is its invariance with respect to a common scaling of the weights (noise variances) of all data types. Note that in the definition of both the measurement and dynamic consider contributions, the inverse of the data noise matrix, R^{-1}, appears in both the numerator and denominator (or their matrix generalizations). Thus, should the data noise matrix R be scaled by some factor, the sensitivity matrix will be invariant with respect to this scaling. This implies that if there is only one data type being used, the consider contribution is independent of the data noise (assuming no a priori information). Thus for single-data-type consider filters the tuning strategy of artificially deweighting the data cannot be used to alter the sensitivity matrix, and data editing must be used instead. This will not be the case when process noise is added to the filter.

If there are two or more data types present, then the consider contribution may be altered by scaling the data weights relative to each other. This may be a useful procedure if one data type induces a large consider contribution with respect to the other data types. Then by deweighting that one data type, the consider contribution of this data type is reduced relative to the other data types. Unfortunately, in performing the deweighting, the computed covariance of the state will grow. If the data type in question is essential to the performance of the filter, then deweighting that data type may lead to poorer performance.

3. Measurement Consider Contribution. The measurement contribution M_k [Eq. (12)] is related to the uncertainty of the measurements with respect to the consider parameters. It has no terms that generically increase with time. While its magnitude may be large for effects such as the tropospheric-calibration-error or station-location-error contribution for the radio Doppler data type, or a range bias uncertainty for the radio range data type, it tends to remain at a fixed level.

4. Dynamic Consider Contribution. The dynamic contribution $\bar{\theta}_k$ [Eq. (13)] is related to the uncertainty induced in the state dynamical model by the consider parameters. As discussed earlier, it may be viewed as an average of the quantities $\theta(t_k, t_j)$ weighted by their respective information matrices $\Lambda(t_j, t_k)$. Recalling the definition of the θ [Eq. (9)], note that it is an integral over the time interval $t_k - t_j$. Hence, if this time interval grows large (and the information content of the matrix Λ does not decrease as the time from the measurement increases) these

individual terms may grow large. Thus, if conditions are right, the mean value of these terms will begin to grow, leading to an increase in the size of the consider covariance if the computed covariance does not decrease swiftly enough. Some specific instances where such a failure may occur are discussed in the next subsection.

C. Some Batch Failure Modes and Their Remedies

The usual remedy for a large consider covariance, as discussed above, is to deweight the data type with the largest consider contribution. This approach will usually allow one to reduce the consider covariance somewhat, yet may not always be the most insightful remedy. It also requires tuning by the navigator, often a difficult and imprecise procedure.

1. High Information Content Divergence. One particular failure mode often encountered in practice is due to measurement data with both a dynamic consider contribution and a large information content. The immediate result of such data is to improve the computed covariance and to have only a small effect on the dynamic consider contribution, as the integral time spans will be short. If the subsequent data have a smaller information content, then the averaging process in computing the dynamic consider contribution will emphasize the terms related to the large-information-content data. As time progresses, these terms will grow at least linearly in time due to the integration effects and may cause the consider covariance to grow in an unbounded manner.

A simple example of this is a planetary (or planetary satellite) flyby navigation scenario, in which the ephemeris errors of the planet are considered and not estimated. While the spacecraft is flying by the planet the information content of data tends to be very large due to the increased dynamics of the spacecraft. However, the considered position errors of the planet will translate directly into an uncertainty in the model of forces acting on the spacecraft. At closest approach the dynamic consider contributions may be negligible, but soon after they will grow with time, as the post-flyby data will contain relatively less information.

The usual remedy to this failure mode is to restart the data arc using the minimized consider covariance at (or soon after) flyby as the a priori covariance matrix. Another remedy is to use the flyby information to estimate the planet position and reduce the otherwise fixed uncertainty in these parameters.

2. Flat-Information-Content Divergence. Another failure mode may be identified with information matrices having a uniformly constant information content. In this instance the dynamic consider contribution weights all its terms equally. Thus the older data will begin to dominate the mean as they grow at least linearly with time.

An example of this effect would be a spacecraft in an outer solar system cruise period, where the nongravitational accelerations acting on the spacecraft are treated as consider parameters. In such a system, without new data with a larger information content to increase the relative size of the information matrices, the consider covariance will grow.

There are several remedies to this mode of failure, the simplest being to discard old data after some time span. Unfortunately, this remedy also discards potentially useful information. Remedies that are more modern are either to model nongravitational accelerations as process noise terms in the dynamics, which forces the filter to give new information relatively greater weight, or to estimate the deterministic model parameters which represent the nongravitational accelerations.

D. Implications of Failure Modes

The fact that, in some cases, the sensitivity matrix may become large due to either the measurement or dynamic contribution indicates something more: that there is a significant information content in the measurements concerning the state estimate which could be exploited if it were not for the effects of the consider parameters. In fact, this is often why these parameters were considered in the first place, as they were estimated too quickly and too well with conventional filters and simplified models.

Some recent investigations into improved filter modeling have turned this fact into a useful strategy [1,2]. This improvement entails building a more sophisticated model for the consider parameter effects, especially the random components, and estimating this expanded parameter set, most of the members of which are treated as stochastic parameters. If appropriate noise variances and correlation times are chosen, the estimation process may not yield significant improvement in the knowledge of these parameters, yet may extract much of the information contained in the measurement partials of these parameters.

That the stochastic parameter uncertainties do not improve significantly in this approach is not an inherent problem, and can be understood in the following context. It is usually the case that these consider parameters are better determined by various intensive, off-line calibration techniques. With regard to station locations, for

example, these techniques would take the form of many careful very long baseline interferometry measurements of quasars. Thus, if the current filter's information matrix with respect to these parameters is weighted against the accumulated data (which defined the a priori uncertainties in these parameters), the current information will have a negligible impact on the uncertainty. The introduction of stochastic parameters to account for modeling imperfections serves as a way to introduce such effects while exploiting the current measurement information to improve the state covariance.

IV. Sequential Mode

Now the case where process noise has been added into the filter is discussed. This alters the covariance computation scheme to a Kalman–Bucy approach. The basic definitions for this type of computation are briefly reviewed.

A. Sequential Consider Covariance Definitions

Assume that the state model is given by

$$x_k = \Phi_{k,k-1} x_{k-1} + \Gamma_{k,k-1} u_{k-1} \tag{16}$$

and that the measurement model is the same as in the batch mode [Eq. (2)]. Then computation of the state covariance when process noise is present is defined as follows:

$$P_k = [I - K_k H_{x_k}] \bar{P}_k \tag{17}$$

$$K_k = \bar{P}_k H_{x_k}^T \left[H_{x_k} \bar{P}_k H_{x_k}^T + R_k \right]^{-1} \tag{18}$$

The quantity P_k is the state (or computed) estimate error covariance computed at time t_k as before. Note that Eq. (17) is only valid when the gain K_k is the optimal gain matrix. The quantity H_{x_k} is the partial derivative of the observation vector at time t_k with respect to the state. The quantity K_k is the Kalman gain matrix computed at time t_k, and R_k is the measurement noise covariance. Finally, the matrix $\bar{P}_k$ is the computed covariance from time t_{k-1} mapped to the present time. Involved in the mapping are both the dynamics of the state and the uncertainty due to process noise in the intervening period:

$$\bar{P}_k = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + \Gamma_{k,k-1} Q_{k-1} \Gamma_{k,k-1}^T \tag{19}$$

Equations (17), (18), and (19) define the matrix Riccati equation for the discrete filter [4]. In Eq. (19) the matrix $\Phi_{k,k-1}$ is the state transition matrix from time t_{k-1} to time t_k, and P_{k-1} is the computed covariance at time t_{k-1}. The matrix $\Gamma_{k,k-1}$ maps the process noise state uncertainty from time t_{k-1} to time t_k. The matrix Q_{k-1} is generally the covariance of a Gaussian white noise sequence u_i where

$$E\left[u_i u_j^T\right] = Q_i \delta_{ij} \tag{20}$$

The term δ_{ij} is the usual Kronecker delta function. The process noise model may also be correlated in time.

The definition of the $\Gamma_{k,k-1}$ matrix is very similar to the $\theta(t_k, t_{k-1})$ function defined previously. Explicitly

$$\Gamma_{k,k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) B_n(\tau) d\tau \tag{21}$$

In this equation $B_n(\tau)$ is the partial derivative of the state dynamics with respect to the process noise parameters.

As before, the consider covariance at time t_k is simply

$$P_{c_k} = P_k + S_k \Pi S_k^T \tag{22}$$

The formula for the covariance P_k is given by Eqs. (17), (18), and (19). The definition of the sensitivity matrix is now given as [3, p. 177]

$$S_k = [I - K_k H_k] \bar{S}_k - K_k H_{c_k} \tag{23}$$

$$\bar{S}_k = \Phi_{k,k-1} S_{k-1} + \theta_{k,k-1} \tag{24}$$

In the above formulae, H_{c_k} is the partial derivative of the observation vector with respect to the consider parameters at time t_k, and $\theta_{k,k-1}$ represents the sensitivity of the spacecraft dynamics due to possible errors in the consider parameters. Both of these terms were defined in the previous section. Note that there is not uniform agreement on the calculation of $\bar{S}_k$. Slightly different forms for this mapping may be derived depending on whether the consider covariance is first computed and then propagated forward in time, or the state estimates are first mapped in time and then the consider covariance is computed. The reason for this difference, which does not exist for the computed covariance, is due to the peculiarities of the consider covariance when mapped. This article follows Ref. [3] and Footnote 1 and uses the first definition given, that the consider covariance is first computed and then mapped. This

definition is more in line with the use of the Kalman–Bucy filter. Note that if no process noise is present in the dynamics ($Q_i \equiv 0$), then the above definitions are equivalent to the definitions given in Section III.A.

B. General Properties of the Filter and Sensitivity Matrix

The Kalman–Bucy sequential filter with process noise exhibits some significant differences in behavior versus the batch filter without process noise. The major difference is that in the sequential filter old data are forgotten as time goes on, as the accumulated effect of process noise is to render measurements made some time ago not valuable. This occurs because the current covariance is continually made relatively larger by the process noise terms in the filter model, and hence the gain matrix K is increased to rely more heavily on the current data.

Some implications of this forgetting process for the sensitivity matrix follow. First of all, this process implies that the sensitivity matrix is no longer invariant under uniform scaling of measurement noise. Thus, data may be deweighted to influence the sensitivity matrix, even if only one data type is used. This was not the case with the batch mode. Another implication is that the dynamic consider contribution, which could lead to unbounded growth of the consider covariance in the batch mode, has a reduced effect. This is related to the forgetting property of the filter. As the time from the measurement increases, and as the potential dynamic consider contribution increases, the relative importance of this measurement is effectively deweighted automatically by the filter. Thus the addition of process noise to the dynamics largely eliminates the dynamic consider contribution, as the filter deweights old data autonomously. The dynamic contribution may still influence the covariance, but its effect should not be as pervasive or long lasting as in the batch mode.

In the sequential mode the major consider contribution comes from the measurement errors or the measurement consider contribution, as described earlier in Section III.B.3. This is seen from the update equation for the sensitivity matrix, Eq. (23), where the measurement partial derivative with respect to the consider parameters is directly scaled by the gain matrix, with no other attenuations acting on this term until it is mapped to the next time step.

C. Saturation of the Sequential Filter

One of the unique features of a sequential filter which contains process noise terms is that after enough measurements have been processed, a quasi-steady state condition will be achieved in which the variances of the estimated parameters reach a lower bound and do not decrease further. This is not the case for a batch filter in which no process noise is present. Assuming that the process noise effects are fairly large, then the elements of the mapped computed covariance, $\bar{P}_k$, may be significantly larger than the corresponding elements of the computed covariance P_{k-1}. If the filter has reached a quasi-steady state, which is assumed, then the update of the mapped computed covariance $\bar{P}_k$ to P_k must reduce the size of P_k to the same order as the previous computed covariance, P_{k-1}. This implies that the gain matrix K_k is relatively large, or that the update matrix $I - K_k H_{x_k}$ is relatively small. In this instance, increasing the data weight (decreasing the assumed measurement noise variance) will reduce the size of the current computed covariance, P_{k-1}, yet will have little effect on the computed covariance as mapped forward in time, $\bar{P}_k$. Thus, while increasing the data weight may yield reduced state uncertainties at the measurement time, it will not reduce the uncertainties after they are mapped forward in time, i.e., the filter is insensitive to the data noise. Of course, if the data weight were decreased by a sufficient amount, a significant degradation would eventually be observed.

D. A Possible Failure Mode

Now consider a case where the process noise is large. From the above discussion, the gain K_k is relatively large and the update matrix $I - K_k H_{x_k}$ is relatively small. Thus, an approximate result for the form of the sensitivity matrix at time t_k is

$$S_k \approx -K_k H_{c_k} + \cdots \tag{25}$$

leading to a consider covariance of

$$P_{c_k} \approx P_k + K_k H_{c_k} \Pi H_{c_k}^T K_k^T + \cdots \tag{26}$$

The neglected terms are of negligible size for a first-order analysis. It is clear now that if the filter places too much emphasis on the current measurement, i.e., increases the size of K_k, then the consider contribution may become unreasonably large. This behavior is due entirely to the measurement consider contribution. Some specific physical effects that would lead to this type of uncertainty when treated as consider parameters would be transmission-media calibration errors or unknown biases in the measurements. The traditional remedy for this failure mode is to deweight the current measurement data. As noted above, in this case it is possible to decrease the consider contribution by deweighting even if only one data type is present,

an option not available in the batch mode. Again, if more than one data type is present, a better choice would be to deweight each measurement datum relative to the others in an attempt to retain precision.

E. Implications of the Failure Mode

One remedy of the failure mode that was discussed in Section IV.D is to deweight the current data, a process which must usually be accomplished through trial and error, especially when multiple data types are present. It would be highly desirable, though, to weight the measurements at their inherent accuracy and employ a filter model which accounts for error sources which are otherwise treated as consider parameters; such a filter will then automatically place the proper amount of emphasis on each measurement, which is determined by the assumed behavior of the modeled error sources. This is the approach behind the enhanced filter model, mentioned earlier, which employs stochastic and deterministic parameters to approximately represent the error sources affecting the measurements.

As an example of the potential benefit of enhanced filter modeling, an error covariance analysis was performed using an enhanced filter to reduce radio Doppler and ranging data in interplanetary orbit determination scenarios derived from the Mars Observer and Mars Environmental Survey Pathfinder missions [2]. In this study, the enhanced filter used to reduce the measurements contained stochastic process models to approximately represent the principal error sources affecting the measurements. The results predicted that orbit-determination accuracy improvement of factors of 2 to 4 could be realized over a conventional approach using a reduced-order filter, in which the measurement error sources were treated as consider parameters. Again, appropriate process noise terms must be added into the estimated-measurement error parameters to represent the actual random and uncertain nature of these error sources, and to prevent the filter from developing an unrealistic knowledge of them.

V. Summary and Conclusion

In this article some of the classical failure modes of reduced-order (consider) orbit-determination filters were discussed. By developing an informal description of these failure modes in terms of the information content of measurements, their error sources, and associated uncertainties, some relatively fundamental aspects of the well-known problems associated with the use of consider filters were established. A detailed description of specific failures of these types of filters is normally given in terms of mission-specific parameters and events; thus, it is difficult to go into greater detail about failure modes without recourse to specific examples.

The traditional motivation for the use of consider parameters was to simplify and speed up the filtering process in an age when computer runs often took hours, and computational efficiency and model simplifications were significant driving conditions on orbit determination procedures. The drawbacks of consider filters, i.e., the non-intuitive behavior and failure modes that were sometimes encountered, were accepted as the cost of using simplified models to increase efficiency. The recent advent of low-cost, high-speed computer workstations makes it possible to eliminate these undesirable consider contributions through the use of more sophisticated filter modeling (i.e., the enhanced filter), an important development. Enhanced filters may yield substantial improvements in orbit determination accuracy, with both existing data types and new data types proposed for future use. In particular, this makes it possible to achieve greater accuracies with simpler data types (such as those generated by the DSN's Doppler system) in future small, low-cost interplanetary missions, for which minimizing the resources and effort needed to support navigation is highly desirable. In the future, enhanced filters may enable very high accuracies for more ambitious missions in which relatively complex navigation techniques and data types may be employed.

Acknowledgment

The author appreciates the discussions with and suggestions by Sam Thurman concerning this article. His input gave a clear focus and intent to this work.

References

[1] S. W. Thurman and J. A. Estefan, "Radio Doppler Navigation of Interplanetary Spacecraft Using Different Data Processing Modes," paper AAS 93-163, presented at the AAS/AIAA Spaceflight Mechanics Meeting, Pasadena, California, February 22–24, 1993.

[2] J. A. Estefan, V. M. Pollmeier, and S. W. Thurman, "Precision X-Band Doppler and Ranging Navigation for Current and Future Mars Exploration Missions," paper AAS 93-250, presented at the AAS/GSFC International Symposium on Space Flight Dynamics, NASA Goddard Space Flight Center, Greenbelt, Maryland, April 26–30, 1993.

[3] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, San Diego, California: Academic Press, 1977.

[4] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, New York: Chelsea Publishers, p. 57, 1987.

Cammatic: An Approach to Improve DSS 13 Antenna Gravity-Loading Performance

R. Levy and D. Strain[1]
Ground Antennas and Facilities Engineering Section

Corrective actuator forces are applied to the DSS 13 antenna structure to improve its accuracy. These forces are calibrated according to the antenna elevation requirements. The calibration is derived from accurate field measurements and is not subject to the errors that result from computer modelling. A successful demonstration has provided significant antenna surface improvement. An improvement of 0.6 to 1 dB is predicted.

I. Introduction

As originally assembled and measured in the summer of 1991, the DSS 13 34-m antenna's performance was inconsistent with and inferior to finite-element-method computer model predictions. Details of this evaluation were presented in a report.[2] Figure 1 shows that the errors derived from actual measurements at Ka-band frequencies (32 GHz) were larger than the errors predicted by either the NASTRAN or IDEAS computer program models.

At that time, the antenna contained both central and off-to-the-side beam waveguide shroud cylinders. As can be seen in Fig. 2, these shrouds are substantial structures and are capable of introducing significant and complex forces that could not readily be modelled accurately by computer. Strain gage measurements on both shrouds indicated force magnitudes of about 75 kN on the bypass shroud and half as much on the central shroud. A computer-modelling analysis showed that structural interaction forces, caused by connecting the shrouds firmly to the remaining structure, could have a significant effect on performance. The study also showed that the bypass shroud forces had more effect than the central shroud forces.

However, an examination of the strain gage measurements showed the direction of the forces to be opposite to those determined from the model investigation. That is, when the model predicted tensile forces on the shroud, the strain gage forces were compressive, and vice versa. Dial-indicator measurements of the relative displacement between shroud and structure were also consistent with the strain gage forces measured on the shrouds. At the time of the report, this reversal of shroud forces was attributed to a dominating eccentricity in the alignment of the shroud bearings relative to the elevation axis.

[1] Planning Research Corporation (PRC).

[2] R. Levy, "DSS 13 Antenna Structure Measurements and Evaluation," JPL Report D-8947 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 1, 1991.

Some time later, it was proposed to remove the bypass shroud—no longer a microwave requirement—to free the structure from the associated interaction forces. There was general agreement that the structural performance would improve by eliminating the bypass. The forces were thought to act unsymmetrically on the structure and therefore were inconsistent with the objective of homologous structural response. Theodolite and holography measurements had clearly established unsymmetrical and non-homologous surface distortion patterns with the bypass shroud in place.

II. History After Bypass Shroud Removal

The bypass was removed in December 1991, and surface panels were reset to accommodate the change in structural configuration. Theodolite, holography, and gain measurements were obtained. These measurements all indicated that the surface was no better, from a microwave standpoint, and more likely somewhat worse, than with the bypass shroud in place. On the other hand, surface deformations were symmetrical and smooth, and a large surface disturbance near the former bypass had disappeared. Furthermore, theodolite measurements of the gross structure deflections were impressively consistent with model predictions. Unfortunately, the much smaller residual deflections obtained by least-square fitting the gross deflections to the nearest homologous surface were larger than anticipated.

III. Evaluation and Proposed Improvement

Since the forces developed by attaching the bypass shroud to the structure could deteriorate the antenna's performance and because these forces were observed to be in the opposite direction to those expected from its attachment to the structure, it now appears that forces from the shroud, although not symmetrical and homologous, were actually helping performance. These forces were explained as accruing from a fortuitous shroud bearing misalignment, such as might result from cam-like effects. It was proposed to recover the beneficial effects of these cammatic forces by applying calibrated active forces through symmetrically placed actuators from the alidade to the surface-supporting structure. Actuators, rather than permanent structural member attachments, were needed because the analysis had demonstrated that passive forces from attachments would act in a direction opposite to that required for improvement.

Figure 3 shows a proposed new actuator structure and system to apply the cam-like forces. This system has the capability of applying forces parallel to the Z and X axes shown in the figure. It incorporates a few new structural bar members (shown in heavy lines) to support the actuators and to apply forces to the antenna structure.

Comprehensive computer modelling studies were made to determine the optimum magnitudes and force directions to be applied at each antenna elevation. The computer effort was based upon a hybrid model: the original antenna's response was derived from a theodolite surface survey, while the additional actuator forces' response was derived from the original finite element model. The computational procedures were based upon [1]. The objective was to obtain the greatest overall improvement.

It was found that, although the greatest improvement could be obtained by optimizing both Z- and X-axis forces, there was only a small reduction in efficiency when the Z-axis actuators were eliminated and only one pair of X-axis actuators was used. This provided a major simplification of the system because the necessary new forces could readily be applied by reacting against the existing massive elevation-bearing housings. Figure 4 shows one side of the antenna with the X-actuator mechanical assembly that was designed and installed. Figure 5 is a photograph of the antenna in which the pair of actuator structures on each side appears darker than the remaining antenna structure in the background.

Figure 6 shows the gain prediction with the proposed X-axis actuators, compared with the previously measured gain curve. For example, an improvement of about 1 dB at 10 deg elevation and about 0.6 dB at 80 deg elevation can be anticipated from the figure. Such changes would provide valuable improvements in the antenna performance. Another favorable feature of the actuator system is that the forces are relatively small, so they are easily accommodated by actuators and the existing structure. Figure 7 shows the magnitudes of the anticipated optimum X-axis actuator forces. At most, the forces would be only two percent of the weight of the structure to which they were applied.

IV. Results of Field Gain Measurements

Gain measurements were made from 4 to 11 p.m. local time on March 18, 1993, by observations at 32 GHz on radio source star 3C84. The star traversed downwards from above 80 deg to just below 10 deg of elevation during this time. The procedure was to observe with and without actuator forces, applied in the range of those in Fig. 7.

The forces were applied by hydraulic jacks, and pressures were adjusted manually according to directions from the control room.

Figure 8 shows the antenna gains computed from the measurements. Solid lines are drawn between the measurements to identify the cases without actuator forces. It can be seen that the gain with actuator forces, with one trivial exception, is always higher than the solid lines. It is not possible to compute the differences in the gains by a direct subtraction because no two measurements of the moving star were made at the same elevation. Nevertheless, since there were almost 20 measurements with no actuator forces, it was reasonable to fit a curve through these points. Therefore a sixth-degree polynomial was developed to provide a close fit as a basis for interpolation at the elevations where the forces were applied.

Figure 9 shows the improvements in gain with the actuator forces, which are obtained by subtracting the interpolated gain without actuator forces. The solid curve in the figure is the predicted improvement, which is equivalent to the differences between the curves in Fig. 6. It can be seen that the measured improvements are distributed both above and below the prediction curve. Actually, more points are above the prediction curve than below.

V. Summary and Conclusions

Measurements have confirmed that this is an effective, simple, and practical way to improve the performance of this antenna. In accordance with the predictions, improvements of about 1 dB at 10 deg elevation and more than 0.6 dB at 75 deg elevation were observed. Furthermore, the applied forces may not have been optimum because it was not possible to provide a careful force calibration procedure during the measurements.

One attractive feature of this approach is that antenna performance data are derived from field measurements and are not subject to traditional differences between model predictions and performance measurements. Models are used only for preliminary calibration and assessment.

Another advantage is the simplicity of the construction and control operations in a system that requires only two actuators. Other proposed antenna actuator systems require tens to hundreds of actuators distributed throughout main reflector or secondary surfaces. The complexity of fabricating, calibrating, and controlling such unproven systems could be overwhelming.

Acknowledgments

Ron Heuerman designed and detailed the actuator assemblies. Michael Britcliffe established the experimental plan and conducted the measurement program. Both made important contributions to the success of the experiment.

Reference

[1] R. Levy,"Reanalysis, Compatibility, and Correlation in Analysis of Modified Antenna Structures," *The Telecommunications and Data Acquisition Progress Report 42-97*, vol. January–March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 367–381, May 15, 1989.

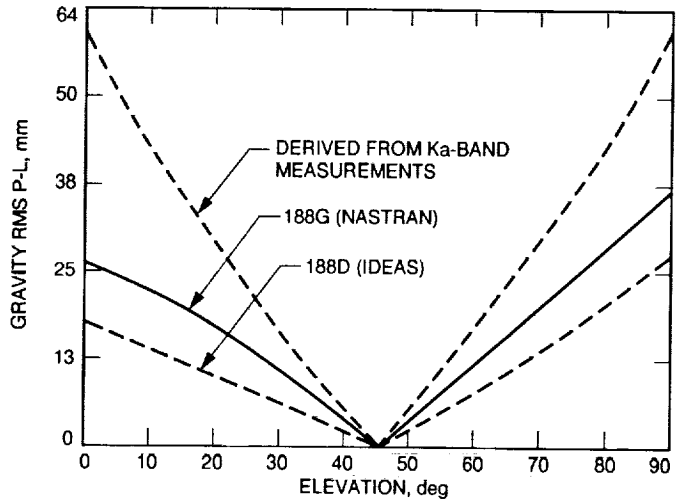


Fig. 1. DSS 13 path-length errors from measurements and computer models.

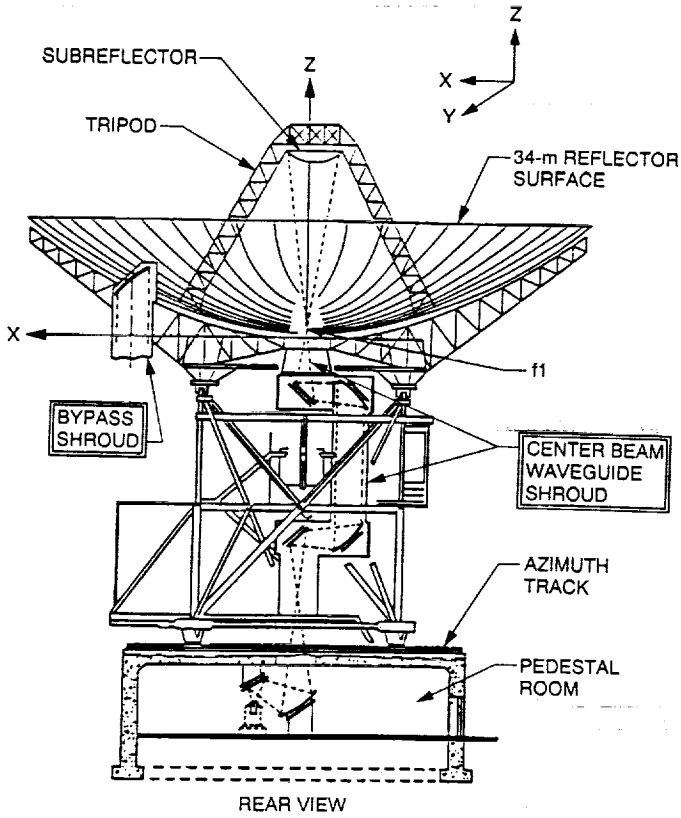


Fig. 2. DSS 13 antenna at zenith (90-deg) elevation.

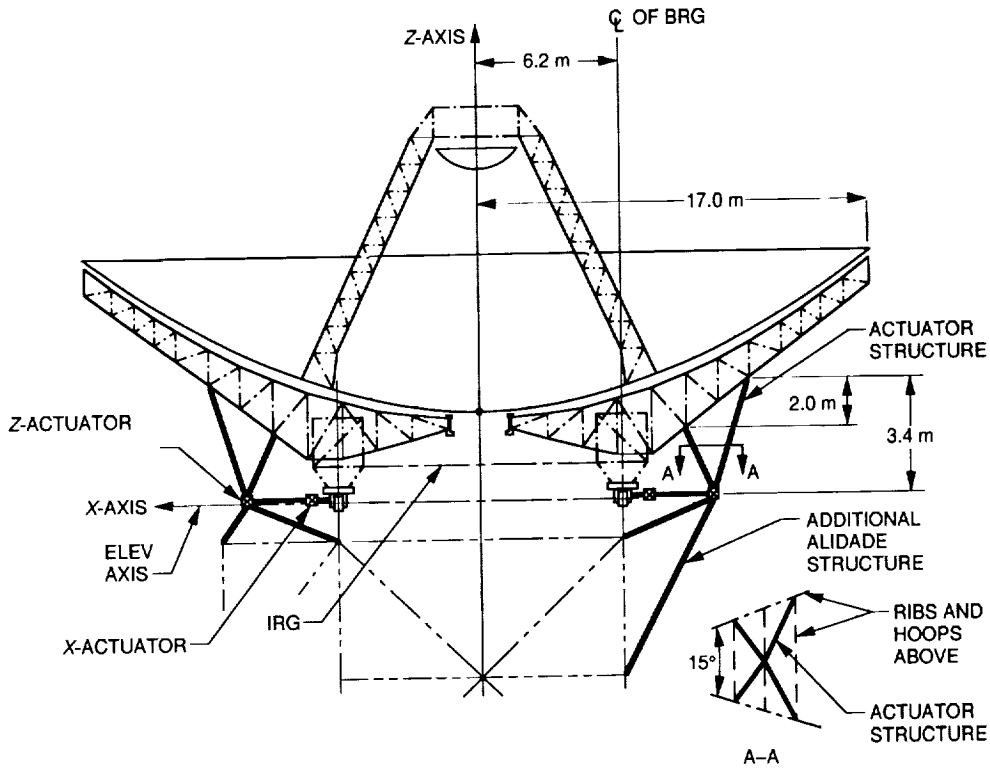


Fig. 3. DSS 13 proposed additional structure–actuator system.

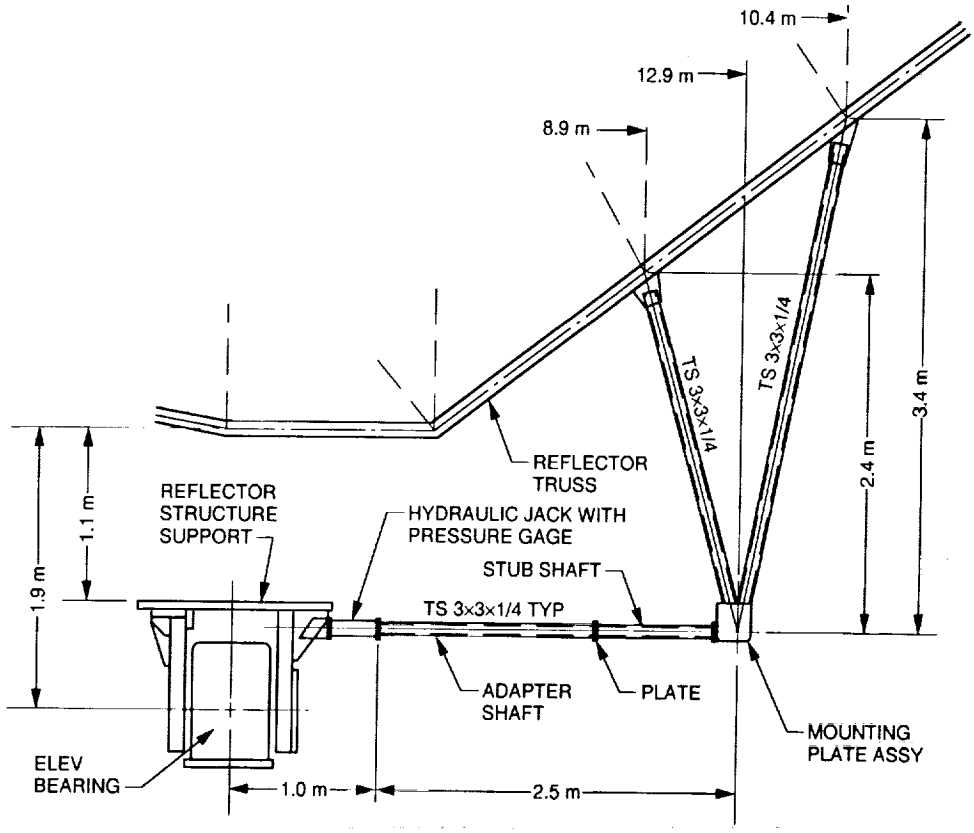


Fig. 4. Elevation view of truss tension/compression.

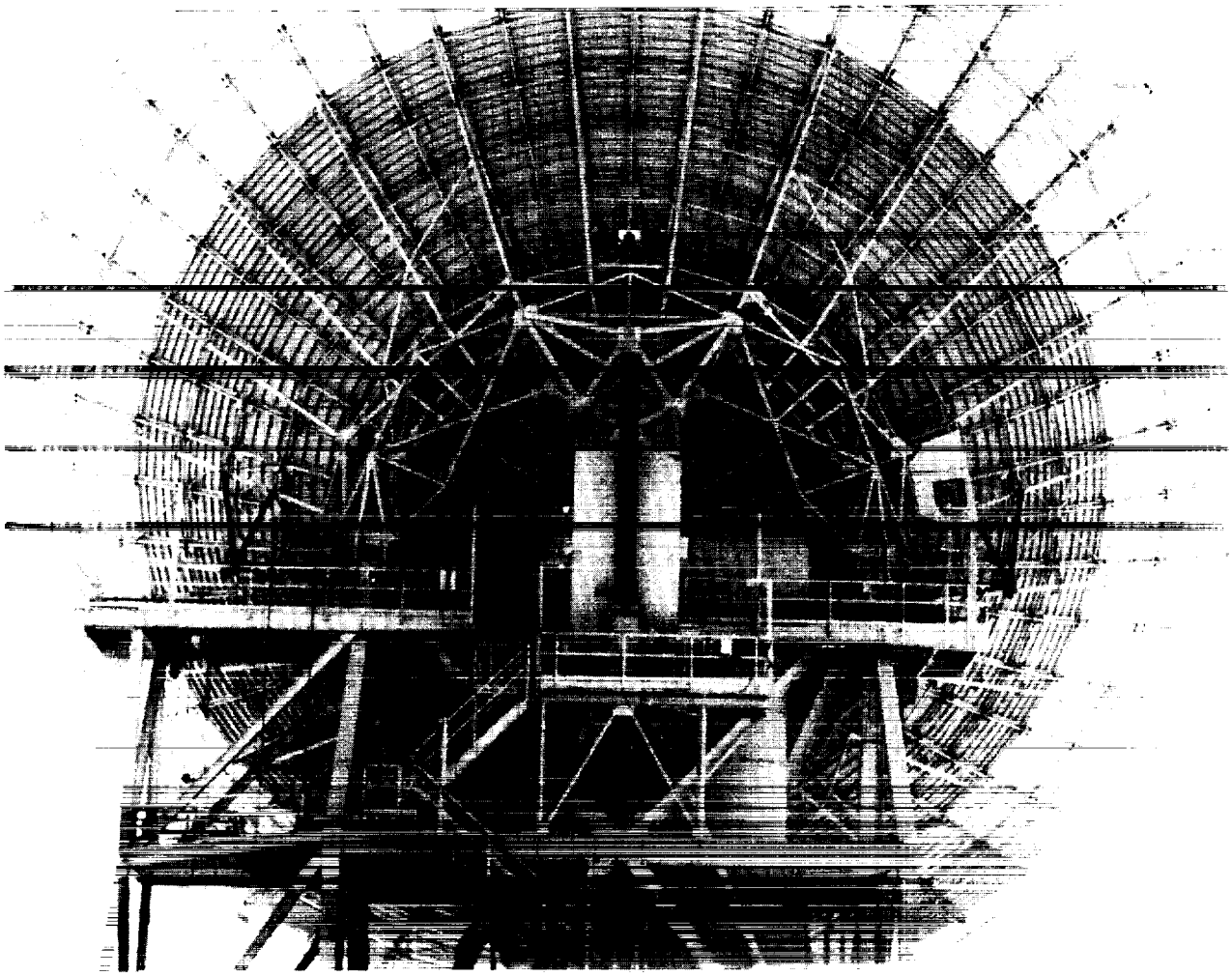


Fig. 5. *X*-actuators in place.

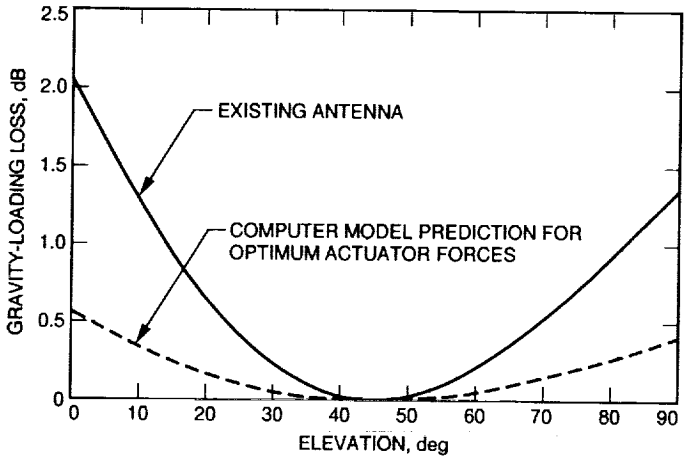


Fig. 6. DSS 13 Ka-band performance with *X*-axis actuators.

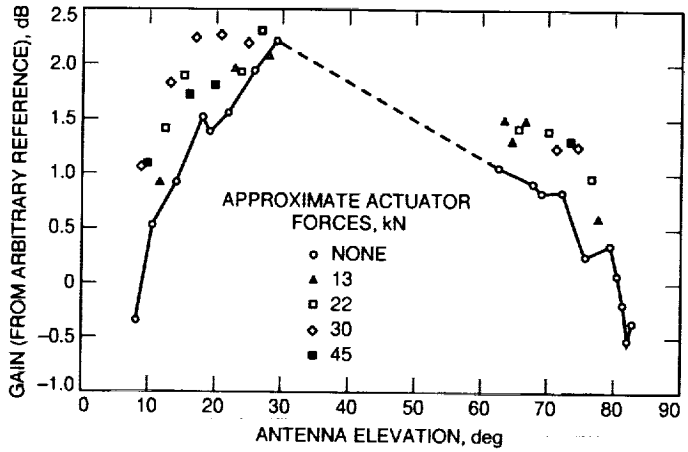


Fig. 8. Preliminary processing of DSS 13 actuator measurements (3-18-93).

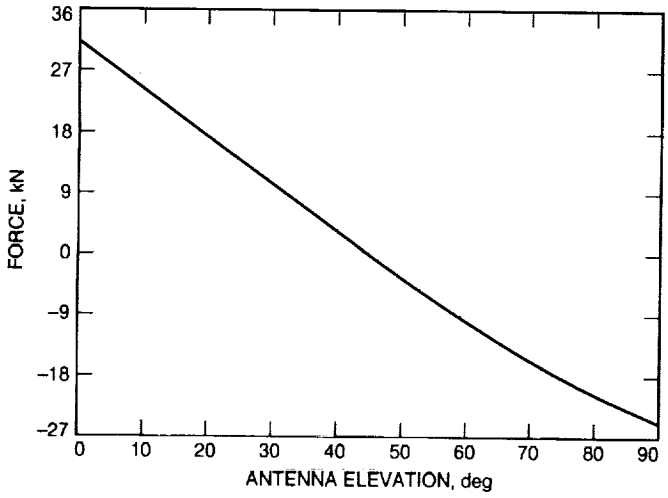


Fig. 7. DSS 13 predicted optimum *X*-actuator forces.

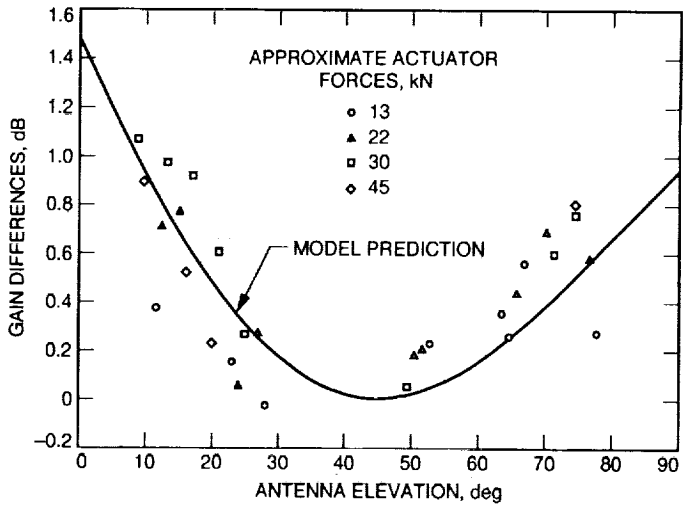


Fig. 9. DSS 13 antenna gain changes with *X*-axis actuators.

A Cryogenic Seven-Element HEMT Front End for DSS 13

J. Bowen and D. Neff
Radio Frequency and Microwave Subsystems Section

A cryogenically cooled Ka-band (33.6-GHz), seven-element front-end array for the DSN has been built and tested. This system uses seven high-electron-mobility-transistor (HEMT) low-noise amplifiers cooled by a two-stage closed-cycle refrigerator. All system components from the polarizers to the output isolators are cooled to a physical temperature between 18 and 35 K. The noise temperatures of the individual elements range from 64 to 84 K over a 2.75-GHz bandwidth.

I. Introduction

There has been a desire to increase the operating frequency of the DSN for the purpose of improving telemetry capabilities to reduce spacecraft costs for a given performance. Although improvements in gain and bandwidth are achievable, problems associated with antenna deformations and focusing become significant. This currently limits the maximum operating frequency of the 70-m antennas to 8.4 GHz [1]. The use of an array of feeds has been suggested [2] and studied [3] as a method for correcting these problems. Digital signal combining techniques can then be utilized to recover the power losses associated with the antenna deformations [4].

This article describes an array front end built to demonstrate this technique at 32 GHz. The front end is described and its performance characteristics are documented.

II. System Description

The input to the system consists of seven 22-dB-gain, smooth dual-mode feedhorns arranged in a triangular grid. The feedhorns with associated mode generators, waveguide tapers, and vacuum windows bolt directly to the top plate of the vacuum jacket (Fig. 1 shows a block diagram of a single channel). Inside the jacket, a round-to-square waveguide transition is located between the feedhorn and the hybrid polarizer. The left circularly polarized (LCP) port of the polarizer is terminated while the right circularly polarized (RCP) port connects to a waveguide isolator that provides input isolation to the low-noise amplifier (LNA). At the output of the LNA is an output isolator similar to the isolator at the input. The output waveguide section is connected to the bottom plate of the vacuum vessel and serves as the connection to an external downconverter assembly, which contains follow-up filters and amplifiers.

These components were assembled in a single vacuum jacket, 35.5 cm in diameter and 91.4 cm long. Cryogenic cooling was provided by a single CTI model 1020 closed-cycle refrigerator (CCR) (CTI-Cryogenics, Waltham, MA). The amplifier bodies were cooled to less than 20 K by thermally connecting them to the CCR cold station with multiple layers of copper shim stock.

Thermal insulation at the input is provided by waveguide sections having a very high thermal resistance but low microwave loss (see Appendix A). The input round-to-square transition is a copper waveguide section suspended in a fiberglass tube. The room-temperature end of the transition is a brass flange bolted to the fiberglass tube. The cold end is a one-piece copper flange and waveguide that has no direct physical connection to the room-temperature flange of the top plate. The RF connection is achieved with a choked circular waveguide flange at the top. Additional heat straps are connected to the polarizers to bring their temperature down to 35 K. With the polarizers at 35 K and the amplifiers at 18 K, the remainder of the circuit components, waveguides, and isolators will all be within the range of 35 to 18 K.

At the output section, thermal insulation is provided by stainless-steel waveguide sections that have a thin copper plating on their interior surfaces. In each output channel are two sections, one that connects the room-temperature flange to the first stage of the refrigerator (70 K) and another that connects the first stage to the LNA output waveguide. A short piece of flexible waveguide is incorporated in each channel to relieve mechanical stress caused by thermal contraction.

A noise diode was provided for system calibration. The output of the noise diode was split by an eight-way power divider to each element of the array and coupled to each LNA by a 20-dB coupler. The noise diode has an excess-noise ratio (ENR) of 23 dB with 32 dB of coupling loss between the diode and each LNA input.

Figure 2 shows the completed front end with the vacuum vessel and radiation shield removed. Instrumentation is also shown.

III. LNA Performance

The high-electron mobility transistor (HEMT) LNA's for the array were obtained from the National Radio Astronomy Observatory (NRAO), Charlottesville, Virginia [5]. These amplifiers are a four-stage single-ended design utilizing Fujitsu (FHR10X) HEMT transistors mounted in a gold-plated brass chassis with microstrip matching circuits on CuFlon. The LNA input and output waveguide transitions are E-plane probes coupled to a K-connector bead which then couples to the microstrip circuit. The LNA DC bias is supplied by a servo power supply externally located to the amplifier and vacuum vessel. This supply permits adjustment of the individual stage drain voltage and regulated drain current but requires individual leads for drain and gate voltages to the amplifier.

At a physical temperature of 20 K, this LNA design provides a gain of 22 to 33 dB in the band of 26 to 36 GHz (26 to 29 dB at 32 GHz). The noise temperature varies between 33 and 63 K in the band of 26 to 36 GHz (52 to 59 K at 32 GHz). Figures 3 and 4 show the maximum, average, and minimum gain, respectively, and noise temperatures for the seven HEMT LNA's.

It should be noted that because of the large number of LNA's required for this project, there was no attempt to optimize the performance of each amplifier. Therefore, the large variation in performance is indicative of the kind of results to be expected in mass production.

IV. System Performance

A. Measurements

Noise-temperature and gain measurements were performed on the system prior to installation in a 34-m beam-waveguide antenna. The method of measurement consisted of hot and cold sources alternately placed over each feedhorn and measurement of the Y-factor (the ratio of the two output noise powers). The hot load consisted of a piece of microwave absorbing pad, the temperature of which was monitored with a mercury thermometer; the cold load was a section of the same type of absorbing pad immersed in liquid nitrogen. A waveguide filter with a passband of 31.20 to 33.95 GHz was placed at the output.

The receiver consisted of a follow-up amplifier, a precision adjustable attenuator, and a second follow-up amplifier. The output of the second follow-up amplifier was sent to the power meter directly without frequency conversion. The attenuator maintained the power level within the optimum range of the power-meter sensor.

The first follow-up amplifier was of the same NRAO design [5] as each of the array amplifiers. This provided a low-noise temperature (285 K at room temperature) and sufficiently broad bandwidth to cover the frequency band of the output filter.

Measurements were also performed to determine the gain of each channel and effective noise input provided to each channel by the noise diode. The gain was determined by placing an absorbing pad at room temperature at the feedhorn aperture and sampling the total output power (subtracting the power of the noise contribution by the

amplifier itself). Noise-diode temperature was determined by again placing an absorbing pad at the feedhorn aperture and sampling the power with the diode on and off. A Y-factor was determined and used, in turn, to calculate the noise temperature of the diode as seen at the LNA input.

B. Results

Table 1 lists the values determined for effective input noise temperature as referenced at the feedhorn aperture, T_e, and the gain of each channel. The values for noise contribution by the front end, T_{fe}, were determined by subtracting the receiver temperature (T_{rcvr} = 285 K) divided by the gain for that channel:

$$T_{fe} = T_e - \frac{T_{rcvr}}{gain}$$

The noise diode added 35.5 K ±1.8 K at the front end.

C. Error Analysis

The equation for noise temperature is

$$T_e = \frac{T_h - Y(T_c)}{Y - 1}$$

where T_e is the equivalent noise temperature of the front end as referenced at the feedhorn aperture, T_h is the physical temperature of the room-temperature pad, and T_c is the physical temperature of the liquid-nitrogen pad. The error in T_e can be characterized by [6]

$$\sigma T_e = \sqrt{\left|\frac{\partial T_e}{\partial T_h} dT_h\right|^2 + \left|\frac{\partial T_e}{\partial T_c} dT_c\right|^2 + \left|\frac{\partial T_e}{\partial Y} dY\right|^2}$$

or, in this case,

$$\sigma T_e = \sqrt{\left[\left|\frac{1}{Y-1}\right| [\sigma T_h]\right]^2 + \left[\left|\frac{Y}{Y-1}\right| [\sigma T_c]\right]^2 + \left[\left|-\frac{T_c}{Y-1} - \frac{T_h - Y(T_c)}{(Y-1)^2}\right| [\sigma Y]\right]^2}$$

Here σT_h is the error in the hot-load physical temperature, σT_c is the error in the cold-load physical temperature, and σY is the error in the Y factor. The hot-load physical temperature is measured by a mercury thermometer inserted into the pad. The ability to accurately read the thermometer is limited to about 0.5 K, but the emissivity of the pad can vary by as much as 1 percent, which leads to an error at room temperature of as much as 3 K. At the liquid-nitrogen temperature, this 1-percent error will be much less significant, but the error in the temperature sensor will be on the order of 1 K.

Since the Y-factor is a ratio of the power measurements, its source of error rests in the errors of the power measurement. The Y-factor is calculated by

$$Y = 10^{\Delta P/10}$$

Here ΔP is the difference between the power resulting from the hot load, P_h, and the power resulting from the cold load, P_c, in dB. The error in the Y-factor is

$$\sigma Y = \left[\frac{\ln(10)}{10}\right] \left[10^{[\Delta P/10]}\right] (\sigma \Delta P)$$

where $\sigma \Delta P$ is then the RMS error generated by the standard deviations in the individual power measurements [6]:

$$\sigma \Delta P = \sqrt{\sigma P_h^2 + \sigma P_c^2}$$

As a result, the error in the noise-temperature measurements is 3.0 K.

V. Summary

A cryogenic seven-element Ka-band array front end has been assembled and tested. The system noise-temperature performance from 31 to 34 GHz varied from 64 to 84 K as referenced to the feedhorns.

This system can be applied to any antenna that suffers performance degradation due to surface deformations. Its relatively wide bandwidth should be attractive to radio astronomy applications as well as applications in deep-space telecommunications.

Acknowledgments

The authors thank Marion Pospieszalski, Bill Lakatosh, and Ron Harris of National Radio Astronomy Observatory for development of LNA's and thermal isolators. Also, the authors thank Gary Bury, Juan Garnica, and Willard Smith for operational support at DSS 13, and Gary Glass for assistance with error analysis.

References

[1] P. W. Cramer, "Initial Studies of Array Feeds for the 70-Meter Antenna at 32 GHz," *TDA Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 50–67, February 15, 1991.

[2] P. D. Potter, "64-Meter Antenna Operation at Ka-Band," *TDA Progress Report 42-57*, vol. March–April 1980, Jet Propulsion Laboratory, Pasadena, California, pp. 65–70, June 15, 1980.

[3] S. J. Blank and W. A. Imbriale, "Array Feed Synthesis for Correction of Reflector Distortion and Vernier Beamsteering," *IEEE Trans. Antennas Propagation*, vol. AP-36, no. 10, pp. 1351–1358, October 1988.

[4] V. A. Vilnrotter, E. R. Rodemich, S. J. Dolinar, Jr., "Real-Time Combining of Residual Carrier Array Signals Using ML Weight Estimates," *IEEE Trans. Communications*, pp. 604–615, March 1992.

[5] M. W. Pospieszalski, J. D. Gallego and W. J. Lakatosh, "Broadband, Low-Noise, Cryogenically Coolable Amplifiers in 1 to 40 GHz Range," *1990 IEEE MTT-S International Microwave Symposium Digest*, pp. 1253–1256, 1990.

[6] J. R. Taylor, *An Introduction to Error Analysis, The Study of Uncertainties in Physical Measurements*, Mill Valley, California: University Science Books, 1982.

Appendix

Waveguide Thermal Insulators for Vacuum Systems

I. Introduction

Waveguide thermal insulators are used in the various frequency bands to connect room-temperature signal waveguides to 15-K HEMT's or 4.2-K maser amplifiers. New devices have been constructed for use at X-, K-, and Ka-bands.

II. Description

A thermal insulator typically consists (see Fig. A-1) of a room-temperature flange, a shim providing an 0.08-mm gap, and a second cooled flange. The assembly is held and supported by an outer shell of G-10 fiberglass tube with a wall thickness of 0.76 mm. The gap dimension can be varied by simply changing the shim to the thickness desired; however, the 0.08-mm shim has proven optimum (at all frequencies) for the RF and thermal characteristics, providing a proper gap for room-temperature operation or when cooled below 15 K.

III. Performance

The performance at room temperature or cooled is typically less than −30-dB return loss and the insertion loss equivalent to a straight waveguide of the same length.

Having an abrupt change in the temperature gradient (hot to cold) simplifies the calculation of noise temperature for the various waveguide components as well as providing the maximum amount of cooled signal path.

The heat leak is difficult to measure; however, it appears to be no worse at 32 GHz than a stainless-steel waveguide of the same length with a thin copper-plated inside diameter (three skin depths).

IV. History

Transitions of this type have been constructed by a number of different organizations for different frequencies and different applications. This adaptation was originally undertaken by Ron Harris and Sander Wienreb of National Radio Astronomy Observatory. Their design used a unit with a fixed gap; setting the gap dimension depended on the skill of the machinist. Setting the gap for use at the higher frequencies becomes much more difficult. The new JPL design makes it much easier to control the gap assembly at any frequency. An effect noticed after repeated cold cycling was the apparent growth of the G-10 fiberglass (from 2.54 to 5.08 μm). This was corrected by simply relapping the flange faces.

The fiberglass is held both with small screws for mechanical strength and an epoxy adhesive (Armstrong A-12) to reduce movement in the joint with repeated cooling cycles.

The latest unit developed for 32 GHz allows an 8.71-mm diameter round waveguide from an antenna horn to be connected—in only 50.8 mm of length—to a 5.64-mm square waveguide feeding a polarizer assembly (the device is constructed using the wire EDM technique). The polarizer output then goes to an HEMT amplifier at 15 K. A typical 4.5-K refrigerator system would require a second thermal isolator between 70 K and 4.2 K.

This design has been used in the 32-GHz reflected wave maser system, for the first 32-GHz R&D HEMT feed package at DSS 13 and the KABLE HEMT 33.6-GHz package recently delivered to DSS 13, in the WR-42 waveguide for the K-band (18- to 26-GHz) masers at the three 70-m sites, and for X-band HEMT testing in a 15-K CCR dewar.

V. Summary

A new component is now available for use with the higher frequency DSN CCR (vacuum) receiver systems; the component offers reliability, low thermal heat leak, and low microwave insertion loss.

Table 1. Noise-temperature and gain values for each element.

Element	Noise temperature with receiver, T_e, K	Gain, dB	Front-end noise temperature T_{fe}, K
1	71.4	25.3	70.6
2	64.2	26.6	63.6
3	77.0	25.3	76.2
4	65.7	26.3	65.0
5	85.2	24.6	84.2
6	80.9	27.1	80.3
7	68.9	28.1	68.5

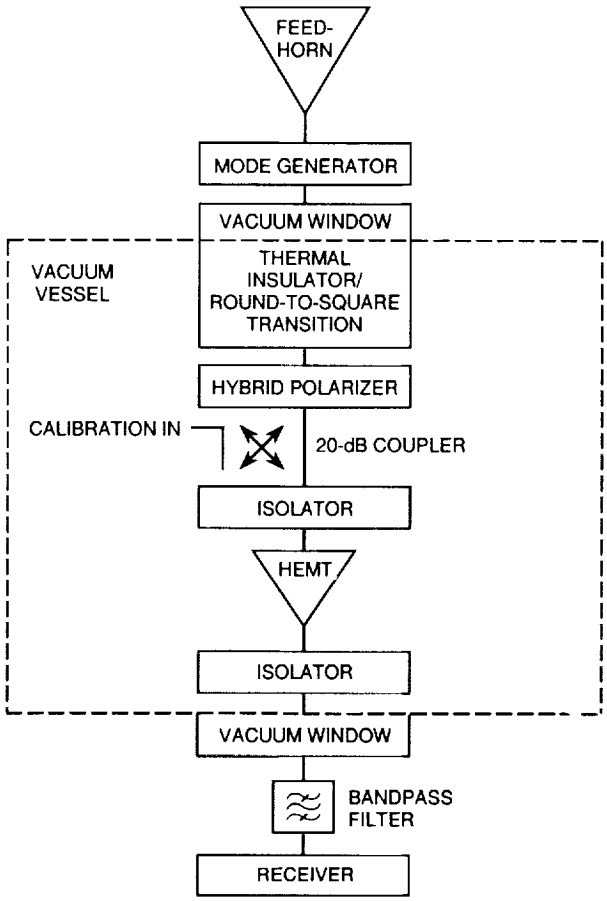


Fig. 1. System block diagram of a single channel.

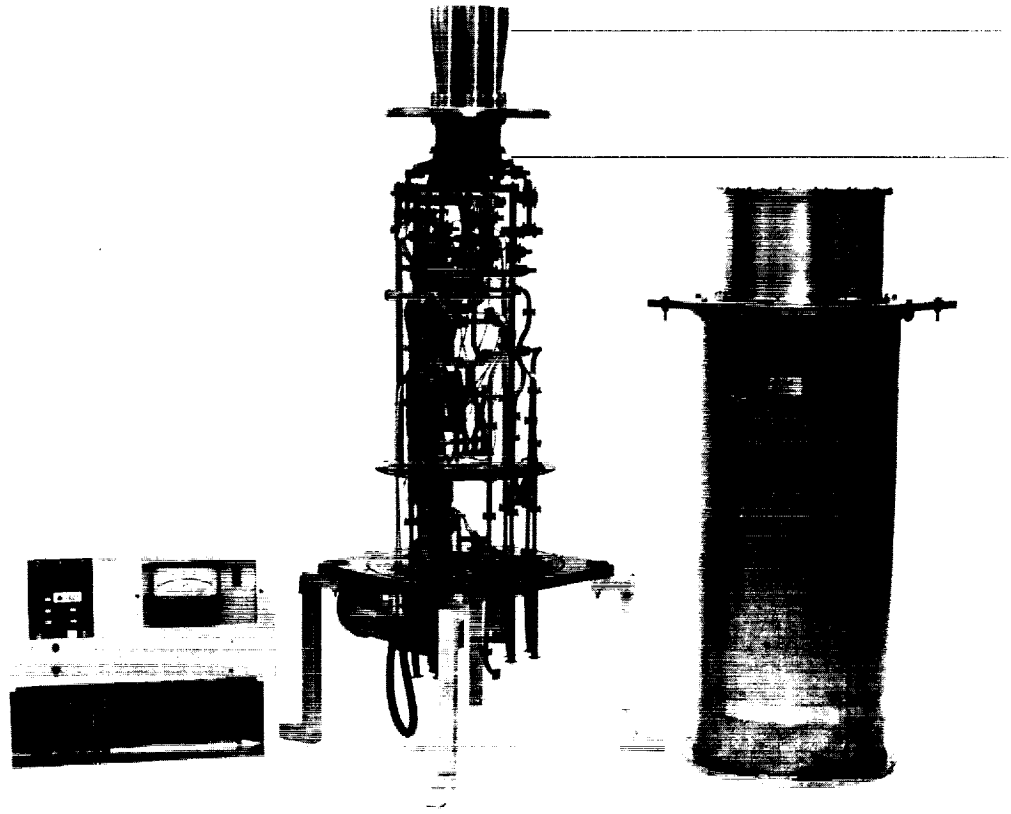


Fig. 2. The front end with vacuum vessel and instrumentation.

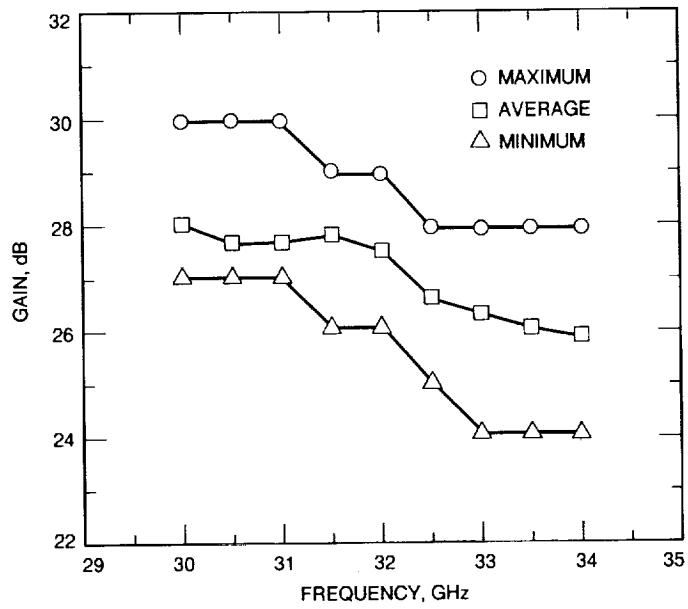


Fig. 3. The maximum, average, and minimum gain of the seven amplifiers as a function of frequency.

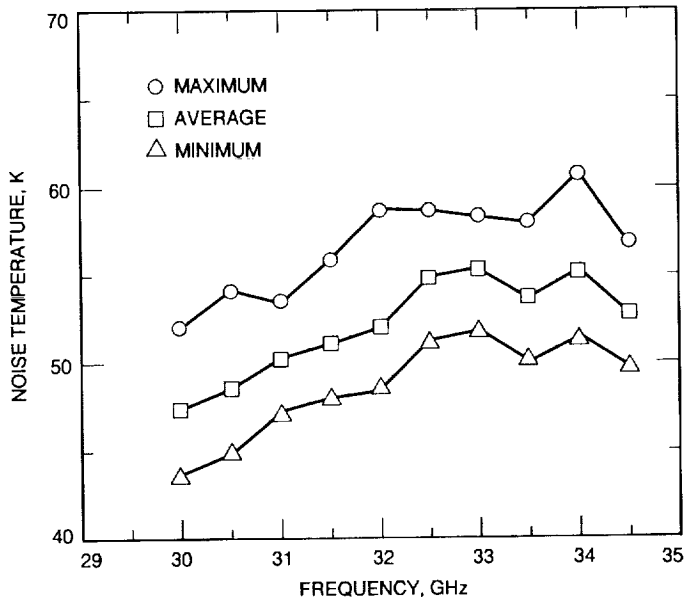


Fig. 4. The maximum, average, and minimum noise temperatures of the seven amplifiers as a function of frequency.

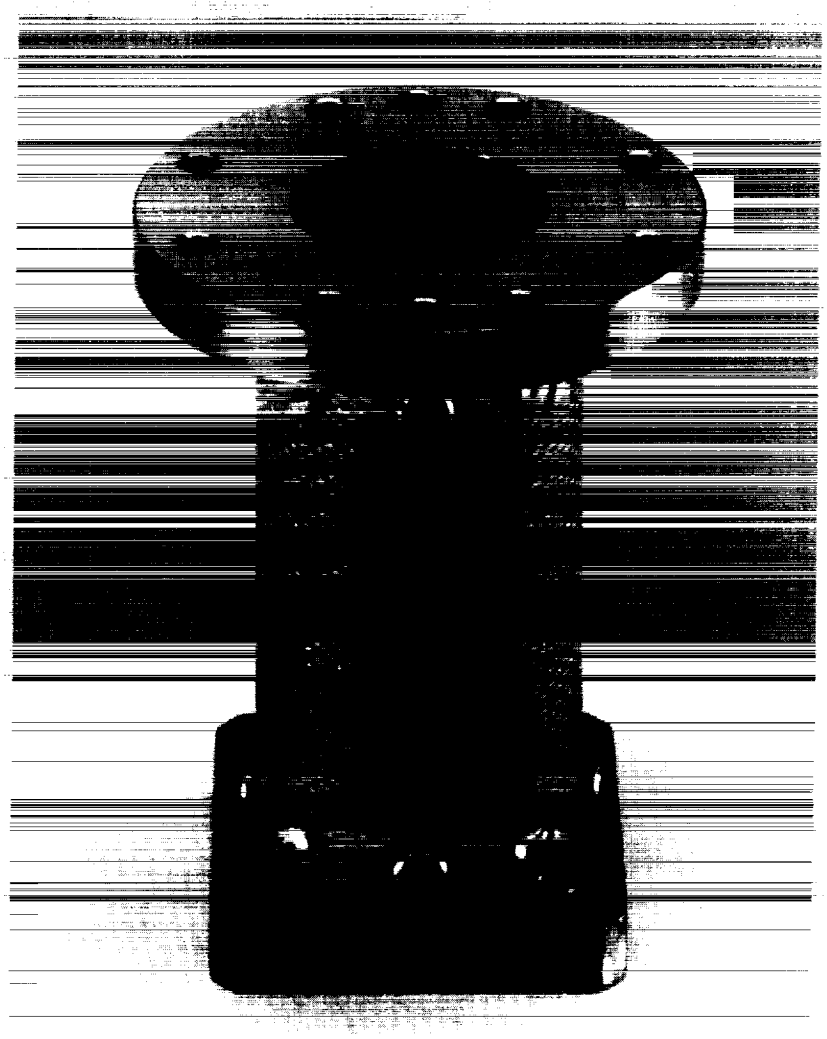


Fig. A-1. Thermal Insulator for 32 GHz.

Insertion Loss and Noise-Temperature Contribution of High-Temperature Superconducting Bandpass Filters Centered at 2.3 and 8.45 GHz

J. L. Prater and J. J. Bautista
Radio Frequency and Microwave Subsystems Section

Two superconducting Tl-Ca-Cu-Ba-O bandpass filters have been fabricated for JPL by Superconductor Technologies Incorporated, Santa Barbara, California. The filters were designed to operate at 2.3 GHz (S-band) with a 0.5-dB bandwidth of 60 MHz and at 8.45 GHz (X-band) with a 0.5-dB bandwidth of 150 MHz. The structure selected for both filters incorporates half-wavelength thin-film resonators in a stripline configuration. The S-band filter uses an edge-coupled interdigital design and the X-band filter uses an end-coupled design. The insertion loss and the noise-temperature contribution were measured at 12 K for both filters.

I. Introduction

The DSN uses extremely sensitive receiver systems to communicate with deep-space probes and to perform radio science experiments. They incorporate ultralow noise amplifiers, such as masers and high-electron mobility transistor (HEMT) amplifiers, at the front end of these systems. Radio frequency interference (RFI) can significantly degrade the performance of these receiver systems. Out-of-band RFI can result in gain compression, noise-temperature increases, and spurious output signals. This in turn can lead to the loss of received data.

Historically, masers have been the primary low-noise front end for deep-space communication, but HEMT amplifiers are now providing similar low-noise results. Both masers and HEMT amplifiers are susceptible to RFI, but HEMT amplifiers can be manufactured and maintained at a fraction of the cost. HEMT amplifiers have another advantage in that they inherently have a much broader bandwidth, but this makes them more susceptible to RFI. Therefore, work on RFI reduction is being conducted to improve the reliability of HEMT amplifiers.

Low-loss bandpass filters fabricated from high-temperature superconducting (HTS) materials can provide out-of-band RFI protection for HEMT low-noise receivers while providing minimum degradation to noise and microwave performance (e.g., return loss and group delay variation). The objective of this article is to demonstrate HTS RFI filters for cryogenic HEMT low-noise amplifiers (LNA's) at 2.3 and 8.45 GHz. The ultimate goal is to provide greater RFI protection without significantly degrading DSN receiver performance.

II. Filter Structure

Both thallium- and yttrium-based superconductors were considered as the superconducting material. The

thallium-based superconductor was chosen over the more widely used yttrium-based one for its higher critical temperature (T_c) of 123 K. Yttrium has a T_c of only 90 K. The possible configurations considered were stripline, microstrip, and coplanar. Calculations determined that the presence of two superconducting ground planes would significantly reduce radiative losses, so the stripline configuration was chosen.

Superconductor Technologies Incorporated (STI) in Santa Barbara, California, designed and fabricated the 2.3-GHz (S-band) and 8.45-GHz (X-band) bandpass filters for JPL. They were designed to have bandwidths of 60 MHz and 150 MHz, respectively, with an insertion loss goal of less than 0.5 dB.

Both filters utilize five, half-wavelength resonators in a stripline configuration with Tl-Ca-Ba-Cu-O thin-film superconductors on a lanthanum aluminate ($LaAlO_3$) substrate [1,2]. At S-band, the filter structure uses a parallel array of broadside-coupled, half-wavelength resonators. Each resonator is offset with respect to its neighbor to create the desired narrow-band response (Fig. 1). The X-band filter uses end-coupled resonators with quarter-wavelength transformers at the input and output to match to 50 ohms (Fig. 2). The dimensions of the finished packages, excluding normal metal connectors, are 3.1 × 2.63 × 1.02 cm for S-band and 3.87 × 1.70 × 3.87 cm for X-band.

III. Insertion Loss Measurements

An automatic microwave network analyzer test set was used to measure the insertion and return losses of the filters. The filters and the necessary microwave circuits were cooled in a two-stage closed-cycle refrigerator (CCR) (Fig. 3).

Two low-loss 7-mm coaxial transmission lines with APC7 connectors and end-on subminiature adapters (SMA's) served to carry the input and output signals from 300 K to 12 K. The end-to-end setup with a bypass coaxial line substituted in place of the filter showed a return loss of better than 16 dB across the frequency bands of interest. Figure 4 is a schematic of the CCR input lines.

IV. Noise Temperature Measurements

S-band and X-band Berkshire (Berkshire Technologies, Berkeley, California) cryogenic HEMT LNA's were used to determine the noise-temperature contribution of the filters. The noise temperatures of the amplifiers were measured using a conventional Y-factor method. A commercial nitrogen cold load at 77 K was used as the cold noise source and an ambient 50-Ω termination was used as the hot noise source. A Hewlett Packard coaxial switch allowed alternation of the noise sources.

The noise sources are amplified by the HEMT amplifier, down-converted, and detected by a power meter (Fig. 5). These three components form the receiver. The ratio of measured noise powers is then substituted into the following expression to obtain the receiver noise temperature (T_r):

$$T_r = \frac{T_{am} - YT_c}{Y - 1} \tag{1}$$

where T_r = noise temperature of the receiver, T_c = noise temperature of the cryogenic load, T_{am} = temperature of the ambient load, and Y = ratio of the system noise power with the hot load to the noise power with the cold load.

The difference between the receiver plus filter noise temperature and the noise temperature of the receiver is taken to be the filter noise-temperature contribution.

V. Results and Discussion

A. Insertion Loss

The filters were designed to have a 0.5-dB bandwidth of 60 MHz at S-band and 150 MHz at X-band with an inband ripple of less than 0.05 dB. The insertion loss at midband for a Tchebyscheff equal-ripple filter is given by the expression [3].

$$L_o(dB) = 8.686\left(\frac{C_n}{WQ_u}\right) \tag{2}$$

where W = fractional bandwidth, Q_u = unloaded resonator Q, and C_n is a coefficient determined by the filter order and its inband ripple value. Based on an optimistic estimate of the unloaded Q (~10,000), the midband insertion loss at 2.3 GHz is expected to be 0.17 dB and 0.24 dB at 8.45 GHz. However, the measured insertion loss is expected to be higher because the filters use normal metal SMA connectors that can contribute up to 0.2 dB of additional loss.

At 12 K, the S-band filter exhibited a bandwidth of 75 MHz with a passband peak-to-peak ripple of 0.874 dB. The maximum insertion loss of 1.07 dB corresponded exactly with the poorest return loss value of 7.8 dB. However,

at frequencies where the filter was well matched, the insertion loss was approximately 0.30 dB. Figure 6 is a plot of the insertion and return loss results.

Thermal-mechanical problems were encountered during the manufacturing of the first S-band filter prototypes. The major source of problems is the very hard substrate material ($LaAlO_3$). The stripline configuration used required that the circuit pattern be etched on the upper and lower substrates. The substrates have to be precisely aligned and properly grounded for optimum performance. The prototype filter that was tested was loaned to JPL and had poor performance near the center frequency. This was probably due to substrate misalignment and poor ground contact.

The X-band filter exhibited a passband peak-to-peak ripple of 0.31 dB with a maximum insertion loss of 0.66 dB and a bandwidth of 160 MHz. Figure 7 is a plot of the insertion and return loss results for this filter.

Taking into account the losses contributed by the connectors, the filter losses agree reasonably well with the measured results. The out-of-band rejection for both filters was greater than 40 dB.

B. Noise-Temperature Contribution

For a low-loss and reasonably matched filter (return loss $\geq$ 15 dB), the effective input noise temperature (T_e) of the filter plus HEMT amplifier can be approximated by the following expression:

$$T_e = \frac{((L-1)+r^2)T_L + T_H}{1-r^2} \tag{3}$$

where T_L = physical temperature of the filter, T_H = noise temperature of the HEMT amplifier, $L = 1/(\text{transmission coefficient})^2$, and r = reflection coefficient [4].

Figures 8 and 9 show the measured and predicted added noise contributions of the 2.3- and 8.45-GHz bandpass filters to the S-band and X-band HEMT LNA's, respectively. At the frequencies where the return loss is better than 20 dB, the noise temperature contributions of the filters are the lowest. The measured and predicted results show good correlation except in regions where the return loss is poor.

VI. Concluding Remarks

Currently, the normal metal evanescent mode and interdigital filters used are larger in bandwidth than the HTS filters tested. Yet the insertion loss of the HTS filters is of the same order of magnitude. The normal metal filters have an insertion loss of about 0.1 dB at cryogenic temperatures at S-band. These HTS prototypes are expected to meet or exceed the low-loss performance of their normal metal counterparts with smaller bandwidths to provide greater RFI protection. This can be achieved through further circuit design iterations and the reduction of connector losses by incorporating the HTS filter and HEMT LNA into a single hybrid package.

References

[1] G. L. Matthaei and G. L. Hey-Shipton, "High Temperature Superconductor 8.45 GHz Bandpass Filter for the Deep Space Network," paper presented at the MTT-S International Microwave Symposium, Atlanta, Georgia, June 14–18, 1993.

[2] G. L. Matthaei and G. L. Hey-Shipton, "Novel Staggered Resonator Array Superconducting 2.3 GHz Bandpass Filter," paper presented at the MTT-S International Microwave Symposium, Atlanta, Georgia, June 14–18, 1993.

[3] G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*, Dedham, Massachusetts: Artech House, Inc., pp. 144-145, 1980.

[4] T. Y. Otoshi, "The Effect of Mismatched Components on Microwave Noise-Temperature Calibrations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-16, no. 9, September 1968, pp. 675–686.

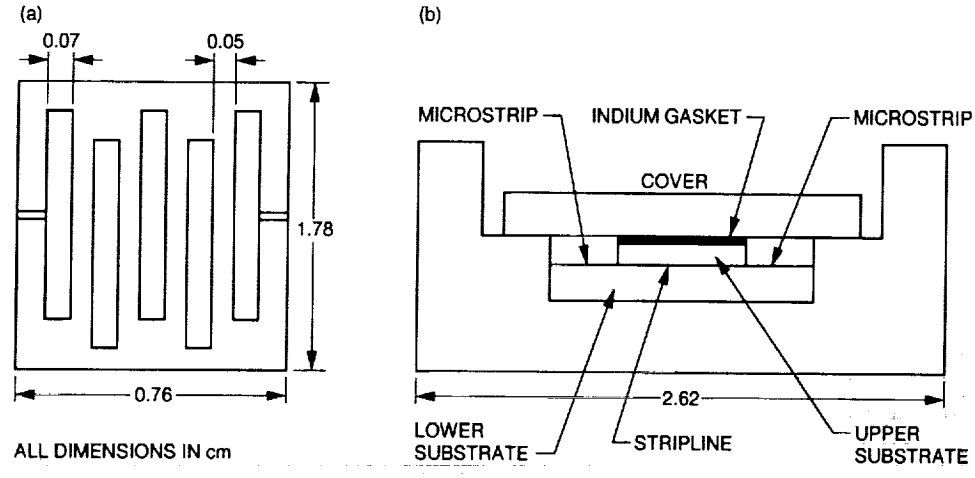


Fig. 1. S-band filter structure: (a) five-resonator parallel-array stripline and (b) cross-sectional view of the filter.

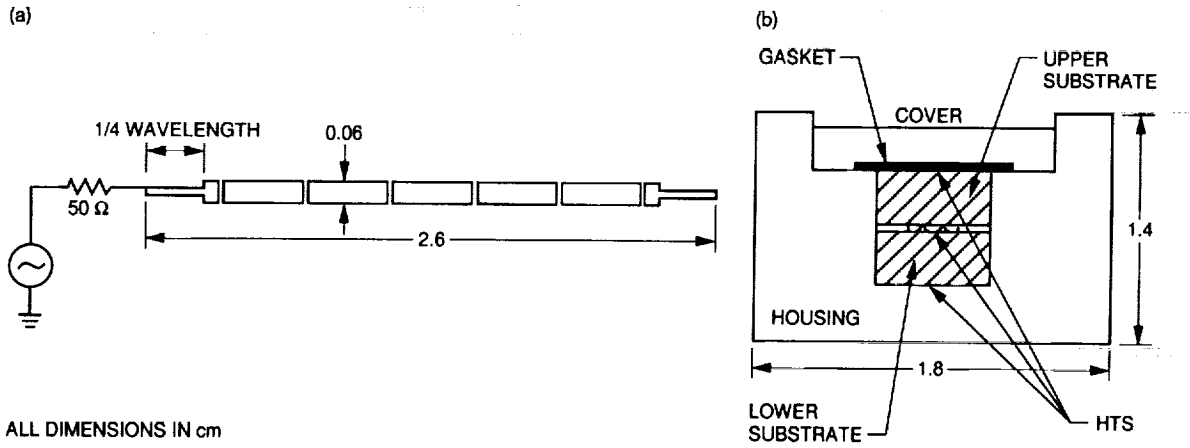


Fig. 2. STI's 8.45-GHz filter structure: (a) five-resonator end-coupled stripline and (b) cross-sectional view of the filter.

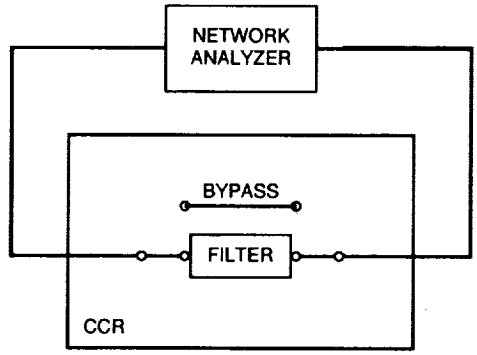


Fig. 3. Insertion-loss measurement setup.

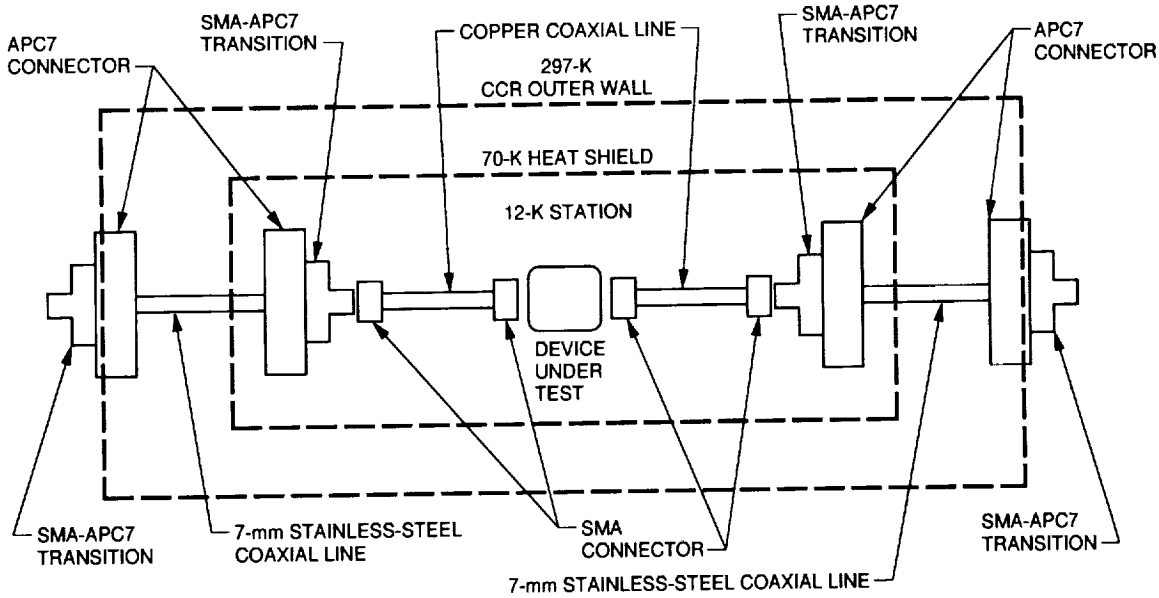


Fig. 4. Coaxial lines for the CCR.

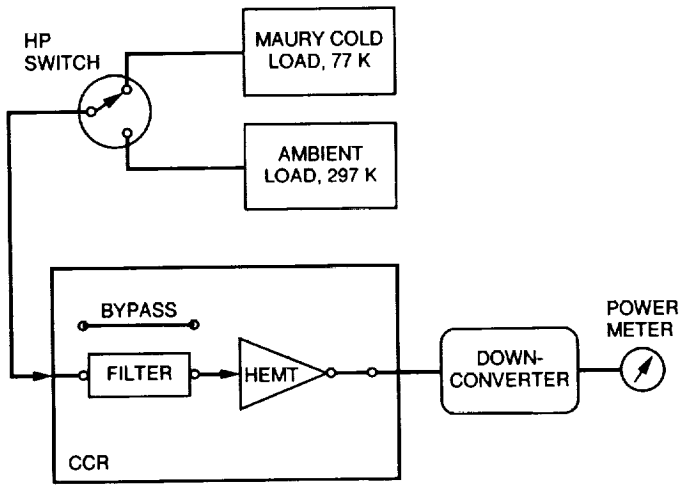


Fig. 5. Noise-temperature measurement setup.

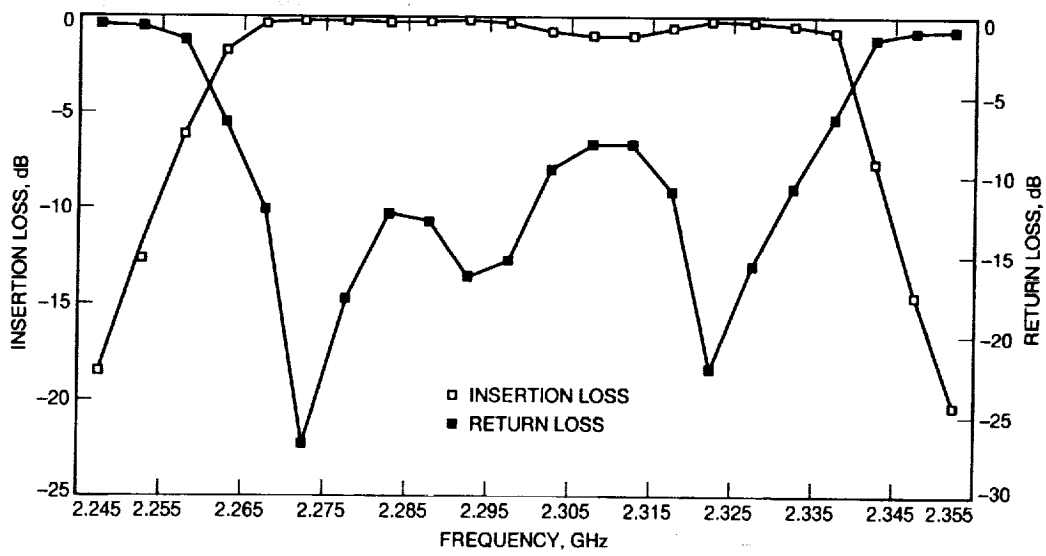


Fig. 6. Insertion loss and return loss of S-band filter.

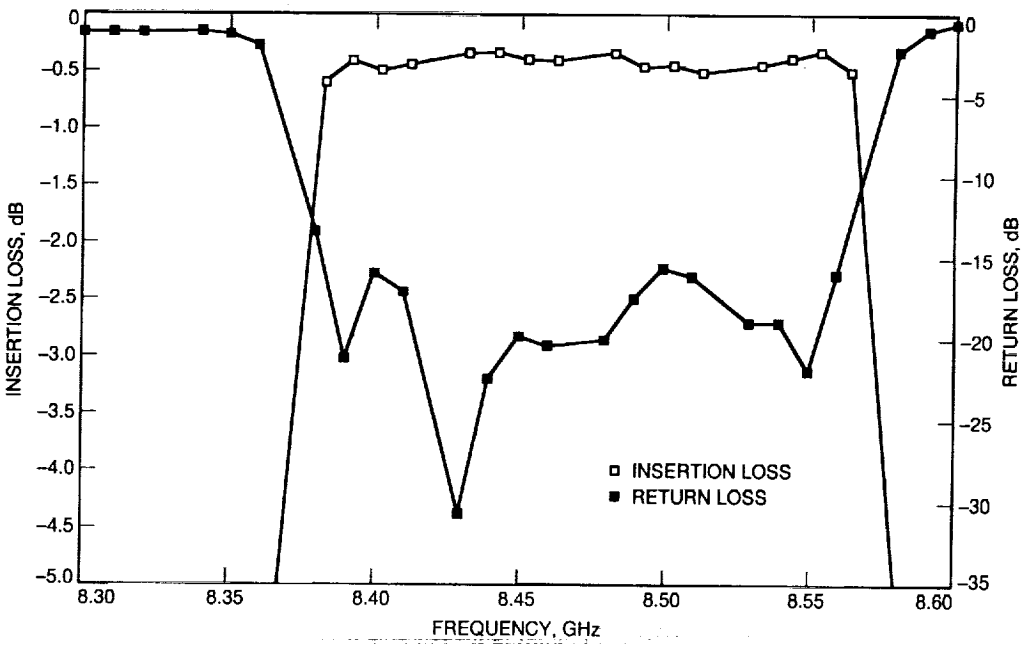


Fig. 7. Insertion loss and return loss of X-band filter.

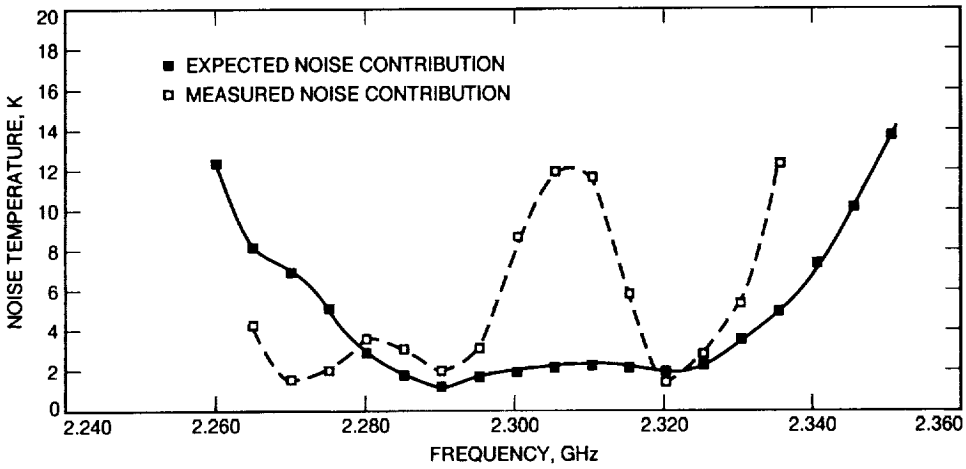


Fig. 8. Measured and expected noise contributions of the S-band filter.

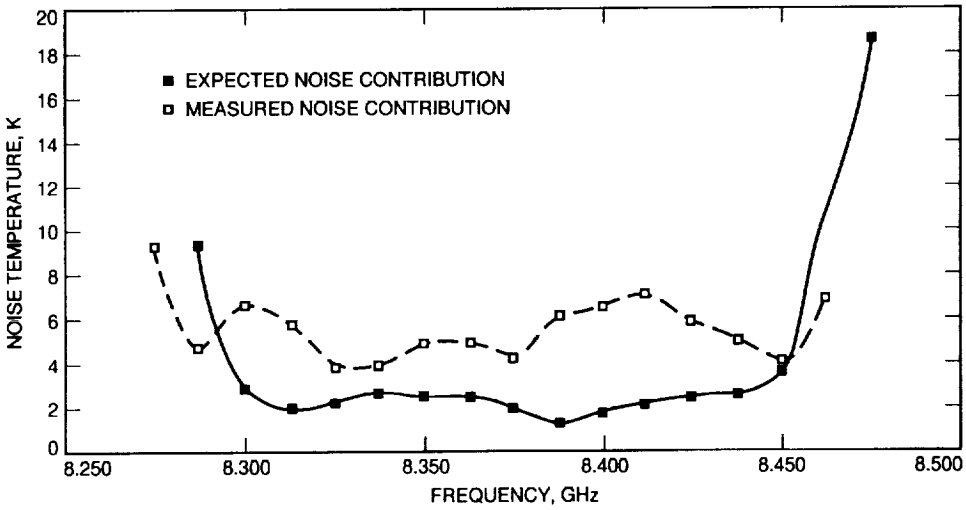


Fig. 9. Measured and expected noise contributions of the X-band filter.

Coding Performance of the Probe–Orbiter–Earth Communication Link

D. Divsalar, S. Dolinar, and F. Pollara
Communications Systems Research Section

The coding performance of the Probe–Orbiter–Earth communication link is analyzed and compared for several cases. It is assumed that the coding system consists of a convolutional code at the Probe, a quantizer and another convolutional code at the Orbiter, and two cascaded Viterbi decoders or a combined decoder on the ground.

I. Introduction

Upon arrival at Jupiter on December 7, 1995, the Galileo spacecraft (herein referred to as the Orbiter) will relay data back to Earth from an atmospheric Probe released 5 months earlier. For about 75 min, data will be transmitted to the Orbiter from the Probe as it descends on a parachute to a pressure depth of 20–30 bars in the Jovian atmosphere. Shortly after the end of Probe relay, the Orbiter will ignite its rocket motor to insert into orbit about Jupiter. The orbital phase of the mission, referred to as the satellite tour, lasts nearly 2 years, during which time Galileo will complete 10 orbits about Jupiter [1].

The Probe–Orbiter–Earth communication link is illustrated in Fig. 1. The data sent from the Probe to the Orbiter are convolutionally encoded with the NASA standard (7,1/2) convolutional code. On board the Orbiter, receivers acquire, track, and package the Probe's data together with radio science and engineering data. The received encoded data are 3-bit quantized but not convolutionally decoded at the Orbiter.

Before the high-gain antenna anomaly occurred, the plan was to send Probe data packets over the Orbiter-to-Earth downlink in real time as the data were sent from the Probe to the Orbiter. Simultaneously, a backup copy of the received Probe data would be recorded via the Orbiter's tape recorder for later playback if needed. If Galileo is forced to use its low-gain antenna for the Orbiter-to-Earth downlink, real-time relay of the Probe's data is no longer possible, and the data playback from the tape recorder is now the primary rather than the backup system for returning the data to Earth.

The data to be played back from the tape recorder consist of 3-bit quantized convolutionally encoded symbols sent over the Probe-to-Orbiter link, and they are packaged together with other data in the Probe's data packet format. In particular, each 3-bit received symbol is padded with a zero fill bit in order to occupy a half-byte in the data packet.

As a backup to the tape recorder system, the Galileo Project now proposes to additionally record a reduced version of the Probe data in the Orbiter's computer memory to protect against catastrophic tape recorder failure. The data saved in this mode are 1-bit quantized received symbol data for a reduced portion of the Probe's total lifetime.

All Probe data sent from the Orbiter to Earth are to be convolutionally encoded with the same NASA standard (7,1/2) convolutional code. Additionally, the tape-recorded data are to be encoded with a (24,12) Golay outer code (as was the original real-time data). However, the data played back from computer memory will not be Golay encoded. On the ground, the plan is to perform all decoding operations independently, i.e., to first decode the Orbiter-to-Earth convolutional code, then the Golay code (if present), and finally the Probe-to-Orbiter code.

This article investigates the performance of such a communication link with 1-bit and 3-bit quantization at the Orbiter. Since the errors at the output of the inner Viterbi decoder come in bursts, the possibility of interleaving the data before the inner convolutional code to enhance the performance is also investigated. For the case of 1-bit quantization at the Orbiter, the possibility of using a single Viterbi decoder on the ground to decode the data encoded by the cascaded convolutional codes is also analyzed. The 3-bit quantization study is treated separately in Appendix B because it is based on a simplified model that does not strictly apply to the actual Galileo Probe coding system.

II. Simulation Study of Coding System "1-bit/Nominal"

The coding system block diagram for coding system "1-bit/Nominal" (nominal system with 1-bit quantization at the Orbiter) is shown in Fig. 2. Both convolutional codes have constraint length $K = 7$ and code rate $r = 1/2$. This case models the nominal backup system for the low-gain mission Probe–Orbiter–Earth link when the received symbols at the Probe are 1-bit quantized.

In this case, a hard limiter is used before the inner convolutional encoder. Let the information data rate at the outer convolutional code be R_{b1} bits per second. Denote the received power and the one-sided noise power spectral density at the input of a hard limiter by P_1 and N_{o1}, respectively. Then the received information bit signal-to-noise ratio BSNR_1 at the input of the hard limiter can be expressed as

$$\mathrm{BSNR}_1 = \frac{E_{b1}}{N_{o1}} = \frac{P_1}{N_{o1}R_{b1}} \tag{1}$$

Let the information data rate at the input of the inner convolutional code be R_{b2}. Let the received power and one-sided power spectral density of the noise at the input of the inner Viterbi decoder be P_2 and N_{o2}, respectively. Then the received bit signal-to-noise ratio at the input of the inner Viterbi decoder can be written as

$$\mathrm{BSNR}_2 = \frac{E_{b2}}{N_{o2}} = \frac{P_2}{N_{o2}R_{b2}} \tag{2}$$

Note that for real-time operation $R_{b2} = 2R_{b1}$. However, the bits from the Probe are stored in the computer memory and therefore R_{b2} may not be related to R_{b1}. Simulation results for end-to-end bit error rate (BER) probability are shown in Figs. 3 and 4.

III. Simulation Study of Coding System "1-bit/Interleaved"

Bit errors at the output of Viterbi decoders occur in bursts. This implies that an outer decoder based on a conventional Viterbi algorithm is not optimal since it is not matched to the statistics of channels with memory. Therefore, the performance of coding system 1-bit/Nominal can be improved by using an interleaver before the inner convolutional code and a deinterleaver after the inner Viterbi decoder to randomize the bit errors at the output of the inner Viterbi decoder. The coding system of Fig. 2 with interleaving added to it is shown in Fig. 5. The end-to-end system bit-error probability is needed for this case (coding system "1-bit/Interleaved").

The channel between the output of the outer convolutional code and the output of the hard limiter can be modeled as a binary symmetric channel (BSC) with transition probability given by

$$\varepsilon_1 = Q\left(\sqrt{\frac{E_{b1}}{N_{o1}}}\right) \tag{3}$$

where E_{b1}/N_{o1} represents the received bit signal-to-noise ratio at the input of the hard limiter.

The channel between the input of the ideal interleaver and the output of the deinterleaver can also be modeled as a memoryless BSC with transition probability ε_2. This channel consists of the inner convolutional code, the additive Gaussian noise channel, and the inner soft Viterbi decoder. The transition probability ε_2 and the overall bit error probability for the cascaded channel can be upper bounded analytically using transfer function bounding

techniques. This analysis is included in Appendix A. For comparison with coding system 1-bit/Nominal, more accurate actual bit-error rate curves are obtained by direct simulation. These simulation results are shown in Figs. 6 and 7.

IV. Simulation Study of Coding System "1-bit/Combined"

In order to improve the performance of the system in Fig. 3 at high BSNR_1, consider the modified system shown in Fig. 8.

In this system, the signal received at the ground station is decoded by a single Viterbi decoder which is designed for two cascaded rate 1/2, $K = 7$ convolutional codes with no noise between the encoders. The equivalent code for these two cascaded convolutional codes is a rate 1/4, $K = 10$ convolutional code that can be decoded with a 512-state decoder. The design of such a decoder requires knowledge of the generator polynomials of the equivalent code, denoted by $h_0(x)$, $h_1(x)$, $h_2(x)$, and $h_3(x)$. These polynomials can be found by the following method, which is a simplification and generalization of the method described in [4].

Consider the cascaded rate 1/2 convolutional codes shown in Fig. 9. In this figure

$$g_0(x) = 1 + x + x^2 + x^3 + x^6$$

$$g_1(x) = 1 + x^2 + x^3 + x^5 + x^6$$

and M represents a time multiplexing operation. One can obtain the generating polynomials of the equivalent code if one can move the middle multiplexer to the output. This can be done if one notes the equivalence between the two circuits shown in Fig. 10. In this figure, $f_e(x)$ and $f_o(x)$ represent the even and odd components of the polynomial $f(x)$. The values $f_e(x)$ and $f_o(x)$ are related to $f(x)$ as

$$f(x) = f_e(x^2) + x f_o(x^2)$$

Using this equivalence, one can move the middle multiplexer to the output and can obtain the circuit shown in Fig. 11, which is equivalent to the circuit in Fig. 9. In Fig. 11, $g_{0e}(x), g_{0o}(x)$ and $g_{1e}(x), g_{1o}(x)$ are the even and odd components of $g_0(x)$ and $g_1(x)$, respectively. This relation is defined as

$$g_0(x) = g_{0e}(x^2) + x g_{0o}(x^2)$$

$$g_1(x) = g_{1e}(x^2) + x g_{1o}(x^2)$$

where

$$g_{0e}(x) = 1 + x + x^3, \quad g_{0o}(x) = 1 + x$$

$$g_{1e}(x) = 1 + x + x^3, \quad g_{1o}(x) = x + x^2$$

Finally, using the equivalence between the two multiplexer circuits shown in Fig. 12, one obtains the structure of the equivalent code shown in Fig. 13. In this figure, $h_0(x)$, $h_1(x)$, $h_2(x)$, and $h_3(x)$ can be obtained as

$$h_0(x) = 1 + x + x^2 + x^6 + x^7 + x^8 + x^9$$

$$h_1(x) = 1 + x^2 + x^4 + x^5 + x^6$$

$$h_2(x) = x + x^2 + x^3 + x^8 + x^9$$

$$h_3(x) = 1 + x^2 + x^3 + x^4 + x^5 + x^6 + x^9$$

The system shown in Fig. 8 with a Viterbi decoder based on the polynomials $h_0(x)$, $h_1(x)$, $h_2(x)$, and $h_3(x)$ was simulated. The results of this simulation were compared to the performance of coding system 1-bit/Nominal, as shown in Fig. 14. As shown in this figure, there is more than a 3-dB gain in BSNR_2 for $\mathrm{BSNR}_1 > 11$ dB.

V. Conclusion

Comparing the BER results for coding systems 1-bit/Nominal, 1-bit/Interleaved, and 1-bit/Combined, the authors conclude that interleaving improves the performance only for $\mathrm{BSNR}_2 \leq 3$ dB. For the coding system 1-bit/Combined, there is a 3-dB improvement in BSNR_2 if BSNR_1 is larger than 11 dB.

These results may be directly applicable to predicting the end-to-end performance of Galileo's Probe data return for the case without the intervening Golay code. They may also be helpful in predicting performance of future missions that might also use cascaded convolutional codes such as Lander–Orbiter–Earth links.

References

[1] T. V. Johnson, C. M. Yeates, and R. Young, "Space Science Review Volume on Galileo Mission Overview," *Space Science Reviews*, vol. 60, no. 1-4, Kluwer Academic Publishers, pp. 3–21, 1992.

[2] E. Biglieri, D. Divsalar, P. McLane, and M. Simon, *Introduction to Trellis Coded Modulation with Applications*, New York: Macmillan, 1991.

[3] D. Forney, "Lower Bounds on Error Probability in the Presence of Large Intersymbol Interference," *IEEE Trans. on Communications*, vol. COM-20, pp. 76–78, February 1972.

[4] F. Pollara and D. Divsalar, "Cascaded Convolutional Codes," *The Telecommunications and Data Acquisition Progress Report 42-110*, vol. April–June 1992, Jet Propulsion Laboratory, Pasadena, California, pp. 202–207, August 15, 1992.

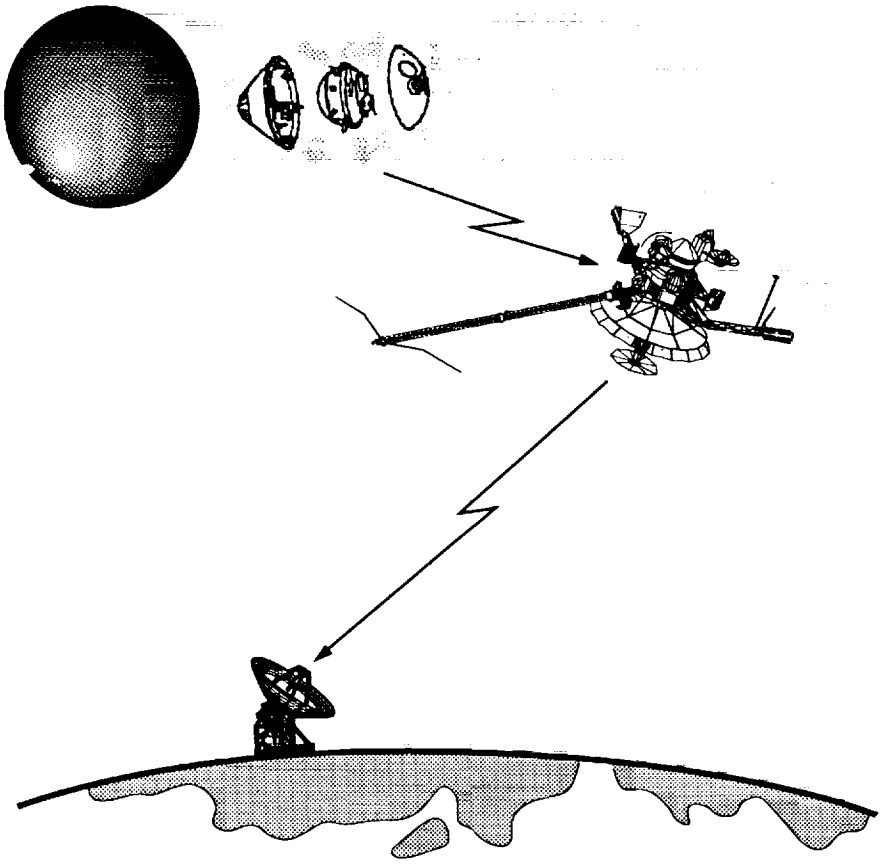


Fig. 1. Probe–Orbiter–Earth communication link.

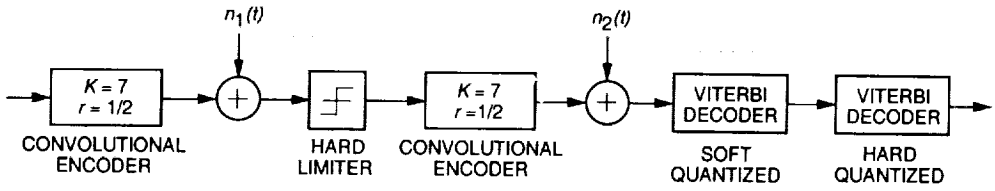


Fig. 2. Concatenated codes without interleaving: coding system 1-bit/Nominal.

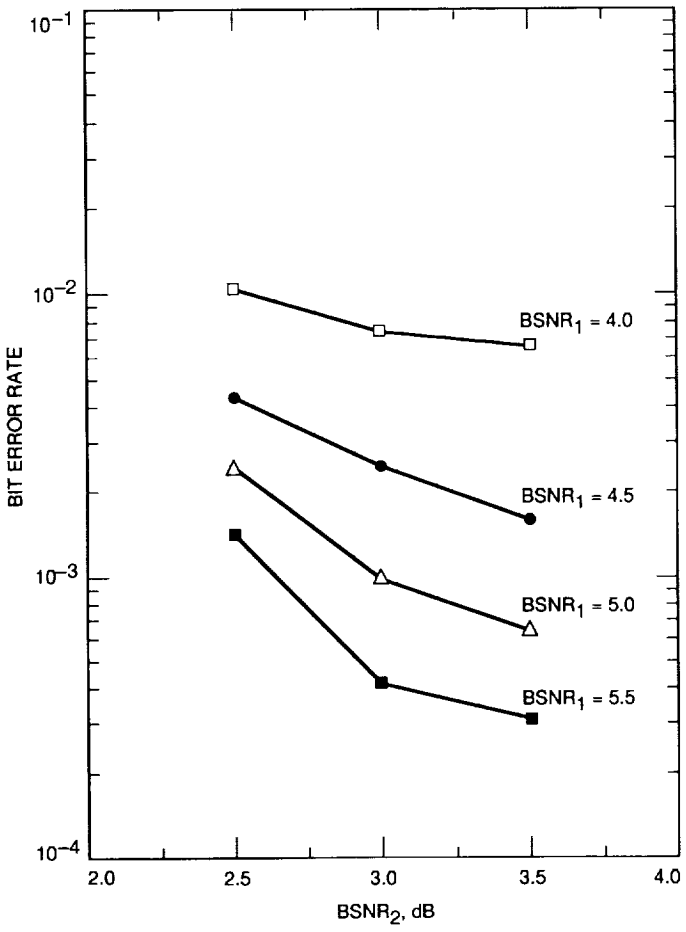


Fig. 3. BER simulation results for the coding system 1-bit/ Nominal (BER versus BSNR$_2$).

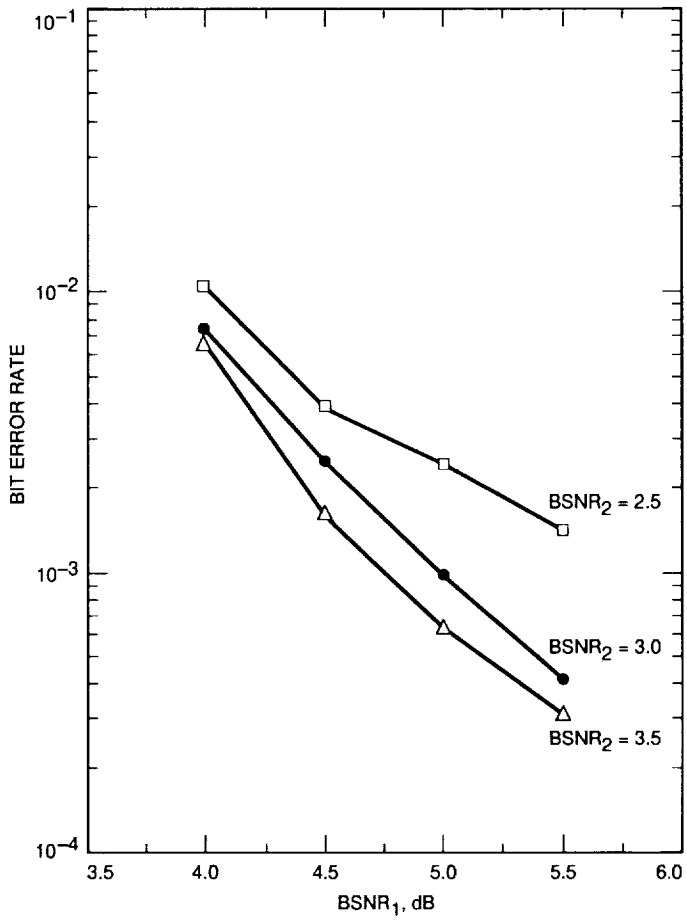


Fig. 4. BER simulation results for the coding system 1-bit/ Nominal (BER versus BSNR$_1$).

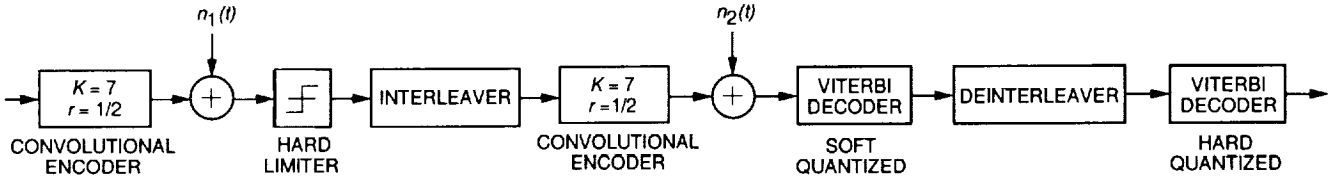


Fig. 5. The coding system 1-bit/Interleaved.

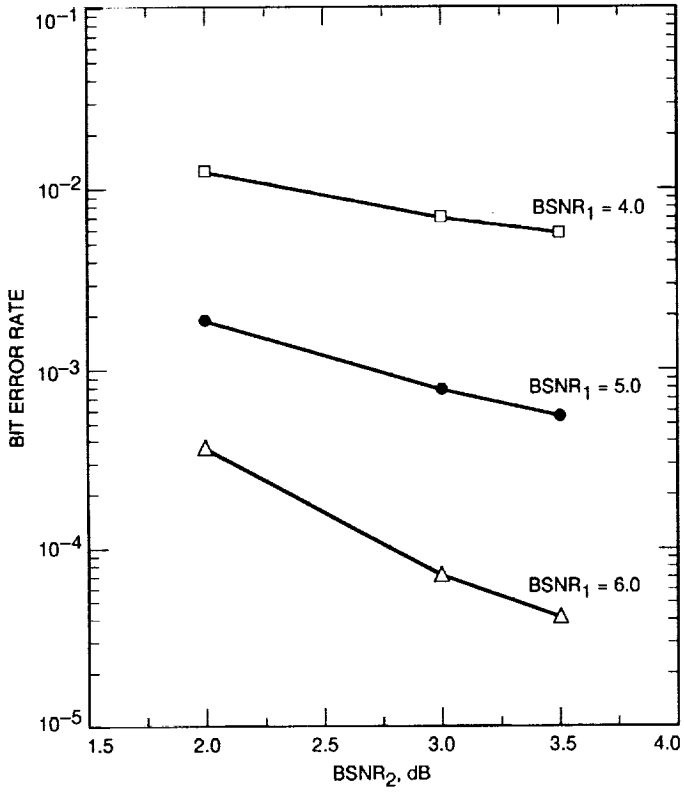


Fig. 6. BER simulation results for the coding system 1-bit/ Interleaved (BER versus BSNR$_2$).

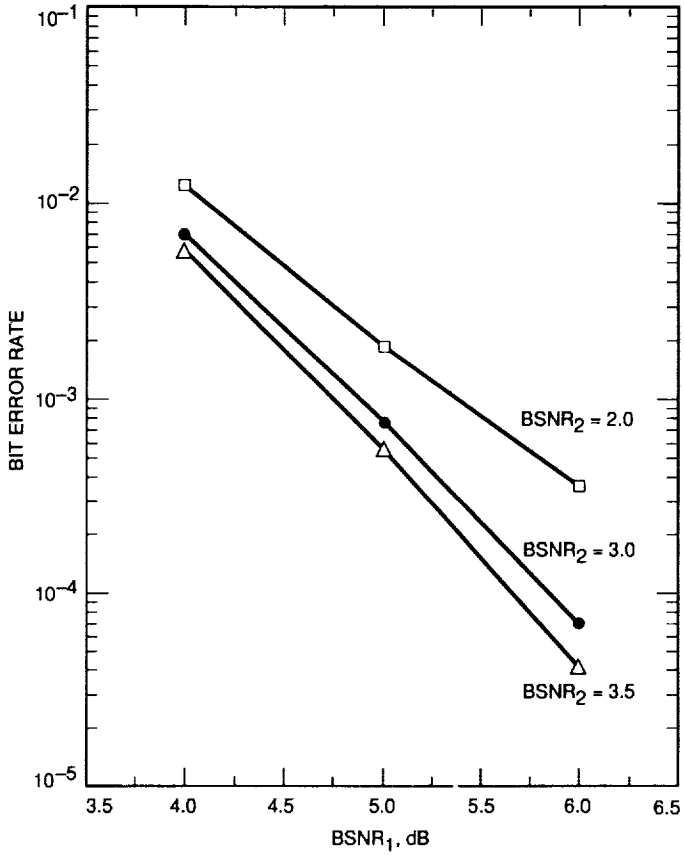


Fig. 7. BER simulation results for the coding system 1-bit/ Interleaved (BER versus BSNR$_1$).

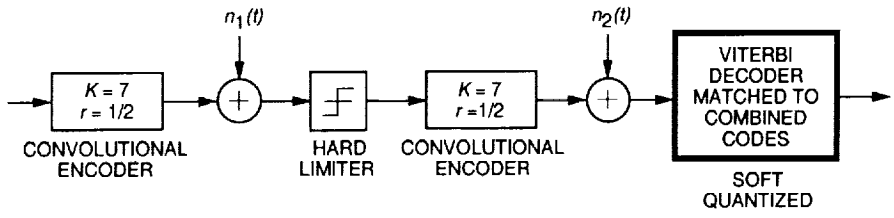


Fig. 8. The coding system 1-bit/Combined.

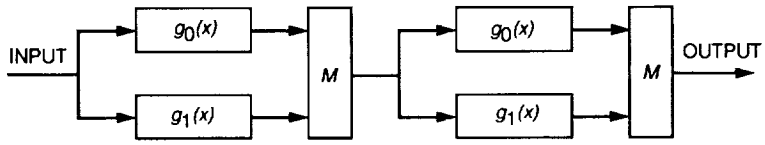


Fig. 9. Cascaded convolutional codes.

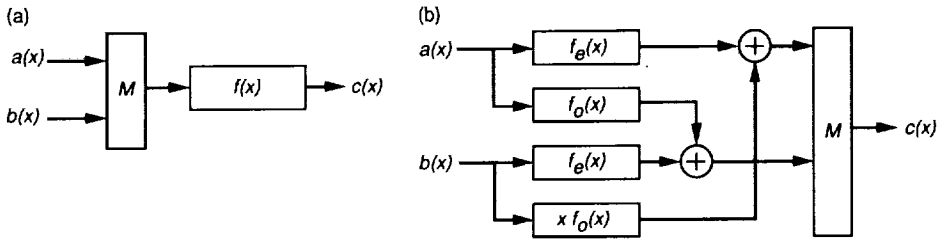


Fig. 10. Equivalent circuits: circuit (a) is equivalent to circuit (b).

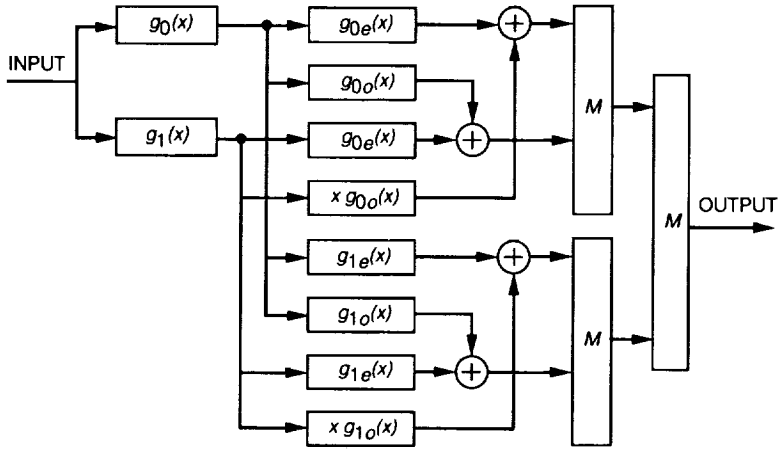


Fig. 11. Equivalent code.

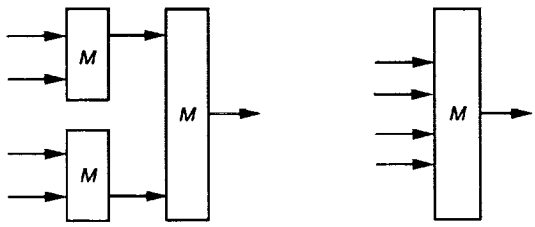


Fig. 12. Equivalent multiplexers.

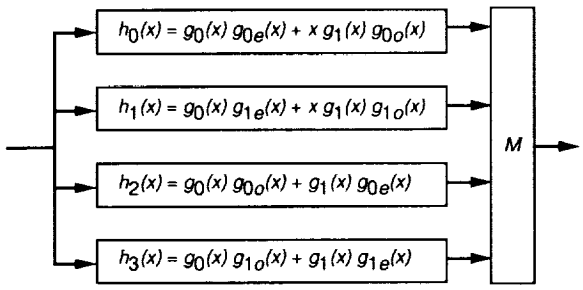


Fig. 13. Simplified equivalent code.

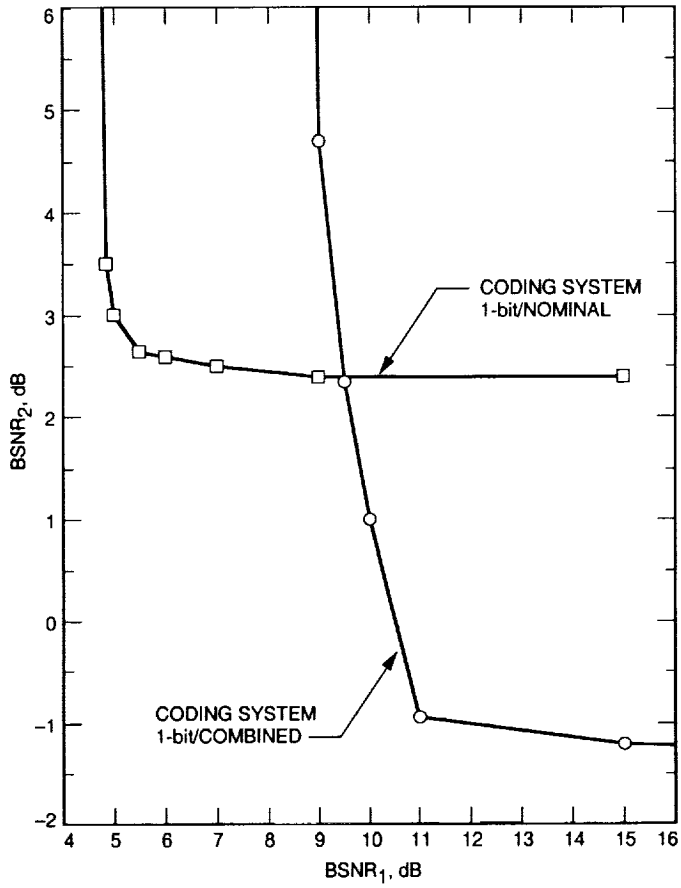


Fig. 14. Simulation comparison of the coding system 1-bit/ Nominal and the coding system 1-bit/Combined for BER = 10^{-3}.

Appendix A

Transfer Function Bounds for Coding System "1-bit/Interleaved"

The transition probability ε_2 can be computed using the transfer function bound on the bit error probability of the inner channel as [2]

$$\varepsilon_2 \leq Q\left(\sqrt{10\frac{E_{b2}}{N_{o2}}}\right) e^{\frac{5E_{b2}}{N_{o2}}} \frac{\partial}{\partial I} T(D_2, I)|_{I=1} \qquad \text{(A-1)}$$

where E_{b2}/N_{o2} represents the received bit signal-to-noise ratio at the input of the inner Viterbi decoder and D_2 is given by

$$D_2 = e^{-\frac{E_{b2}}{2N_{o2}}} \qquad \text{(A-2)}$$

The coefficients β_d of the transfer function bound

$$\frac{\partial}{\partial I} T(D_2, I)|_{I=1} = \sum_{d=d_f}^{\infty} \beta_d D^d \qquad \text{(A-3)}$$

(truncated to 15 terms) for the $K = 7$, $r = 1/2$ convolutional code with free distance $d_f = 10$ are given in Fig. A-1.

A tighter upper bound for the transition probability ε_2 can be obtained as [2]

$$\varepsilon_2 \leq \sum_{d=d_f}^{\infty} \beta_d Q\left(\sqrt{\frac{dE_{b2}}{N_{o2}}}\right) \qquad \text{(A-4)}$$

where d's and β_d's are given in Fig. A-1. The two cascaded BSC's, with transition probabilities ε_1 and ε_2, can be modeled as a single BSC with transition probability ε, which is related to ε_1 and ε_2 as

$$\varepsilon = \varepsilon_1 + \varepsilon_2 - 2\varepsilon_1\varepsilon_2 \qquad \text{(A-5)}$$

The cascaded channels and the equivalent BSC are shown in Fig. A-2.

Now it can be assumed that the outer convolutional code and its Viterbi decoder are operating over this single BSC with transition probability ε. An upper bound for the end-to-end system bit error probability P_b can be obtained once again by using the transfer function bound and Fig. A-1. Then the transfer function bound on P_b can be expressed as [2]

$$P_b \leq \frac{1}{2} \sum_{d=d_f}^{\infty} \beta_d D_1^d \qquad \text{(A-6)}$$

where

$$D_1 = \sqrt{4\varepsilon(1-\varepsilon)} \qquad \text{(A-7)}$$

The upper bound on the end-to-end bit error probability versus $\text{BSNR}_1 = E_{b1}/N_{o1}$ and $\text{BSNR}_2 = E_{b2}/N_{o2}$ is shown in Figs. A-3 and A-4. The required BSNR_1 and BSNR_2 to achieve bit error probability of 10^{-3} using the upper bound in Eq. (A-6) is shown in Fig. A-5.

A simple lower bound on the bit error probability can be obtained by using an argument discussed in [3] that is based on a "genie" providing side information to the decoder. The performance of such a genie-aided decoder will be better than that of an actual decoder. Thus, this argument provides a lower bound on the performance of an actual decoder. For linear convolutional codes, the genie observes the transmitted sequence $\mathbf{x}$ and then reports to the receiver that the transmitted sequence was either $\mathbf{x}$ or a sequence $\hat{\mathbf{x}}$ within the distance d_f (the free distance of the code) from $\mathbf{x}$. Then a lower bound on the bit error probability of rate $1/n$ codes can be obtained as

$$P_b \geq P_{d_f} \qquad \text{(A-8)}$$

where P_{d_f} represents the pair-wise error probability between $\mathbf{x}$ and $\hat{\mathbf{x}}$, separated by d_f. For this case then the lower bound on ε_2 is

$$\varepsilon_2 \geq Q\left(\sqrt{\frac{dE_{b2}}{N_{o2}}}\right) \qquad \text{(A-9)}$$

Finally, for even d_f one can obtain a lower bound on the end-to-end bit error probability as

$$P_b \geq \sum_{k=d_f/2+1}^{d_f} \binom{d_f}{k} \varepsilon^k (1-\varepsilon)^{d_f-k} + \frac{1}{2}\binom{d_f}{d_f/2} \varepsilon^{d_f/2}(1-\varepsilon)^{d_f/2} \qquad \text{(A-10)}$$

where $d_f = 10$ and ε is given by Eq. (A-5) with ε_2 replaced by the lower bound in Eq. (A-9). Unfortunately, the lower bound on the bit error probability is not tight. For high signal-to-noise ratios, the lower bound is 1/36th of the upper bound. For this reason, simulation results are also provided for this case, as shown earlier in Figs. 6 and 7.

C-2

d	β_d
10	36
11	0
12	211
13	0
14	1404
15	0
16	11633
17	0
18	77433
19	0
20	502690
21	0
22	3322763
23	0
24	21292910

Fig. A-1. Transfer function bound coefficients for the $K = 7, r = 1/2$ convolutional code.

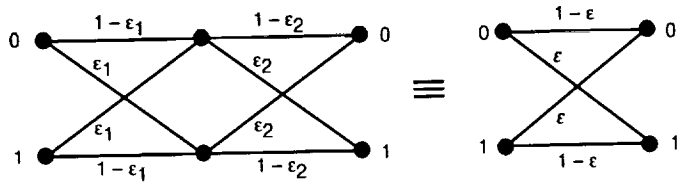


Fig. A-2. Cascaded BSC channels.

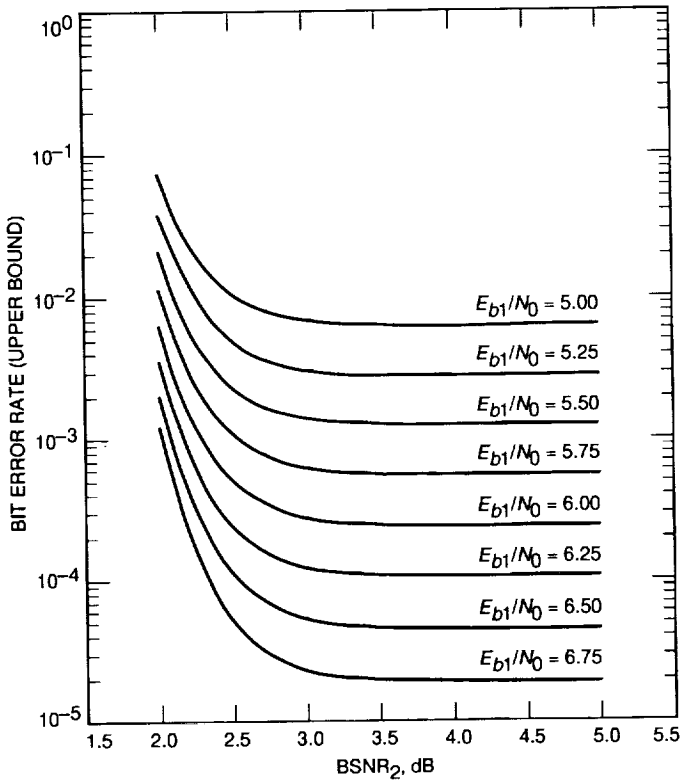


Fig. A-3. BER performance analysis for the coding system 1-bit/ Interleaved (BER upper bound versus $BSNR_2$).

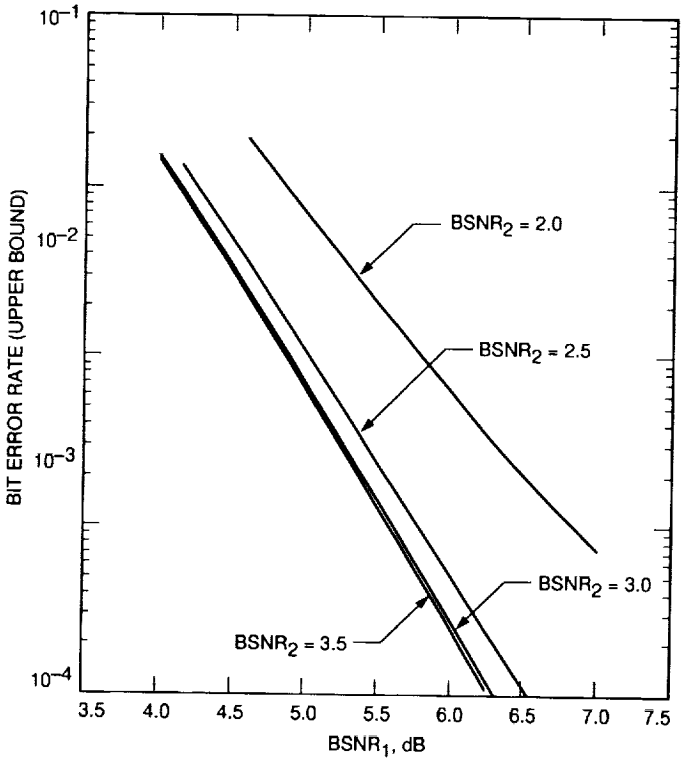


Fig. A-4. BER performance analysis for the coding system 1-bit/ Interleaved (BER upper bound versus BSNR$_1$).

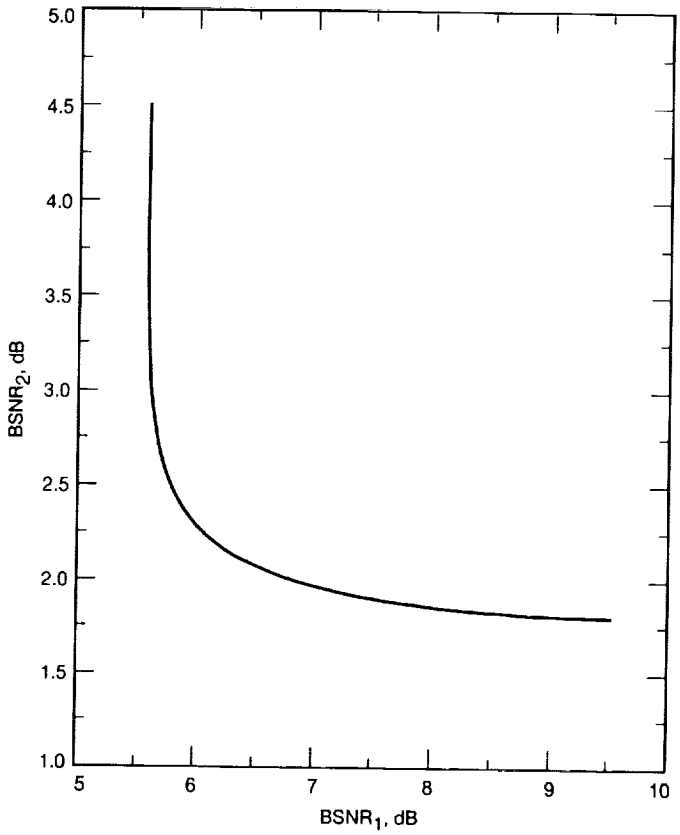


Fig. A-5. Required BSNR$_1$ and BSNR$_2$ to achieve BER = 10^{-3} (upper bound).

Appendix B

Simulation Studies of 3-bit Quantized Coding Systems

In this appendix, two different analogues of the coding system 1-bit/Nominal are analyzed for the case of 3-bit quantization at the Orbiter. Neither one of these cases models the actual Galileo Probe coding system using 3-bit symbols stored in the Orbiter tape recorder, because the effects of the Golay code are not included in this analysis.

I. Simulation Study of Coding System 3-bit/ No-Fill

The block diagram for coding system "3-bit/No-Fill," which includes separate decoders and 3-bit quantization at the Orbiter, is shown in Fig. B-1.

In this case, a 3-bit quantizer is used before the inner convolutional code. The received information bit signal-to-noise ratio at the input of the quantizer is given by Eq. (1). The received information bit signal-to-noise ratio for the inner convolutional code is given by Eq. (2).

Simulation results for end-to-end bit error probability are shown in Figs. B-2 and B-3.

II. Simulation Study of Coding System 3-bit/ Fill

The block diagram for coding system "3-bit/Fill," which includes 3-bit quantization and an added fill bit at the Orbiter, is shown in Fig. B-4.

In this case again, a 3-bit quantizer is used before the inner convolutional code. But now for every 3 bits out of the quantizer, a "0" fill bit is inserted. The received information bit signal-to-noise ratio at the input of quantizer is given by Eq. (1). The received information bit signal-to-noise ratio for the inner convolutional code is given by Eq. (2).

Since the inserted 0's are known at the inner Viterbi decoder, one can use a known-state forcing algorithm to enhance the performance. The known-state forcing algorithm simply adds a vector $(N, 0, N, 0, \ldots)$ to the state metric vector, periodically, at the times corresponding to zero bit fills. The value N can be chosen appropriately to prevent buffer overflow. No zero component of this vector corresponds to states having a zero in the least significant place. Simulation results for end-to-end bit error probability are shown in Figs. B-5 and B-6.

Comparing coding systems 3-bit/No-Fill and 3-bit/Fill, there is a loss of about $10 \log_{10} 4/3 \approx 1.25$ dB for the coding system 3-bit/Fill due to the change of transmission rate, but there is a gain of about 1 dB by using known state forcing. Therefore, it seems that the overall performances of coding systems 3-bit/No-Fill and 3-bit/Fill are very close to each other. Comparing coding system 3-bit/No-Fill with coding system 1-bit/Nominal, one concludes that coding system 3-bit/No-Fill offers about 2 dB of performance improvement.

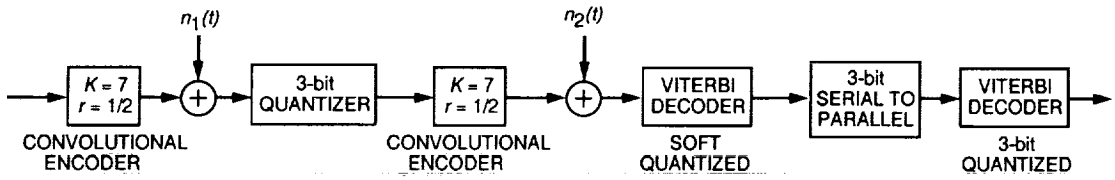


Fig. B-1. Block diagram for the coding system 3-bit/No-fill.

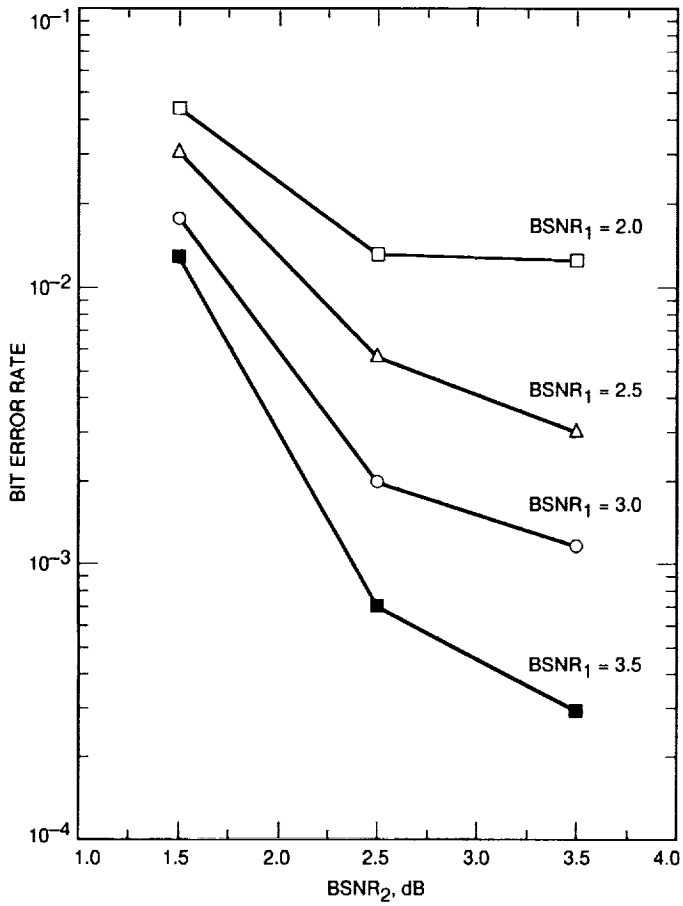


Fig. B-2. BER simulation results for the coding system 3-bit/No-fill (BER versus $BSNR_2$).

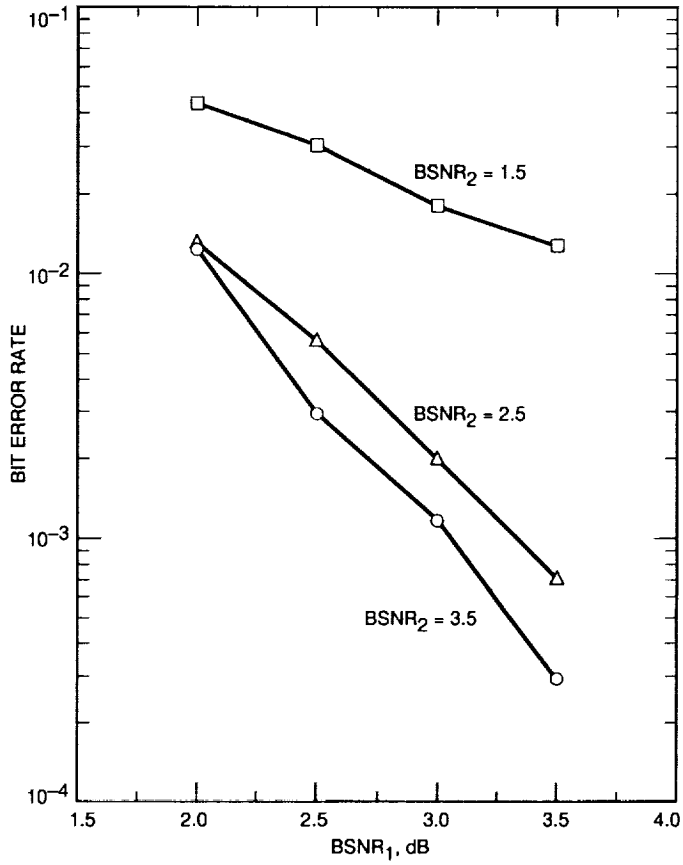


Fig. B-3. BER simulation results for the coding system 3-bit/No-fill (BER versus $BSNR_1$).

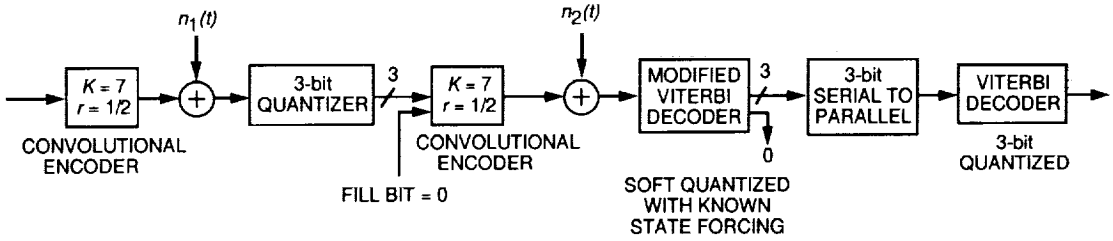


Fig. B-4. Block diagram for the coding system 3-bit/Fill.

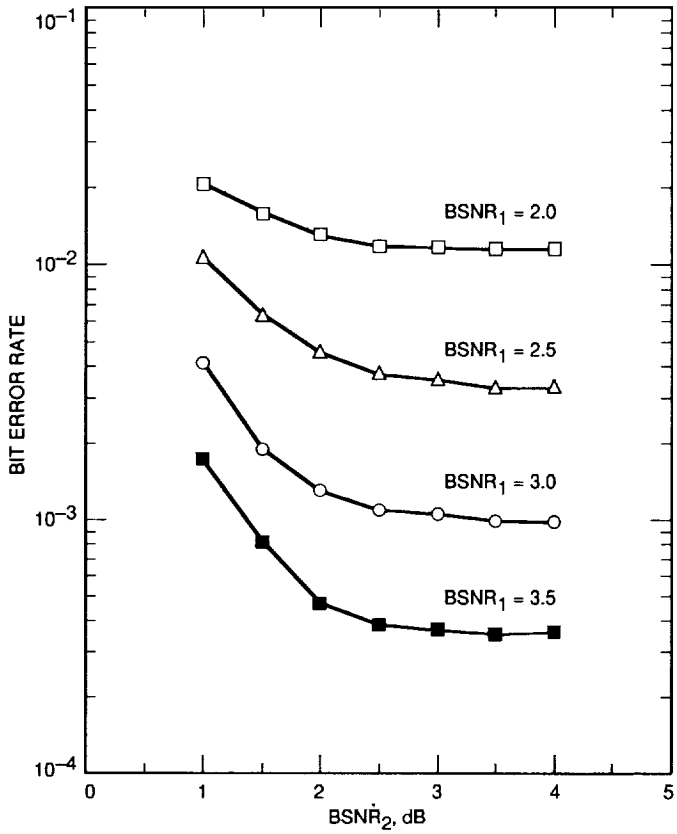


Fig. B-5. BER simulation results for the coding system 3-bit/Fill (BER versus BSNR$_2$).

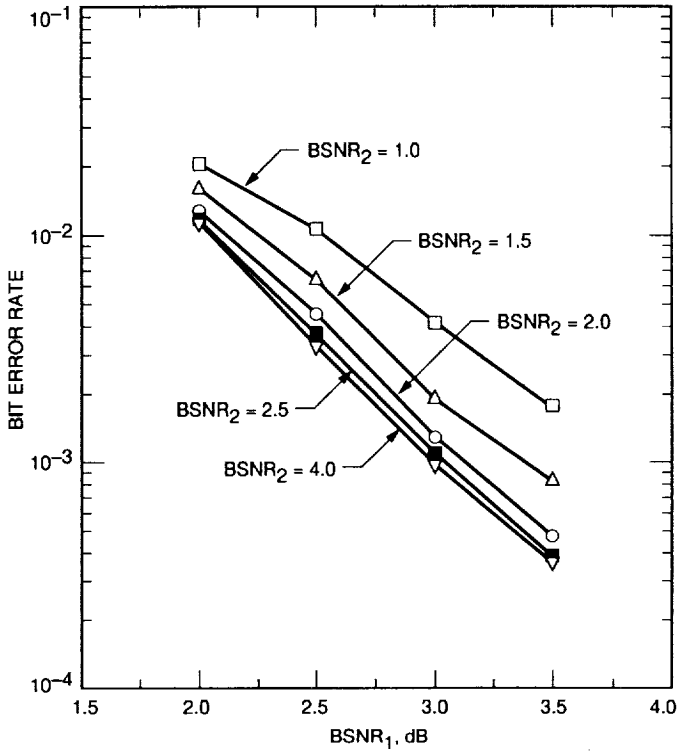


Fig. B-6. BER simulation results for the coding system 3-bit/Fill (BER versus BSNR$_1$).

Exploiting the Cannibalistic Traits of Reed-Solomon Codes

O. Collins[1]
Johns Hopkins University

In Reed-Solomon codes and all other maximum distance separable codes, there is an intrinsic relationship between the size of the symbols in a codeword and the length of the codeword. Increasing the number of symbols in a codeword to improve the efficiency of the coding system thus requires using a larger set of symbols. However, long Reed-Solomon codes are difficult to implement and many communications or storage systems cannot easily accommodate an increased symbol size, e.g., M-ary frequency shift keying (FSK) and photon-counting pulse-position modulation demand a fixed symbol size. This article describes a technique for sharing redundancy among many different Reed-Solomon codewords to achieve the efficiency attainable in long Reed-Solomon codes without increasing the symbol size. The article presents techniques both for calculating the performance of these new codes and for determining their encoder and decoder complexities. These complexities are usually found to be substantially lower than conventional Reed-Solomon codes of similar performance.

I. Introduction

This article examines a new class of codes derived from Reed-Solomon codes that captures the essential power of long Reed-Solomon codes even though the symbol size remains small. In most cases, the decoding complexity also remains small. One explanation of the increased performance of longer Reed-Solomon codes in a communications or storage system is that for a given rate they have a higher minimum distance. However, this rationale applies rigorously only when the symbol error probability is minute. Another explanation for the performance of long Reed-Solomon codes is that as sequences become longer the law of large numbers begins to take hold, e.g., very long codewords are likely to have the typical number of errors. An elaboration of this statement will serve as a good introduction to the fundamental concept in this article.

A Reed-Solomon code that experiences independent errors with symbol error probability e and erasure probability f must have a rate less than $1 - 2e - f$, the expected fraction of parity check symbols required, if the probability of its failing to decode is to be low. For a given decoder failure probability, the longer the code, the more closely its redundancy can approach this value. Any size code will have a 50 percent chance of correct decoding if its distance is exactly equal to one more than the expected number of erasures plus twice the expected number of errors. A very long code will operate in the region where the law of large numbers applies and experience a sudden, precipitous drop

[1] Work supported under a contract between Johns Hopkins University and JPL.

in the probability of decoder failure as the amount of redundancy provided increases from less than $1 - 2e - f$ to greater than $1 - 2e - f$. A shorter code will experience a more gradual drop. Figure 1 shows an example of a 5-bit extended Reed-Solomon code operating over the independent symbol erasure channel with 1 percent symbol erasure probability.

If a 5-bit Reed-Solomon code experiences an independent symbol erasure rate of 1 percent, and if the probability of decoder failure must be less than 5×10^{-6}, then the code must able to correct 5 erasures. The probability of having six or more erasures, $E_{32}(0.01, 06)$, is actually 7.24×10^{-7}. There is, however, less than a 0.4 percent chance of having more than two erasures in a codeword. Thus, there is the possibility of sharing among many different codewords the three parity check symbols that are needed only 0.4 percent of the time. One way of achieving this would be through a return channel over which the receiver informs the transmitter of the particular codewords that require the extra parity checks. The transmitter computes all five parity checks, but, in the beginning, sends over the channel only the first two. The receiver attempts decoding of the shortened codewords by declaring the symbols that were not sent to be erasures. Then, $E_{29}(0.01, 3) = 0.0030$ of these shortened codewords will fail to decode. (Since this example is using the erasure channel, the probability of incorrect decoding is undefined.) The receiver then requests the extra parity checks for those few words that failed to decode. The decoder failure probability is unchanged from the original system, but only 2.009 parity checks per codeword are required on average, i.e., the average number of parity checks per codeword has been cut by more than half.

The essential concept in this article is that the same type of parity check sharing can be accomplished without the reverse channel. The next section will show how this is accomplished by completing the introductory example and will explain techniques for computing the performance of codes used on the erasure channel.

II. Codes for Erasure Correction

The parity check symbols will be shared among a set of codewords by using the same Reed-Solomon encoder that produced them. The third, fourth, and fifth parity symbols from 27 different codewords will be fed back into the encoder to form the vertical codewords shown in Fig. 2.

None of the symbols intersected by both Reed-Solomon codewords is sent over the channel. In fact, only the first four parity check symbols of the vertical codewords will be sent over the channel; the fifth symbol of each of the vertical codewords will be discarded.

Now that the scheme for sharing parity check symbols among the codewords has been precisely described; its operation will be proven by showing that the decoder failure probability of the horizontal codewords has decreased. Consider a horizontal codeword that has more than two erasures and so has failed to decode. The probability that three or more of the remaining horizontal codewords have also failed to decode is $E_{26}(0.0030, 3) = 6.67 \times 10^{-5}$. Should this event occur, decoder failure can be declared since doing so will contribute only $(0.003)(6.67 \times 10^{-5}) = 2 \times 10^{-7}$ to the overall probability of failure to decode. The probabilities multiply since the initial failure of a single codeword is independent of the success or failure of any or all of the others.

If the 26 other codewords have only two failures among them, then each of the vertical codewords will decode if no more than one channel erasure has occurred in its set of four parity symbols, i.e., the symbols actually sent over the channel. The probability of more than one channel erasure in a set of four is $E_4(0.01, 2) = 0.0006$. Thus, if three horizontal codewords have failed, the vertical codewords will fail independently with probability 0.0006; this number is an upper bound on the erasure rate experienced by the rightmost three symbols of each horizontal codeword if failure has not already been declared.

An upper bound on the failure of each horizontal codeword is thus obtained by assuming independent erasures on all symbols with a 0.01 rate for the first 29 and a 0.0006 rate for the last three and then adding 2×10^{-7} to account for the probability of declaring vertical codeword failure. This bound is negligibly less than the original failure probability based on a uniform 1 percent erasure rate. The complexity increase at the encoding end is 3/27, since for every set of 27 codewords, three more will be needed. The decoder complexity increase will never be more than 6/27, the fractional contribution of the vertical codewords plus the fractional contribution of the three possible horizontal redecodings. The average work performed by the decoder can, of course, be less, e.g., often complete decoding can be accomplished without using the vertical codewords and so their information may simply be discarded.

III. Single Field Codes for Error Correction

This section will explain the design of single stage combined Reed-Solomon codes for the independent symbol error channel by presenting an example based on the

NASA Standard (255,223) code. The channel error rate is 2 percent which produces a decoder failure probability of $E_{255}(0.02, 17) = 1.9 \times 10^{-5}$; this is almost equal to the design symbol error rate in the Voyager spacecraft communications system. For implementation economy, the vertical and horizontal codewords will again be identical. The first 16 redundant symbols of each horizontal codeword will be sent over the channel in the conventional manner. The remaining 16 will be incorporated into vertical codewords as shown in Fig. 3. The probability that one of the horizontal codewords will have more than eight errors, and so fail, is $E_{239}(0.02, 9) = 0.0528$. The structure of Reed-Solomon codes guarantees that almost all of these excess error patterns can be recognized, i.e., the probability of incorrect decoding of one of the horizontal codewords is negligible [1]. The 223 information symbols of the vertical codewords will thus experience an erasure rate of 0.0528.

Consider a codeword that has experienced 9 or more errors. The probability of 28 or more of the other codewords failing is $E_{222}(0.0528, 28) = 2.00669 \times 10^{-5}$. If 29 or more of the horizontal codewords have failed, the decision not to attempt decoding of the vertical codewords contributes $(0.0528)(2.00669 \times 10^{-5}) = 1 \times 10^{-6}$ to the failure probability of each of the horizontal codewords. The decision never to attempt decoding of the vertical codewords with more than 28 declared erasures is quite sound in this case since it keeps the error probability of the vertical codewords negligible [1].

If a vertical codeword has 28 erasures, then the probability of its not decoding successfully is $E_{32}(0.02, 3) = 0.0257$. Thus, each of the horizontal codewords will experience an independent symbol erasure rate of no more than 0.0257 on the last 16 symbols if an excess of erasures has not already caused vertical codeword failure to be declared. The failure probability of the horizontal codewords has again improved since the drop in the error (now erasure) rate experienced by the last 16 symbols more than offsets the 10^{-6} chance of decoder failure caused by an excess of erasures in the vertical codewords.

The average redundancy of each codeword in the block is $16 + 16(32/223) = 18.29$, which gives a code rate for the block of $223/(223 + 18.29) = 0.9242$. The rate of a 10-bit Reed-Solomon code that experiences a 2 percent symbol error rate and is able to achieve the same failure probability is 0.917.

The encoder complexity increase produced by using the scheme shown in Fig. 3 will be the fractional increase in the number of codewords required, 16/223, just as it was in the case of erasure. The decoder may now, however, experience a decrease in complexity since it never needs to cope with the situation where more than 28 of the horizontal codewords have more than eight errors. The design of Reed-Solomon decoders often exploits the typically small number of errors per codeword by making the decoding time a random variable and employing a buffer. However, the scheme presented here allows an explicit upper bound on the required buffer size. This decoder simplification is another reason for declaring failure before exhausting all possibility of success.

Sections II and III have presented techniques for analyzing the decoder failure probabilities of single-stage arrays of codes when the symbol error rate is known. The essential technique used in those sections was the division of the vertical codeword failure mechanisms into two groups: those that affect the entire block of codewords and those that affect each codeword independently. For economy of expression, the division was made complete by the use of a union bound. A more refined horizontal decoder failure probability estimate would have to consider all different possible numbers of erasures in the vertical codewords; the methods involved are straightforward but lengthy extensions of the arguments in Sections II and III.

IV. Optimum-Distance Single Field Codes

The approach of Sections II and III was to take an existing coding system and improve its rate without substantially increasing its complexity. The total distance of the entire code block never factored into the design because the channel error and erasure rates were substantial. However, the designs developed did maintain the free distance of the entire block. This section considers to what extent the number of redundant symbols in a block of crossed maximum distance separable (MDS) codes can be reduced if the only requirement is that a minimum distance be maintained for the entire block. This section will present optimum constructions for single- and double-error correcting codes. In addition to quiet communications channels, such low distance codes are important for disk drive arrays and computer memory applications. The techniques used can be extended to higher distances; however, different constructions not based on crossed sets of MDS codes can yield higher rates when the distance is greater than five. Nevertheless, [3] shows that other types of multilevel codes may still be useful for high distance applications because of their economy of implementation and adaptability to channel error statistics.

Figure 4 shows a means of constructing a single-error or double-erasure correcting code with $(N-2)^2$ information symbols and three redundant symbols out of length

N Reed-Solomon codes. The distance of the code is most easily demonstrated by the erasure decoding algorithm. If the two possible erasures are in different horizontal codewords, then the first vertical codeword will experience two erasures and will decode. Each of the two horizontal codewords with an erasure will receive one symbol of redundancy from the first vertical word and, so, will be able to decode. The second vertical codeword provides for the case where both erasures occur in the same horizontal codeword. Thus, since the code can correct two erasures, it has distance three.

The procedure for correcting a single error is slightly more complex. As before, the first step involves calculating the information symbols of the vertical codewords by re-encoding the information in the horizontal codewords. If there has only been one error, then all of the symbols in the first vertical codeword except one are correct. Furthermore, the symbol in the first vertical codeword coming from the horizontal codeword containing the error is guaranteed to be incorrect since the first redundant symbol together with the information symbols of each horizontal codeword form a code of distance 2, which is single-error detecting. Thus, the decoding of the first vertical codeword pinpoints the horizontal codeword with the error. The second vertical codeword can now be decoded by declaring an erasure in the marked position and, so, both redundant symbols will be available to the damaged horizontal codeword. The construction in Fig. 4 is asymptotically close to being a perfect code since $[(N-2)^2(N-1)]/N^3$ approaches one as N becomes large. Figure 5 shows a construction for a distance five code. Its efficiency is, however, no better than can be achieved by concatenating symbols and using double-length Reed-Solomon codes; its advantage is easy encoding and decoding.

V. Summary and Discussion

The technique presented in this article allows a clear improvement in code rate of symbol-error-correcting and symbol-erasure-correcting codes for any given decode failure probability. Moreover, the computational cost to both the encoder and decoder is negligible. The only price to be paid is in interleaving and buffering, i.e., decoding cannot be completed until all of the codewords in an entire array are received.

An encoder will customarily use interleaving anyway to make symbol errors independent. Some number of Reed-Solomon codewords, e.g., eight 255 symbol codewords for the Galileo S-band, will be interleaved to form an interleaving block. All codewords in an interleaving block must come from different arrays to preserve independence. The amount of storage required at the encoder is not large, however, since only the vertical redundancy needs to be saved until the entire interleaved collection of arrays has been sent.

The overall encoding and decoding cost of the class of codes presented in this article proves to be substantially superior to that of longer Reed-Solomon codes even if large and small symbols experience the same error rate. Real communications channels, e.g., those employing determinate-state decoding [2], favor small symbols. For these channels, cannibalistic Reed-Solomon codes can offer very large improvements.

References

[1] R. J. McEliece and L. Swanson, "On the Decoder Error Probability for Reed-Solomon Codes," *IEEE Transactions on Information Theory*, IT-32, pp. 701–703, September 1986.

[2] O. Collins and M. Hizlan, "Determinate State Convolutional Codes," *The Telecommunications and Data Acquisition Progress Report 42-107, vol. July–September*, Jet Propulsion Laboratory, Pasadena, California, pp. 36–56, November 15, 1991.

[3] K. Abdel-Ghaffar and M. Hassner, "Multilevel Codes for Data Storage Channels," *IEEE Transactions on Information Theory*, IT-37, pp. 735–741, May 1991.

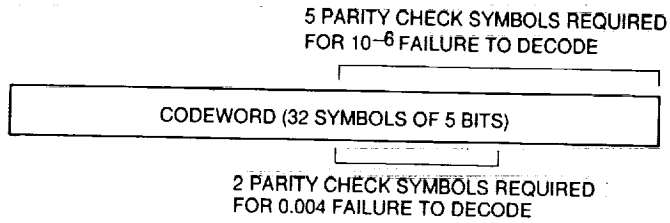


Fig. 1. Five-bit extended Reed-Solomon code.

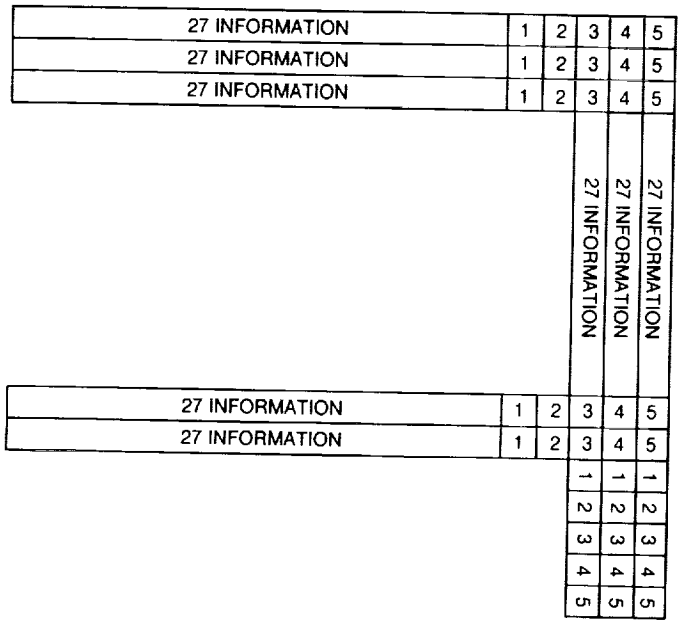


Fig. 2. Redundancy sharing.

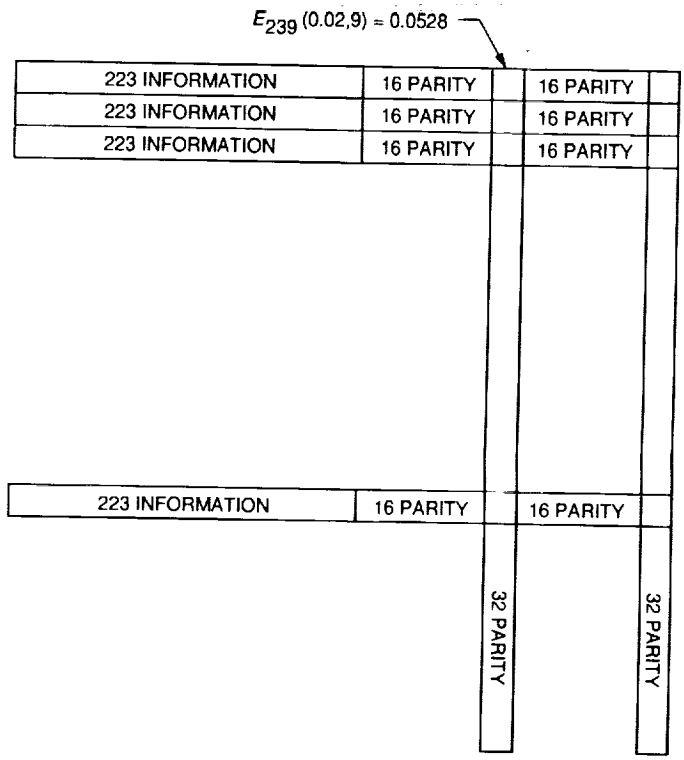


Fig. 3. Using the NASA Standard (255,223) code.

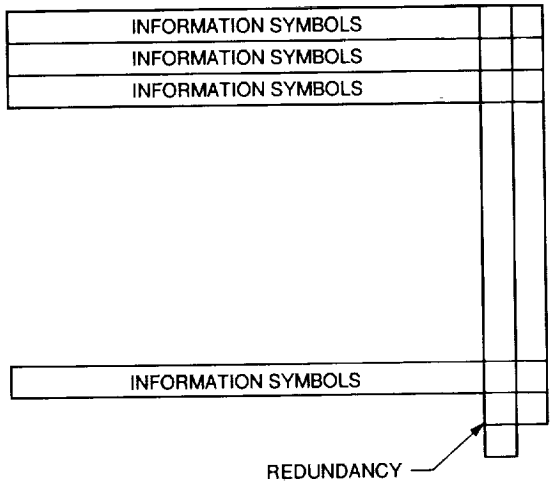


Fig. 4. A means of constructing a single-error or double-erasure correcting code.

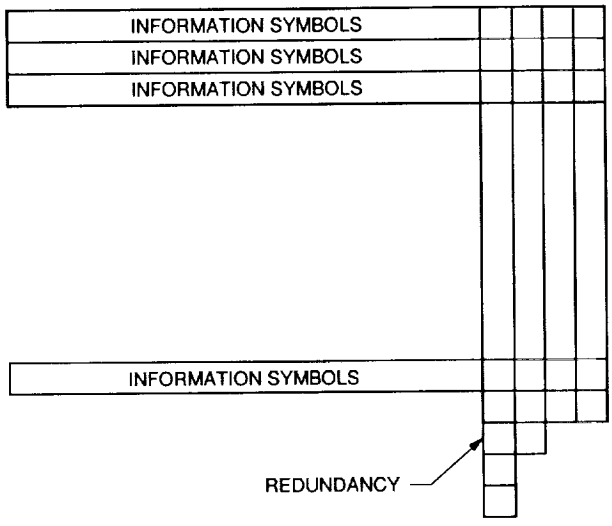


Fig. 5. A construction for a distance five code.

TDA Progress Report 42-114

August 15, 1993

Estimating the Size of Huffman Code Preambles

R. J. McEliece
Communications Systems Research Section

T. H. Palmatier
California Institute of Technology

In this article, data compression via block-adaptive Huffman coding is considered. The compressor consecutively processes blocks of N *data symbols, estimates source statistics by computing the relative frequencies of each source symbol in the block, and then synthesizes a Huffman code based on these estimates. In order to let the decompressor know which Huffman code is being used, the compressor must begin the transmission of each compressed block with a short preamble or header file. This file is an encoding of the list* $\mathbf{n} = (n_1, n_2, \ldots, n_m)$, *where* n_i *is the length of the Huffman codeword associated with the* i*th source symbol. A simple method of doing this encoding is to individually encode each* n_i *into a fixed-length binary word of length* $\log_2 l$, *where* l *is an a priori upper bound on the codeword length. This method produces a maximum preamble length of* $m \log_2 l$ *bits. The object of this article is to show that, in most cases, no substantially shorter header of any kind is possible.*

I. Introduction

Huffman data compression is optimal for sources with known statistics ([4], Chapter 10). However, in adaptive implementations, in which the Huffman code is determined empirically by the data, the recipient of the compressed data will not know which code is being used. One way for the transmitter to identify the code is to prefix the encoded data with an ordered list of the codeword lengths being used. The receiver can then synthesize a variable-length code with these lengths, using a prearranged algorithm. (For example, the decoder can use a "greedy" algorithm, in which the shortest codewords are generated first, then the next shortest words, etc.) The object of this article is to demonstrate that, in many cases, this simple scheme is near optimal, by showing that any scheme for specifying the code will use almost as many bits as the simple preamble scheme just described.

Throughout, it will be assumed that both the compressor and decompressor know that there are m codewords, and that each codeword has a maximum length of l. One such bound can be obtained by observing that no codeword in a Huffman code with m words can be longer than $m-1$. However, in many situations, this bound can be improved upon. For example, suppose the compressor works by partitioning the source sequence into blocks of N symbols, and then estimates the source statistics as $p_i = N_i/N$, where N_i denotes the number of times the ith source symbol occurs in the block. A Huffman code for the p_i's is then synthesized and used to compress the block. The maximum

Huffman codeword length corresponds to the least probability in the source probability table, which is at least $1/N$. In a recent paper [2], Abu-Mostafa and McEliece showed that the longest Huffman codeword has a maximum length of $1.44 \log_2 p^{-1}$ for a source whose smallest probability is p, and that no better bound is possible. Thus, for an adaptive Huffman scheme, each Huffman codeword will have a maximum length of $1.44 \log_2 N$, which is often considerably smaller than $m - 1$.

In the simple scheme, the code preamble will be an ordered list $(n_1, n_2, \ldots, n_m)$, where n_i is the length of the Huffman codeword for the ith symbol, and m is the number of symbols in the source alphabet. Since $n_i \leq l$, for $i = 1, 2, \ldots, m$, each n_i can be represented by a $\lceil \log_2 l \rceil$-bit binary word, the length of the preamble in bits will be $m \lceil \log_2 l \rceil$. In a slightly more sophisticated scheme, the list $(n_1, n_2, \ldots, n_m)$ can be regarded as an l-ary representation of a large integer and converted to binary, which will require at most $\lceil \log_2 l^m \rceil = \lceil m \log_2 l \rceil$ bits. In either case, one can say that the simple preamble requires around $m \log_2 l$ bits.

On the other hand, if $N(m, l)$ denotes the total number of lists $(n_1, n_2, \ldots, n_m)$ that can possibly occur as length lists for Huffman codes whose codeword lengths are at most l, then at least $\log_2 N(m, l)$ bits are needed to specify one of them. In the next section it will be shown that $\log_2 N(m, l)$ is near $m \log l$ in many cases, which implies that any scheme for specifying the Huffman code must use almost as many bits as the simple preamble scheme.

II. Main Result

It is well-known ([4], Chapter 10) that the codeword lengths in a Huffman code must satisfy the Kraft–McMillan equation

$$\sum_{i=1}^{m} 2^{-n_i} = 1 \tag{1}$$

In this section, an estimate will be obtained for the number of ordered solutions of Eq. (1), where each n_i is further restricted to lie in the range $1 \leq n_i \leq l$. For future reference, denote this number by $N(m, l)$. A first observation is that since there are l possible values for each of the n_i's, then $N(m, l) \leq l^m$. Equivalently, if one defines $B(m, l) = \log_2 N(m, l)$ (the number of bits required to specify an arbitrary ordered solution to Eq. (1), if $1 \leq n_i \leq l$), one has

$$B(m, l) \leq m \log_2 l \tag{2}$$

In a sense, the object of this article is to show that the upper bound in Eq. (2) is quite good; thus, in the rest of the section, lower bounds on $B(m, l)$ will be considered. The key to these bounds is the observation that if $\mathbf{n} = (n_1, \ldots, n_m)$ is a particular solution to Eq. (1), any permutation of the components of $\mathbf{n}$ will also be a solution. Indeed, if g_j denotes the multiplicity of the integer j as a component of $\mathbf{n}$, then there are exactly

$$\binom{m}{g_1, g_2, \ldots, g_l} = \frac{m!}{g_1! g_2! \cdots g_l!} \tag{3}$$

distinct solutions to Eq. (1) that can be generated by permuting the components of $\mathbf{n}$. For example, with $m = 4$, $l = 3$, the unordered solution $(1, 3, 3, 3, 3)$ to Eq. (1) yields 5 ordered solutions, and the unordered solution $(2, 2, 2, 3, 3)$ yields 10 unordered solutions, so that $B(5, 3) = 15$.

As a first step towards the general results, consider solutions to Eq. (1) with no restrictions on l. Since the longest word in a Huffman code with m words is $m - 1$, this is equivalent to taking $l = m - 1$. In this case, the upper bound in Eq. (2) is $m \log_2(m - 1)$. On the other hand, the particular unordered solution $(1, 2, 3, \ldots, m-2, m-1, m-1)$ to Eq. (1) yields, upon permutation of its components, $m!/2$ ordered solutions, so that $B(m, m-1) \geq \log_2 m!/2$. Thus

$$\log_2(m!/2) \leq B(m, m-1) \leq m \log_2(m-1) \tag{4}$$

It follows from Stirling's approximation to the factorial ([3], Section 1.2.11) that

$$\lim_{m \to \infty} \frac{\log(m!/2)}{m \log(m-1)} = 1$$

which means that for large m, $B(m, m-1) \sim m \log_2 m$ (see Theorem 1, below).

For restricted solutions to Eq. (1), i.e., cases when $l < m - 1$, one can do something very similar. The idea is again to find a particular solution to Eq. (1) with as many distinct permutations as possible. To facilitate the discussion, now rewrite Eq. (1) as

$$\sum_{j=0}^{l} g_j 2^{-j} = 1 \tag{5}$$

where g_j denotes the multiplicity of the integer j in the list $\mathbf{n} = (n_1, \ldots, n_m)$. The goal is to maximize the multinomial coefficient of Eq. (3), subject to Eq. (5).

The following construction yields a family of particular solutions to Eq. (5), for which the multinomial coefficient of Eq. (3) is relatively large, and which therefore provides reasonably good lower bounds for $B(m,l)$. For a given value of l, choose integers u, r, and s such that

$$3 \le u \le l-2, \quad 1 \le r \le u-2, \quad 1 \le s \le 2^r - 2 \tag{6}$$

Now define $g_0, g_1, \ldots, g_l$ as follows:

$$\begin{aligned} g_0 &= \cdots = g_{r-1} = 0 \\ g_r &= s \\ g_{r+1} &= \cdots = g_{u-1} = 2^r - s \\ g_u &= 2^r - s + 1 \\ g_{u+1} &= \cdots = g_{l-1} = 2^r - s - 1 \\ g_l &= 2(2^r - s - 1) \end{aligned} \tag{7}$$

It is then routine but tedious to verify algebraically that Eq. (5) holds. However, it is much easier to see that this is true by visualizing a binary tree with g_j external nodes at level j, for $j = 0, 1, \ldots, l$. Figure 1 is such a tree, for $l = 7$, $u = 5$, $r = 2$, and $s = 1$. In general, such a tree has

$$m = \sum_{j=0}^{l} g_j = s + (2^r - s)(l + 1 - r) - (l - u) \tag{8}$$

external nodes. For example, in Fig. 1 there are $m = 17$ external nodes. It thus follows that for any choice of l, u, r, and s satisfying (6), the numbers $(g_0, \ldots, g_l)$ defined in Eq. (7) give a particular solution to Eq. (5), and so

$$B(m,l) \ge \log_2 \binom{m}{g_1, g_2, \ldots, g_l}$$

where m is given by Eq. (8).

For example, since there are $m = 17$ external nodes on the Fig. 1 tree, it follows that

$$B(17,7) \ge \log_2 \binom{17}{1,3,3,4,2,4} = 33.00$$

On the other hand, from Eq. (2), $B(17,7) \le 16 \log_2 7 = 44.92$ so that, at least on a logarithmic scale, the particular solution to Eq. (5) represented by the binary tree depicted in Fig. 1, together with its permutations, accounts for a substantial fraction of the total number of ordered solutions to the Kraft–McMillan [Eq. (1)].

The same kind of thing happens in general. That is, the largest multinomial coefficient of the form of Eq. (3), where the g_j's are given by Eqs. (6) and (7), is nearly always close to $m \log_2 l$. To see why this is so, one further specializes the solution to Eq. (5) given by Eq. (7). First, notice that Eq. (8) implies that

$$m \le 2^r (l + 1 - r) \tag{9}$$

Choose r to be the least integer such that Eq. (9) holds. Second, having chosen r, notice that Eq. (8) implies that

$$m \le s + (2^r - s)(l + 1 - r) \tag{10}$$

Thus, choose s to be the largest integer that Eq. (6) holds. Explicitly,

$$s = \left\lfloor \frac{2^r (l + 1 - r) - m}{l - r} \right\rfloor \tag{11}$$

Finally, having chosen both r and s, u is determined by Eq. (8), i.e.,

$$u = m + l - s - (2^r - s)(l + 1 - r) \tag{12}$$

In this way, the numbers u, r, and s are uniquely determined by m and l, as are the g_j's in Eq. (7), which in turn define the multinomial coefficient of Eq. (3). Define the logarithm of this multinomial coefficient as $B'(m,l)$. Thus, from the foregoing discussion

$$B(m,l) \ge B'(m,l)$$

For example, if $m = 20$ and $l = 8$, the least value of r satisfying Eq. (9) is $r = 2$. Then from Eq. (11), one finds that $s = 1$, and from Eq. (12), that $u = 6$. Thus from Eq. (7), $g_0 = g_1 = 0$, $g_2 = 1$, $g_3 = g_4 = g_5 = 3$, $g_6 = 4$, $g_7 = 2$, $g_8 = 4$. Finally

$$B'(20,8) = \log_2 \binom{20}{1,3,3,3,4,2,4} = 43.15$$

as compared to the upper bound of Eq. (2) $B(20,8) \leq 20 \log_2 8 = 60$.

A theorem will next be presented to illustrate what happens if l is a fixed fraction of m (e.g., $l = \beta m$ and $m \to \infty$), to further substantiate the claim that the upper bound in Eq. (2) is usually fairly tight.

Theorem 1. *For any fixed β, with $0 < \beta \leq 1$, there are positive constants K_1 and K_2 (dependent on β but independent of m) such that for sufficiently large m,*

$$m \log_2 m - K_2 m \leq B(m, \beta m) \leq m \log_2 m - K_1 m \quad (13)$$

Proof: The upper bound from Eq. (2), $m \log \beta m$, is asymptotically given by

$$m \log \beta m = m \log m - \log \beta^{-1} m$$

which proves the upper bound in Eq. (13), with $K_1 = \log_2 \beta^{-1}$.

To obtain an asymptotic lower bound on $B(m, \beta m)$, the reasoning is as follows. For sufficiently large m, the smallest integer solution r to Eq. (9) (using $l = \beta m$) is

$$r = r(\beta) = \lfloor \log_2 \beta^{-1} \rfloor + 1 \quad (14)$$

which is a constant, independent of m. Note then from Eq. (7) that

$$g_j \leq 2^r \qquad \text{for } j = r, r+1, \ldots, l-1$$

$$g_l \leq 2^{r+1}$$

Thus

$$\binom{m}{g_1, \ldots, g_l} \geq \frac{m!}{(2^r!)^{l-r}(2^{r+1}!)}$$

This means that one can estimate $B(m, \beta m)$ using Stirling's formula, as follows:

$$B(m, \beta m) \geq \log_2 \binom{m}{g_1, \ldots, g_l} \geq \log m!$$

$$- (l-r) \log 2^r! - \log 2^{r+1}!$$

$$\sim m \log m - (1 + \beta \log 2^r!) m + O(\log m)$$

which gives a lower bound of the form promised in Eq. (13).

In conclusion, some brief remarks about the constants in Eq. (13) will be provided. When $\beta = 1$, i.e., when there are no restrictions on the length of the codewords, then the lower bound derived in Eq. (4) implies

$$B(m, m-1) \geq m \log_2 m - \log_2 em = m \log_2 m - 1.4427m$$

for sufficiently large m. On the other hand, [1] shows that T_m, the total number of *unordered* solutions to Eq. (1), satisfies $\log_2 T_m/m \to \lambda = 1.794....$ But the total number of ordered solutions to Eq. (1) cannot exceed $m!T_m$, which means (by Stirling's formula)

$$B(m, m-1) \leq m \log_2 m - \log_2 \frac{e}{\lambda} m = m \log_2 m - 0.5994m$$

For general β, from the proof of Theorem 1, it follows that K_1 can be taken to be

$$K_1(\beta) = \log \beta^{-1} \quad (15)$$

and that K_2 can be taken as

$$K_2(\beta) = 1 + \beta \log_2(2^r!) \quad (16)$$

where r is given by Eq. (14). Using the same method, but with a little more work, K_2 can be improved. Indeed, it can be shown that

$$K_2(\beta) = 1 + \alpha \log(2^r - s)! + (\beta - \alpha) \log(2^r - s - 1)!$$

where

$$r = \lfloor \log b^{-1} \rfloor + 1$$

$$s = 2^r - \lfloor b^{-1} \rfloor$$

$$\alpha = \beta - \left(1 - \frac{\lfloor b^{-1} \rfloor}{2^r}\right)$$

For example, with $\beta = 1/3$, Eqs. (15) and (16) give

$$K_1(1/3) = 1.58, \quad K_2(1/3) = 1.908$$

The determination of the best possible constants in Eq. (13) is an interesting and important unsolved problem.

References

[1] D. W. Boyd, "The Asymptotic Number of Solutions of a Diophantine Equation from Coding Theory," *J. Comb. Theory*, vol. (A) 18, pp. 210–215, 1975.

[2] Y. S. Abu-Mostafa and R. J. McEliece, "Maximal Codeword Lengths in Huffman Codes," *The Telecommunications and Data Acquisition Progress Report, 42-110, vol. April–June 1992*, Jet Propulsion Laboratory, Pasadena, California pp. 188–193, August 15, 1992.

[3] D. E. Knuth, *The Art of Computer Programming, vol. 1: Fundamental Algorithms*, Reading, Massachusetts: Addison-Wesley, 1968.

[4] R. J. McEliece, *The Theory of Information and Coding*, Reading, Massachusetts: Addison-Wesley, 1977.

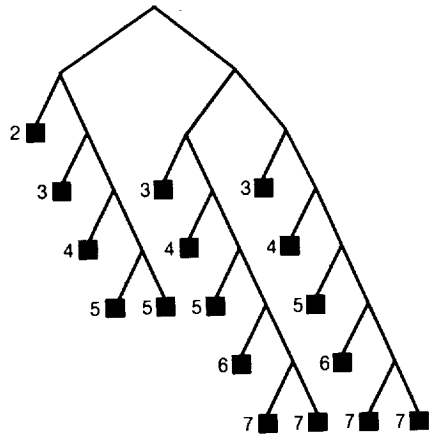


Fig. 1. A high entropy binary tree, for $m = 17$, $l = 7$. (Here $u = 5$, $r = 2$, and $s = 1$. Refer to Eq. (7).) The level of each external vertex is indicated.

TDA Progress Report 42-114

August 15, 1993

Enhanced Decoding for the Galileo S-Band Mission

S. Dolinar and M. Belongie
Communications Systems Research Section

A coding system under consideration for the Galileo S-band low-gain antenna mission is a concatenated system using a variable redundancy Reed-Solomon outer code and a (14,1/4) convolutional inner code. The 8-bit Reed-Solomon symbols are interleaved to depth 8, and the eight 255-symbol codewords in each interleaved block have redundancies 64, 20, 20, 20, 64, 20, 20, and 20, respectively (or equivalently, the codewords have 191, 235, 235, 235, 191, 235, 235, and 235 8-bit information symbols, respectively). This concatenated code is to be decoded by an enhanced decoder that utilizes (1) a maximum likelihood (Viterbi) convolutional decoder; (2) a Reed-Solomon decoder capable of processing erasures; (3) an algorithm for declaring erasures in undecoded codewords based on known erroneous symbols in neighboring decodable words; (4) a second Viterbi decoding operation (redecoding) constrained to follow only paths consistent with the known symbols from previously decodable Reed-Solomon codewords; and (5) a second Reed-Solomon decoding operation using the output from the Viterbi redecoder and additional erasure declarations to the extent possible.

It is estimated that this code and decoder can achieve a decoded bit error rate of 1×10^{-7} at a concatenated code signal-to-noise ratio of 0.76 dB. By comparison, a threshold of 1.17 dB is required for a baseline coding system consisting of the same (14,1/4) convolutional code, a (255,223) Reed-Solomon code with constant redundancy 32 also interleaved to depth 8, a one-pass Viterbi decoder, and a Reed-Solomon decoder incapable of declaring or utilizing erasures. The relative gain of the enhanced system is thus 0.41 dB.

It is predicted from analysis based on an assumption of infinite interleaving that the coding gain could be further improved by approximately 0.2 dB if four stages of Viterbi decoding and four levels of Reed-Solomon redundancy are permitted. Confirmation of this effect and specification of the optimum four-level redundancy profile for depth-8 interleaving is currently being done.

I. Introduction

This article looks at the performance of a Reed-Solomon plus convolutional concatenated coding system with enhanced decoding as planned for the Galileo S-band low-gain antenna (LGA) mission. The baseline system without enhanced decoding uses a (255,223) Reed-Solomon outer code concatenated with a (14,1/4) convolutional inner code, and interleaves the Reed-Solomon symbols to depth 8. The convolutionally encoded symbols are decoded by maximum likelihood (Viterbi) decoding, and each Reed-Solomon codeword is decoded algebraically to correct a maximum of sixteen 8-bit symbols per 255-symbol word. There are two types of decoding enhancements planned for the Galileo LGA mission: *Reed-Solomon redecoding* using erasure declarations and *Viterbi redecoding* using Reed-Solomon corrected symbols.

Reed-Solomon redecoding is possible when at least one but fewer than eight of the codewords within a block of eight interleaved words is decodable (correctable). The Reed-Solomon decoder can then extrapolate the locations of corrected errors in the decodable word(s) to neighboring locations in adjacent undecodable word(s) and declare the corresponding symbols to be erased. The basis for this error extrapolation or *error forecasting* is that errors from the Viterbi decoder tend to occur in bursts that are often longer than the 8-bit Reed-Solomon symbols. If the erased symbols are highly likely to be erroneous, then the undecoded words might be decoded by a second try at Reed-Solomon decoding that utilizes the erasure information.

Viterbi redecoding consists of an extra pass through a maximum likelihood decoder that is now constrained to follow only paths consistent with the known symbols from previously decodable Reed-Solomon codewords. The Viterbi redecoder is much less likely to choose a long erroneous path because any path under consideration is pinned to coincide with the correct path at the location(s) of the known symbols.

Both redecoding processes may be repeated an arbitrary number of times (with diminishing returns). Repetitions of Reed-Solomon redecoding allow the testing of several different combinations of reasonable error extrapolations from the known error locations, and can be accomplished relatively cheaply, as Reed-Solomon decoding is much faster than Viterbi decoding. Each repetition of Viterbi redecoding requires a whole new *decoding stage*: it must begin with the output from the previous Reed-Solomon decoding stage and must feed its output to another following Reed-Solomon stage.

With both types of redecoding, it usually pays to put different amounts of redundancy in neighboring Reed-Solomon codewords. Words with high redundancy can be counted on to decode during an initial decoding try, and the information from these decoded words can be used to assist the decoding of codewords with lower redundancy later.

In this article, the stages of the enhanced decoding process are denoted as follows:

VIT-1 the first pass of the raw data through the Viterbi decoder.

RS-1 the first pass of the Viterbi decoded bits through the Reed-Solomon decoder, including the possibility of several trials per codeword using erasure declarations forecast from previously decoded codewords.

VIT-2 the second pass of the raw data through the Viterbi decoder, aided this time by known bits from codewords successfully decoded in RS-1.

RS-2 the pass of the Viterbi redecoded bits from VIT-2 through the Reed-Solomon decoder, again using erasure declarations to assist in decoding previously undecodable codewords.

⋮ ⋮

VIT-n the *nth* pass of the raw data through the Viterbi decoder, aided this time by known bits from codewords successfully decoded in RS-$(n-1)$.

RS-n the pass of the Viterbi redecoded bits from VIT-n through the Reed-Solomon decoder, again using erasure declarations to assist in decoding previously undecodable codewords.

II. Three Analysis Approaches

The analysis in this article is first performed using an assumption of infinite interleaving and no Reed-Solomon erasure declarations. This is done for one, two, and four Viterbi decoding stages. A calculation is made of the maximum coding gain obtainable by adding the extra stages. With the infinite interleaving assumption, concatenated system performance can be accurately obtained to error rates of 1×10^{-7} or lower, based on megabits of simulated Viterbi decoder data.

The analysis then continues for the actual Galileo conditions with depth-8 interleaving but still no Reed-Solomon erasure declarations. This analysis cannot directly verify 1×10^{-7} error rates even with gigabits of Viterbi decoded data, but extrapolated performance curves for depth-8 interleaving can be relied on with high confidence because of their similarity to certain corresponding curves derived under the infinite interleaving assumption.

Finally, the depth-8 interleaving analysis is extended to allow Reed-Solomon erasure declarations using a simple error forecasting strategy.

The analysis in this article characterizes the concatenated decoding system based on the usual assumption that the received symbols are corrupted by stationary additive white Gaussian noise. No attempt has been made to refine the decoder models to account for effects due to such factors as imperfect arraying, phase drifts, or time-varying signal-to-noise ratio.

III. Analysis Based on Infinite Interleaving and No Erasure Declarations

This section analyzes the gains possible from Viterbi redecoding alone, without Reed-Solomon erasure declarations, but using an assumption of infinite interleaving. It begins with a baseline analysis of one-stage decoding (no redecoding) and proceeds to evaluate the gains possible from using two stages and four stages.

A. One-Stage Decoding

One-stage decoding for this case consists of one pass through the Viterbi decoder followed by one pass without erasure declarations through the Reed-Solomon decoder. The performance of the Viterbi decoder is evaluated by simulation, and that of the Reed-Solomon decoder is calculated from a formula.

1. Stage VIT-1. The Viterbi decoder for the Galileo (14,1/4) convolutional code is characterized by the performance curves given in Fig. 1. The curves show Viterbi decoder bit error rate BER and 8-bit Reed-Solomon symbol error rate SER as functions of the signal-to-noise ratio E_b/N_o of the convolutional code. The points on these curves are based on about 2 Gbits of data each in the range from 0.0 to 0.5 dB, and on more than 400 Mbits each in the range from 0.6 dB to 1.0 dB (but 2 Gbits at 0.7 dB). All of the data were decoded by the Big Viterbi Decoder (BVD).

Figure 1 shows that the (14,1/4) Galileo LGA code achieves a BER of 5×10^{-3} at an E_b/N_o of 0.65 dB. By comparison, Tables 1 and 2 of [1] show that the (15,1/4) Galileo high-gain antenna (HGA) code achieves a BER of 5×10^{-3} at an E_b/N_o of 0.52 dB. The (14,1/4) LGA code is 0.13 dB inferior to the (15,1/4) HGA code at this error rate.

2. Stage RS-1. Figure 2 shows the performance of the Reed-Solomon decoder under an assumption of infinite interleaving and no erasure declarations, but with varying amounts of error correction E (or codeword redundancy $2E$). The E values for the various curves are labeled along the right and bottom edges of the graph. These curves are based on the same BVD simulation runs described for Stage VIT-1. On this figure, and on all succeeding figures, the signal-to-noise ratio E_b/N_o on the x-axis is the same as that in Fig. 1, i.e., the signal-to-noise ratio for the inner convolutional code alone. Thus, these curves cannot be used for a direct reading of the optimum value of E for single-stage Reed-Solomon decoding because they do not reflect the increasing dilution of signal energy due to increasing redundancy of the Reed-Solomon code. However, this latter adjustment is a simple numerical computation, and presenting a parametric family of performance curves as those in Fig. 2 leads to effective trial-and-error optimization methods for the multistage decoding processes analyzed below.

The curves in Fig. 2, and in all succeeding figures, show the Reed-Solomon decoded 8-bit SER rather than the decoded BER. This is a convenience because the analysis of future decoding stages depends primarily on SER rather than BER. The BER is slightly less than half the corresponding SER, as can be seen by comparison of the BER and SER curves in Fig. 1.

3. Performance of One-Stage Decoding. For the baseline coding system using the (255,223) Reed-Solomon code, E is 16 and a one-stage decoding system with infinite interleaving can achieve an SER of 2×10^{-7} (i.e, a BER of $\sim 1 \times 10^{-7}$) at convolutional code $E_b/N_o = 0.52$ dB. The corresponding concatenated code signal-to-noise ratio is $E_b/N_o = 1.10$ dB, accounting for the 0.58 dB redundancy of the (255,223) Reed-Solomon code.

Alternatively, if a (255,231) Reed-Solomon code with $E = 12$ had been used, the required convolutional code operating point would move to $E_b/N_o = 0.67$ dB, but the corresponding concatenated code signal-to-noise ratio would stay at $E_b/N_o = 1.10$ dB, because the redundancy of the (255,231) Reed-Solomon code is only 0.43 dB. In fact, the lowest required concatenated code signal-to-noise

ratio of approximately 1.10 dB is achieved over a range of values of E from 12 to 16, or equivalently over a range of convolutional code operating points E_b/N_o from 0.52 dB to 0.67 dB.

B. Two-Stage Decoding

Two-stage decoding consists of two passes each through the Viterbi decoder and the Reed-Solomon decoder. The only difference in the decoding algorithms is that the Viterbi decoder on the second pass is constrained to follow only paths consistent with known symbols from Reed-Solomon codewords that were decoded on the first pass. The Reed-Solomon decoding algorithm is unchanged for both passes.

1. Stage VIT-1. The output of Stage VIT-1 is characterized by the same performance curves given in Fig. 1.

2. Stage RS-1. The output of Stage RS-1 is characterized by the same performance curves given in Fig. 2. However, now it is important to utilize the full set of information contained in all the curves in Fig. 2 rather than just to focus on the values of E from 12 to 16 that give optimum one-stage performance.

3. Stage VIT-2. The performance of Stage VIT-2 is dependent on how often known 8-bit symbols from the previous Reed-Solomon decoding are inserted into the Viterbi decoder data stream. If the Reed-Solomon code has constant redundancy, the known symbols will occur with a certain average frequency (determined by the corresponding SER curve in Fig. 2), but not according to any repeating pattern. If the Viterbi redecoder had to contend with only randomly occurring unknown symbols interspersed with probability $SER < 10^{-1}$ into a stream of mostly known symbols, Stage VIT-2 would produce an almost perfect output at any E_b/N_o shown in Fig. 2. However, even with infinite interleaving, Stage RS-1 decoding failures will be correlated from one codeword to the next due to the burstiness of the errors from Stage VIT-1. Therefore, the Viterbi redecoder is likely to encounter strings of unknown symbols of unpredictable length. Long strings of unknown symbols will locally negate the value of the otherwise highly likely known symbols and cause correspondingly long strings of redecoded errors. No effective analytical techniques have been developed to quantitatively assess the performance of Stage VIT-2 after decoding a constant redundancy code in Stage RS-1.

If the Reed-Solomon codewords have variable redundancy and there is an appreciable difference in redundancy values, then it is possible (and reasonable) to set the E_b/N_o operating point so that the most highly redundant codewords are almost certain to decode, while the less redundant codewords are undecodable with probability much higher than the target error probability for the multistage decoding process. Then, to a good conservative approximation, it can be assumed that the known symbols from Stage RS-1 will occur in at least the same repeating pattern as the pattern of the most highly redundant codewords. Under these conditions, the performance of Stage VIT-2 can be effectively parameterized by feeding the Viterbi redecoder various repeating patterns of known symbols.

Viterbi redecoded BER and SER tables were presented in [2] for three different known symbol patterns (one in two, one in four, and one in eight) and three different convolutional codes (the (7,1/2) NASA standard code, an (11,1/4) code, and the (15,1/4) Galileo HGA code). A similar analysis for the (14,1/4) Galileo LGA code was performed for this article, and the SER results are shown in Fig. 3. These curves are based on slow-running software simulations of only about 1 Mbit each, because the high-rate BVD used to generate the 2-Gbit samples for Stage VIT-1 could not without modification incorporate the known symbol information into its decoding process.

Figure 3 shows multiple curves for the one-in-four and one-in-eight known symbol cases, because, as noted in [2], the redecoded SER depends on the symbol's phase relative to the position of the nearest known symbol in the pattern. It can be argued theoretically and has been confirmed by simulation that phases $\pm n$ have approximately the same SER, so the curves in Fig. 3 represent averages of the simulated SER values for $\pm n$ ($n = 1, 2$ for the one-in-four case and $n = 1, 2, 3, 4$ for the one-in-eight case).

Figure 3 has two different labels for its y-axis. The label on the left is the actual SER observed from the simulation runs. The label on the right is the equivalent convolutional code E_b/N_o that produces the same SER values without any known symbols, i.e., the convolutional code signal-to-noise ratio that would have been required to produce the same SER with one-stage Viterbi decoding.

4. Stage RS-2. The output of Stage RS-2 is characterized by the same family of performance curves given in Fig. 2. Under the infinite interleaving assumption, all symbols in a given codeword are independent, so the decoding error rate depends only on the SER from Viterbi decoding during Stage VIT-2. Thus, the curves in Fig. 2 are directly applicable to the determination of the performance of Stage RS-2 if the equivalent one-stage E_b/N_o values are used to characterize the output of Stage VIT-2.

For example, suppose that every fourth codeword is *strong* (i.e., highly redundant) and can be presumed to decode with high probability during Stage RS-1. Then for Stage VIT-2, there are known symbols spaced four symbols apart, and the curves labeled 1/4 and 2/4 in Fig. 3 are applicable. If the convolutional code operating point is $E_b/N_o = 0.20$ dB, then the equivalent Stage VIT-1 operating point is 0.85 dB for symbols at phases ± 1 and 0.77 dB for symbols at phase ± 2. If a Reed-Solomon code with $E = 9$ is used for the remaining weak (i.e., less redundant) codewords, Fig. 2 shows that the SER for the weak codewords is approximately 1×10^{-7} for codewords at phases ± 1 and 1×10^{-6} for codewords at phase ± 2.

5. Performance of Two-Stage Decoding. The overall performance of two-stage decoding depends on the likelihood of decoding both the strong codewords during Stage RS-1 and the weak codewords during Stage RS-2. The SER of the strong codewords can be conservatively presumed to equal the SER from Stage RS-1 for a code with the redundancy of the strong codewords. There are two contributions to the SER of the weak codewords. Weak codewords will fail to decode when the number of symbol errors from the redecoding stage exceeds the weak codewords' correction capability. Furthermore, weak codewords may not successfully decode whenever neighboring strong codewords do not decode during Stage RS-1. In this article, it is assumed that this second contribution to the SER of weak codewords equals the SER of the strong codewords. This leads to an approximate formula for the overall SER of the form $SER = SER_s(1) + f_w SER_w(2)$, where $SER_s(1)$ is the strong codewords' symbol error rate after Stage RS-1; $SER_w(2)$ is the weak codewords' symbol error rate after Stage RS-2, assuming known symbols during Stage VIT-2 from neighboring decodable strong codewords; and f_w is the fraction of weak codewords. If the weak codewords do not all have equivalent performance, the second term of this formula can be generalized to include a sum of contributions from each individual type of weak codeword, weighted by the fraction of each type.

If the target SER is 2×10^{-7}, then both $SER_s(1)$ and $SER_w(2)$ should be reduced to roughly 1×10^{-7}. Building upon the example introduced in the previous section with one strong codeword every four, Fig. 2 shows that a Stage RS-1 correction capability of $E = 30$ is required to bring $SER_s(1)$ to 1×10^{-7} at a convolutional code operating point of $E_b/N_o = 0.20$ dB. Similarly, the discussion of the previous example noted that $SER_w(2)$ could be brought to 1×10^{-7} for the weak codewords at phases ± 1 by using a Reed-Solomon code with a correction capability of $E = 9$. For the weak codewords at phase ± 2, $E = 9$ is insufficient because the resulting $SER_w(2)$ of 1×10^{-6} would raise the overall SER above 2×10^{-7} all by itself. However, $E = 10$ yields $SER_w(2) = 2 \times 10^{-7}$ for the weak codewords at phase ± 2, resulting in an overall $SER = 1 \times 10^{-7} + (1/2)1 \times 10^{-7} + (1/4)2 \times 10^{-7} = 2 \times 10^{-7}$. The *redundancy profile* for this codeword set is $2E = (60, 18, 20, 18)$, which corresponds to an average code rate of 229/255, or 0.52 dB of overhead for the Reed-Solomon code. The corresponding concatenated code operating point is $E_b/N_o = 0.20$ dB + 0.52 dB = 0.72 dB. This represents an improvement of 0.38 dB relative to the required operating point for the one-stage decoder.

If the target SER is still 2×10^{-7} but the convolutional code operating point is moved to 0.10 dB, the required strong codeword correction capability to achieve $SER = 1 \times 10^{-7}$ in Stage RS-1 is now $E = 36$. After Viterbi redecoding with one in four symbols known, the SER-equivalent Stage VIT-1 operating points from Fig. 2 are 0.79 dB for phases ± 1 and 0.72 dB for phase ± 2. Weak-codeword correction capabilities of $E = 10$ and $E = 11$ are sufficient to bring the corresponding $SER_w(2)$ values to 1×10^{-7} and 2×10^{-7}, respectively. The overall $SER = 1 \times 10^{-7} + (1/2)1 \times 10^{-7} + (1/4)2 \times 10^{-7} = 2 \times 10^{-7}$, as before. The redundancy profile of this codeword set is $2E = (72, 20, 22, 20)$, which corresponds to an average code rate of 443/510, or 0.61 dB of overhead for the Reed-Solomon code. The corresponding concatenated code operating point is $E_b/N_o = 0.10$ dB + 0.61 dB = 0.71 dB, which is almost exactly the same as before. As for the case of one-stage decoding, there is a range of convolutional code operating points over which the same near-optimal concatenated code performance can be achieved by appropriately adjusting the Reed-Solomon redundancies.

The case of one strong codeword every two is not directly analyzable from Fig. 2 because the equivalent convolutional code E_b/N_o is outside the range of values measured for Fig. 2. However, calculations show that for SER in the range of 0.0006 to 0.0009, a Reed-Solomon code with $E = 4$ will bring the SER for Stage RS-2 under 1×10^{-7}, but with $E = 3$ it is above 3×10^{-7}. Referring to Fig. 3 for the Stage VIT-2 SER values after redecoding with one in two symbols known, $E = 4$ will be required for the weak codewords if the convolutional code operating point is beyond about 0.1 dB. Because of the large spacing between the $E = 4$ and the $E = 3$ curves, it can be argued that $E = 3$ remains inadequate for a few tenths of a dB beyond 0.1 dB (and beyond the region in Fig. 3 for which simulation results have been obtained thus far). The choice of E for the strong codewords depends on the convolutional code operating point. For example, $E = 30$ is required at 0.20 dB, while $E = 25$ is sufficient at 0.30 dB. The corresponding redundancy profiles (60,8) and (50,8) con-

tribute Reed-Solomon coding overheads of 0.62 dB and 0.52 dB, respectively, and both of the resulting concatenated code operating points are $E_b/N_o = 0.82$ dB. This is about 0.1 dB worse than the best operating point obtainable with the scheme that uses only one strong codeword every four. It reflects the inefficiency of investing high amounts of Reed-Solomon redundancy in codewords spaced so closely together that the Stage VIT-2 SER in Fig. 3 is driven lower than necessary.

The case of one strong codeword every eight blurs the distinction between strong and weak codewords because of the relatively wide variation in Stage VIT-2 SER among the various possible symbol phases. For example, at 0.20 dB a code with $E = 30$ will bring the strong codeword SER to 1×10^{-7}. From Fig. 3, the SER-equivalent operating points for the weak codewords after Viterbi redecoding are 0.67 dB, 0.51 dB, 0.43 dB, and 0.41 dB for phases $\pm 1, \pm 2, \pm 3$, and ± 4, respectively, requiring $E = 12, 17, 19$, and 20, respectively, to bring the individual $SER_w(2)$ values to the range of 1×10^{-7} to 2×10^{-7}, and the resulting overall SER to approximately 2×10^{-7}. The corresponding redundancy profile (60, 24, 34, 38, 40, 38, 34, 24) costs 0.67 dB, which moves the concatenated code operating point to 0.87 dB. This is inferior to both the cases of one-in-four and the one-in-two strong codewords.

All of the performance results discussed in this section for two-stage decoding have been obtained by trial-and-error optimization using Figs. 2 and 3, and they may not always reflect the exact optimum performance. However, because the concatenated code operating point stays relatively constant (within a couple hundredths of a dB) over a relatively large range of convolutional code operating points (at least a tenth of a dB), the trial-and-error method gives results accurate to a few hundredths of a dB wherever data are available.

C. Four-Stage Decoding

Four-stage decoding consists of four passes each through the Viterbi decoder and the Reed-Solomon decoder. The main difference relative to two-stage decoding is that now it is profitable to design four separate levels of Reed-Solomon codeword redundancies rather than two. The analysis proceeds as in the two-stage case.

1. Stage VIT-1. The output of Stage VIT-1 is characterized by the performance curves given in Fig. 1.

2. Stages RS-1, RS-2, RS-3, and and RS-4. The outputs of Stages RS-1, RS-2, RS-3, and RS-4 are characterized by the performance curves given in Fig. 2.

3. Stages VIT-2, VIT-3, and VIT-4. The outputs of Stages VIT-2, VIT-3, and VIT-4 are characterized by the performance curves given in Fig. 3.

4. Performance of Four-Stage Decoding. With four-stage decoding, it pays to design four different categories of Reed-Solomon codewords, not just strong and weak. As an example, one codeword out of every eight could be assigned the highest redundancy, and these codewords could be expected to decode during Stage RS-1; one of eight could be assigned the next highest redundancy (spaced halfway between the highest redundancy words) and these could be expected to decode during Stage RS-2 after Stage VIT-2 Viterbi redecoding, using a pattern of one known symbol every eight. Then the third highest redundancy words (one of every four spaced halfway between the two types of higher redundancy words) will be decoded in Stage RS-3 after Stage VIT-3 Viterbi redecoding with the help of one known symbol every four; finally, the lowest redundancy words (one of every two) will be decoded in Stage RS-4 after Stage VIT-4 Viterbi redecoding using one known symbol every two. A scheme of this sort was proposed in [2] for the Galileo HGA convolutional code and other codes.

The overall SER for this scheme can be calculated as in the two-stage case by assuming that the SER for the codewords that are supposed to decode during a particular stage also contributes to the SER for all weaker codewords. The resulting formula for the overall SER is of the form $SER = SER_a(1) + 7/8\ SER_b(2) + 3/4\ SER_c(3) + 1/2\ SER_d(4)$, where the indices a, b, c, d, refer to the strongest, next strongest, third strongest, and weakest codewords and (n) refers to decoding during Stage RS-$n, n = 1, 2, 3, 4$.

To achieve an overall SER of 2×10^{-7}, it is necessary that all four terms in the expression for SER be driven to approximately 10^{-7} or lower. At a convolutional code operating point of 0.10 dB, the equivalent operating points for Stages VIT-2 and VIT-3 are 0.33 dB and 0.72 dB, respectively, from the curves for phase ± 4 of 8 and phase ± 2 of 4 in Fig. 3. An appropriate set of E values to drive $SER_a(1)$, $SER_b(2)$, and $SER_c(3)$ each below 1×10^{-7} is $E = 36$, 24, and 12. As in the earlier discussion, the SER-equivalent output of Stage VIT-4 is beyond the scale of Fig. 2, but separate calcuations show that $E = 4$ is sufficient to bring $SER_d(4)$ under 1×10^{-7} and, together with the three aforementioned values, to achieve an overall SER of about 2×10^{-7}. The corresponding redundancy profile (72, 8, 24, 8, 48, 8, 24, 8) costs 0.45 dB, and so the concatenated code operating point is $E_b/N_o = 0.10$ dB + 0.45 dB = 0.55 dB.

If the convolutional code operating point is lowered to 0.00 dB, the SER for phase ± 1 of 2 is no longer small enough to permit $E = 4$ for the lowest redundancy codewords, but $E = 5$ is sufficient. From Fig. 3 the intermediate SER-equivalent convolutional code operating points are 0.27 dB and 0.66 dB. An appropriate set of E values to achieve an overall SER of about 2×10^{-7} is $E = 43$, 27, 13, and 5. The corresponding redundancy profile (86, 10, 26, 10, 54, 10, 26, 10) costs 0.52 dB, and the concatenated code operating point is E_b/N_o = 0.00 dB + 0.52 dB = 0.52 dB. This represents a slight improvement over the 0.10 dB convolutional code operating point, and there may be a slightly better operating point in the unsimulated region below 0.00 dB.

IV. Analysis Based on Depth-8 Interleaving and No Erasure Declarations

The analysis now continues for the actual Galileo conditions with depth-8 interleaving. As with the infinite interleaving analysis presented earlier, this section begins with a baseline analysis of one-stage decoding (no Viterbi redecoding) and proceeds to evaluate the gains possible from using a second stage. Results are not presented for more than two decoding stages.

A. One-Stage Decoding

For one-stage decoding with depth-8 interleaving, the performance of the Viterbi decoder is the same as before. The performance of the Reed-Solomon decoder must be evaluated by directly measuring the probability of decoder failure on simulated interleaved output from the Viterbi decoder.

1. Stage VIT-1. The output of Stage VIT-1 is characterized by the performance curves given in Fig. 1.

2. Stage RS-1. Figure 4 shows the performance of the Reed-Solomon decoder under an assumption of depth-8 interleaving and no erasure declarations, with varying amounts of error correction $E = 4$, 10, 16, 32, and 34. The depth-8 interleaving curves are plotted against a backdrop "grid" of infinite interleaving curves from Fig. 2.

In Fig. 4, the performance curves for depth-8 interleaving are shifted slightly to the right with respect to the corresponding infinite interleaving curves, showing a slight degradation due to nonideal interleaving of less than 0.05 dB at SER values around 1×10^{-5} or so. Equivalently, this degradation may be characterized as an approximate reduction in the error correction capacity of the Reed-Solomon code. For example, the depth-8 curve for $E = 16$ starts out coincident with the $E = 16$ infinite interleaving curve for high SER, but gradually drifts to cross the $E = 15$ infinite interleaving curve at an SER of $\sim 1 \times 10^{-3}$, and appears headed to cross the $E = 14$ infinite interleaving curve at an SER of $\sim 1 \times 10^{-7}$. Thus, depth-8 interleaving effectively decreases the error correction capacity of the (255,223) Reed-Solomon code by approximately two errors at an SER of $\sim 1 \times 10^{-7}$.

For the curves in Fig. 4, as well as succeeding figures below, the results are only statistically meaningful down to about 1×10^{-5} or 1×10^{-6} SER for the 2-Gbit datasets from Stage VIT-1. However, due to the almost parallel behavior of the depth-8 interleaving curves relative to the family of infinite interleaving curves, the depth-8 curves can be extended to 1×10^{-7} SER with high confidence by extrapolating along the appropriate infinite interleaving curve(s).

3. Performance of One-Stage Decoding. For the baseline coding system using the (255,223) Reed-Solomon code, E is 16 and a one-stage decoding system with depth-8 interleaving can achieve an SER of 2×10^{-7} (i.e., a BER of $\sim 1 \times 10^{-7}$) at convolutional code E_b/N_o = 0.59 dB. This calculation assumes an extrapolation along the $E = 14$ infinite interleaving curve as discussed above. The corresponding concatenated code signal-to-noise ratio is E_b/N_o = 1.17 dB, accounting for the 0.58 dB redundancy of the (255,223) Reed-Solomon code.

B. Two-Stage Decoding

For two-stage decoding with depth-8 interleaving, the performance of the Viterbi decoder in both stages is the same as that for the infinite interleaving analysis. The Reed-Solomon decoding performance is the same as that for one-stage decoding with depth-8 interleaving, to a degree of approximation noted below.

1. Stage VIT-1. The output of Stage VIT-1 is characterized by the performance curves given in Fig. 1.

2. Stage RS-1. The output of Stage RS-1 is characterized by the performance curves given in Fig. 4.

3. Stage VIT-2. The output of Stage VIT-2 is characterized by the same performance curves as given in Fig. 3.

4. Stage RS-2. The output of Stage RS-2 can be conservatively approximated by the performance curves given in Fig. 4, after using the SER-equivalent Stage VIT-1 operating points to characterize the output of Stage VIT-2. For the case of noninfinite interleaving, this equivalence

is not exact, because the output of Stage RS-2 depends not only on the input SER but also on the statistics of the error bursts that contribute to the average SER. The typical error bursts output from Stage VIT-2 were demonstrated to be much shorter than those output from Stage VIT-1 at the same value of SER. Shorter bursts are less likely to contribute two or more symbol errors to a given interleaved codeword, and hence the performance of Stage RS-2 should be slightly better (i.e., closer to infinite interleaving performance) than that of Stage RS-1 at the same average SER.

5. Performance of Two-Stage Decoding. If one of every four codewords is strong, extrapolation of the curves in Fig. 4 shows that a correction capability approximately halfway between $E = 32$ and $E = 34$ for the strong words is sufficient to drive $SER_s(1)$ to about 1×10^{-7} at a convolutional code operating point of 0.20 dB. According to Figs. 3 and 4, $E = 10$ is sufficient to push $SER_w(2)$ to 1×10^{-7} for the weak codewords at phases ± 1 and $E = 11$ can achieve $SER_w(2) = 2 \times 10^{-7}$ for the weak codewords at phase ± 2. The overall SER for this scheme is approximately $1 \times 10^{-7} + (1/2)1 \times 10^{-7} + (1/4)2 \times 10^{-7} = 2 \times 10^{-7}$. The redundancy profile is (66,20,22,20) and costs 0.58 dB of overhead. The resulting concatenated code operating point is 0.20 dB + 0.58 dB = 0.78 dB.

V. Analysis Based on Depth-8 Interleaving and a Simple Erasure Declaration Rule

This section analyzes the gains possible from applying a combination of Viterbi redecoding and Reed-Solomon redecoding using erasure declarations to codewords interleaved to depth 8. The erasure declaration rule is a simple strategy modeled after one of the rules given by Belongie and Arnold[1]; it is effective but not optimized. A symbol is always erased (a "double-sided" erasure) if it falls between two known erroneous symbols no more than eight symbols apart from each other. If this does not suffice to make the codeword decodable, a symbol is also erased (a "single-sided" erasure) if it is immediately adjacent to one known erroneous symbol. Single-sided erasures are extended one symbol farther (into the next adjacent codeword) every time another codeword successfully decodes. The erasure declaration process is repeated until no additional codewords are decodable within the block of eight interleaved words.

[1] M. Belongie and S. Arnold, "Error Forecasting Performance (Preliminary)," JPL Interoffice Memorandum 331-92.2-142 (internal document), Jet Propulsion Laboratory, Pasadena, California, January 4, 1993.

As with the analyses presented earlier, this section begins with a baseline analysis of one-stage decoding (no Viterbi redecoding) and proceeds to evaluate the gains possible from using a second stage. Results are not presented for more than two decoding stages.

A. One-Stage Decoding

One-stage decoding consists of one pass through the Viterbi decoder followed by one or more passes through the Reed-Solomon decoder to allow the various sequences of erasure declarations described above to be tried. The performance of the Viterbi decoder is the same as before. The performance of the Reed-Solomon decoder must be evaluated by directly measuring the ultimate probability of decoder failure after applying the rule for erasure declarations, using simulated interleaved output from the Viterbi decoder.

1. Stage VIT-1 The output of Stage VIT-1 is characterized by the performance curves given in Fig. 1.

2. Stage RS-1. Figure 5 shows the performance of the Reed-Solomon decoder under an assumption of depth-8 interleaving and the simple erasure declaration algorithm described above. For the solid curves in Fig. 5, it is assumed that all eight codewords in a frame have the same error correction capacity E, and performance is plotted for $E = 10, 16$. The $E = 16$ performance curve in Fig. 5 shows a performance gain of about 0.10 dB with respect to the corresponding curve in Fig. 4 and a gain of about 0.05 dB with respect to the curve in Fig. 2. The erasure declaration algorithm is powerful enough not only to overcome the degradation due to finite interleaving but also to outperform decoding with ideal interleaving but no erasure declarations. At low SER, the $E = 16$ depth-8 interleaving curve with erasure declarations is almost coincident with the $E = 18$ infinite interleaving curve with no erasure declarations. Thus, depth-8 interleaving combined with Reed-Solomon redecoding using a simple erasure declaration rule effectively increases the error correction capacity of the (255,223) Reed-Solomon code at an SER of $\sim 10^{-7}$ by approximately two errors relative to infinite interleaving with no erasure declarations, or by approximately four errors relative to depth-8 interleaving with no erasure declarations.

The dot-dash curve in Fig. 5 shows the performance of the Reed-Solomon decoder using the same erasure declaration algorithm but allowing the codewords to have the variable redundancy profile (44,28,28,28) discussed by Belongie and Arnold.[2] This profile has the same average

[2] Ibid.

redundancy as the constant redundancy code with $E = 16$ and therefore the same redundancy overhead. The SER curve for this variable redundancy scheme lies just under the infinite interleaving curve for $E = 20$. The effective average error correction capability is improved over the constant redundancy scheme by approximately two additional errors, or equivalently by about 0.05 dB at an SER of $\sim 10^{-5}$.

3. Performance of One-Stage Decoding. Extrapolation of the $E = 16$ curve in Fig. 5 yields a required operating point of 0.46 dB to achieve an $SER = 2 \times 10^{-7}$. This corresponds to a concatenated code operating point of 0.46 dB + 0.58 dB = 1.04 dB. The $E = 10$ curve requires 0.67 dB plus 0.35 dB of overhead or 1.02 dB for the concatenated code. The variable redundancy (44,28,28,28) code requires 0.40 dB + 0.58 dB = 0.98 dB.

B. Two-Stage Decoding

For two-stage decoding with depth-8 interleaving and Reed-Solomon erasure declarations, the performance of the Viterbi decoder in both stages is the same as before. The evaluation of Reed-Solomon decoding performance must be modified somewhat to account for the fact that erasure declarations in strong codewords in Stage RS-1 must typically be forecast only from other strong codewords (because weak codewords are not highly likely to decode during Stage RS-1), and erasures in weak codewords in Stage RS-2 must be forecast from other weak codewords (because there are no symbol errors remaining in strong codewords in Stage RS-2 from which to forecast neighboring symbol errors).

1. Stage VIT-1. The output of Stage VIT-1 is characterized by the performance curves given in Fig. 1.

2. Stage RS-1. The output of Stage RS-1 is characterized by the performance curves given in Fig. 6. These curves show the SER that is obtained when strong and weak codewords are intermixed within a frame, but one attempts to decode only the strong words (and declares only those erasures based on decodable strong codewords). The plotted SER is the average SER obtained by accumulating symbol errors from all the codewords in frames that contain at least one strong codeword error. The dashed curves in Fig. 6 show SER for a redundancy profile of one strong word every two codewords, and the thick solid curves show SER for one strong word every four codewords. The error correction values noted for these curves ($E_s = 26, 28$ for the first case and $E_s = 30, 32, 34$ for the second case) pertain to the strong codewords only.

3. Stage VIT-2. The output of Stage VIT-2 is characterized by the same performance curves given in Fig. 3.

4. Stage RS-2. At present no curves directly characterize the output of Stage RS-2. The SER from Stage RS-2 can certainly be upper bounded by the SER obtainable without erasure declarations, as shown in Fig. 4. The incremental effect of Reed-Solomon redecoding with erasure declarations can be estimated to some degree by comparing curves in Fig. 4 (without erasure declarations) with corresponding curves in Fig. 5 (with erasure declarations), but this is not an exact procedure because the effect of erasure declarations is different after Viterbi redecoding than after Viterbi decoding the first time. In particular, the shorter bursts from Stage VIT-2 should lead to less efficacious error forecasting than the longer bursts from the "equivalent" Stage VIT-1 operating point.

5. Performance of Two-Stage Decoding. Assume that one of every four codewords is strong, and the convolutional code operating point is 0.20 dB. Extrapolation of the curves in Fig. 6 shows that a correction capability of $E = 32$ for the strong words is sufficient to drive $SER_s(1)$ to about 1×10^{-7}. Similarly, according to Figs. 3 and 4, $E = 10$ is sufficient to push $SER_w(2)$ to the same level for the weak codewords at phases ± 1 even without any assistance from erasure declarations. For the overall SER to reach 2×10^{-7}, the remaining weak codewords at phase ± 2 must achieve $SER_w(2)$ of 2×10^{-7} or lower. From Figs. 3 and 4 this can be approximately accomplished without any erasure declarations by a code with $E = 11$. However, it is estimated that this required value of E can also be reduced to $E = 10$ by using erasure declarations from the codewords at phases ± 1 to help decode the words in between at phase ± 2. The redundancy profile for this case is (64,20,20,20) and it costs 0.56 dB of overhead. The resulting concatenated code operating point is 0.20 dB + 0.56 dB = 0.76 dB.

If the convolutional code operating point is moved to 0.10 dB, the corresponding required values of E for the strong and weak codewords are approximately $E = 38$ and $E = 11$, respectively, but there are insufficient data plotted in Figs. 4 and 6 to confirm this directly. The corresponding redundancy profile is (76,22,22,22) and the concatenated code operating point is 0.75 dB, which is almost identical to the previous value. Again this indicates an insensitivity of the optimum performance to small variations on the order of 0.1 dB in the convolutional code operating point. All other things being equal, there is a slight preference to operate at the higher convolutional code E_b/N_o in order to raise the channel symbol signal-to-noise ratio.

VI. Performance Comparisons

Tables 1 and 2 summarize the optimal or near-optimal (within a few hundredths of a dB) concatenated code decoding thresholds for achieving an overall SER of 2×10^{-7} (a BER of $\sim 1 \times 10^{-7}$) for the various decoding alternatives considered in the preceding sections.

Several conclusions can be drawn from these two tables. Based on the infinite interleaving analysis, a second stage of Viterbi decoding without any Reed-Solomon redecoding is worth about 0.38 dB at $SER = 2 \times 10^{-7}$. The same comparison for depth-8 interleaving yields essentially the same answer, 0.39 dB. Adding Reed-Solomon redecoding using erasure declarations increases this value only marginally to 0.41 dB for the case of depth-8 interleaving. This tiny incremental value of Reed-Solomon redecoding when used in conjunction with Viterbi redecoding is probably somewhat underestimated due to the conservative approximations made in the foregoing analysis, but the true incremental value is unlikely to be more than several hundredths of a dB. On the other hand, Reed-Solomon redecoding without Viterbi redecoding is worth about 0.19 dB for depth-8 interleaving.

Allowing four stages of Viterbi decoding (and four levels of Reed-Solomon redundancy) is worth about 0.58 dB for infinite interleaving and no erasure declarations. Data are not available to confirm the corresponding result for depth-8 interleaving, but a similar effect is expected.

Tables 1 and 2 also give a measure of the amount by which depth-8 interleaving is nonideal. Compared to infinite interleaving, depth-8 interleaving costs 0.06 or 0.07 dB for both one- and two-stage decoding without erasure declarations.

VII. Choosing a Coding System for Best Performance

If the objective for the Galileo LGA code is to maximize the allowable data rate, this is accomplished by lowering the concatenated code's decoding threshold as far as possible. Under a constraint of no more than two decoding stages and two levels of codeword redundancy, the best performance is obtained by using two-stage decoding with one strong codeword every four. The redundancy profile (64,20,20,20) achieves a near-optimum concatenated code decoding threshold of 0.76 dB while requiring the convolutional code to operate at 0.20 dB.

Based on the infinite interleaving analysis reported in Table 1, performance may be further improved by approximately 0.2 dB if four stages of Viterbi decoding and four levels of Reed-Solomon redundancy are permitted. Confirmation of this effect and specification of the optimum four-level redundancy profile for depth-8 interleaving will be the subject of a future article.

References

[1] K.-M. Cheung and S. J. Dolinar, Jr., "Performance of Galileo's Concatenated Codes with Nonideal Interleaving," *The Telecommunications and Data Acquisition Progress Report 42-95, vol. July–September 1988*, Jet Propulsion Laboratory, Pasadena, California, pp. 148–152, November 15, 1988.

[2] O. Collins and M. Hizlan, "Determinate-State Convolutional Codes," *The Telecommunications and Data Acquisition Progress Report 42-107, vol. July–September 1991*, Jet Propulsion Laboratory, Pasadena, California, pp. 36–56, November 15, 1991.

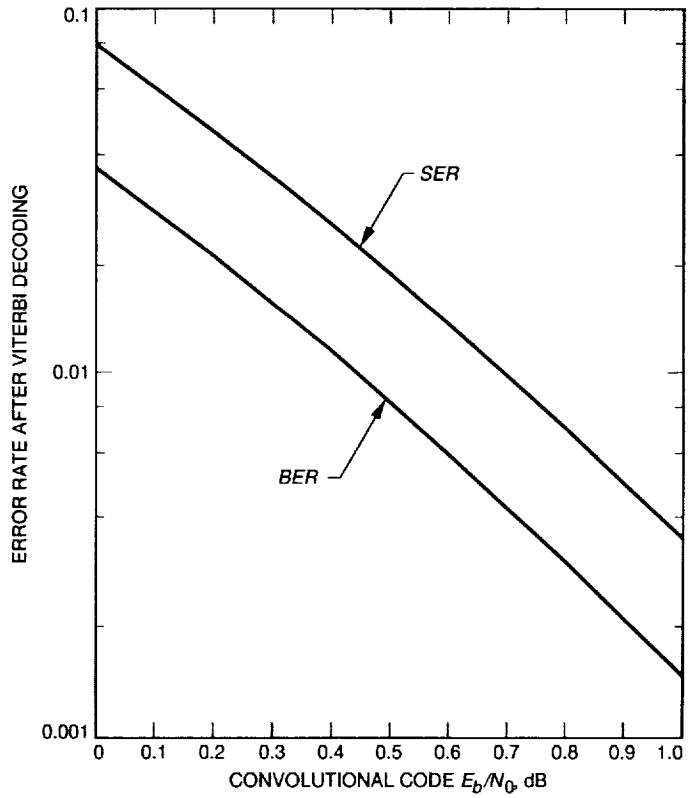


Fig. 1. Bit error rate and symbol error rate after Viterbi decoding of the (14, 1/4) Galileo LGA convolutional code.

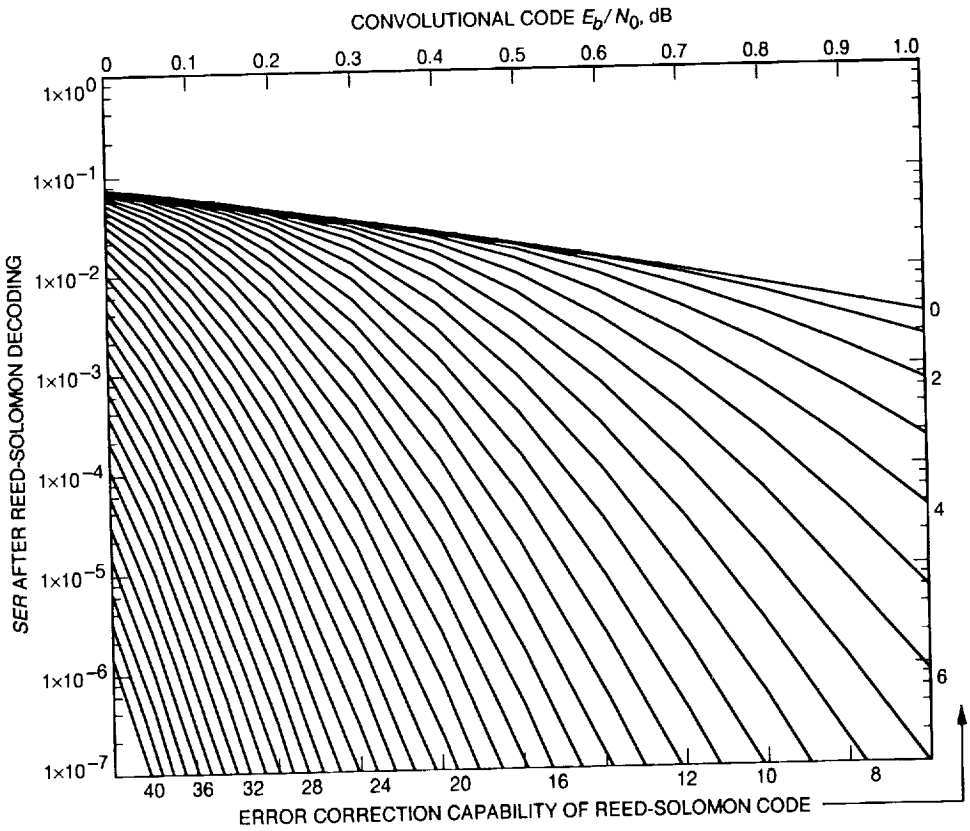


Fig. 2. Symbol error rate after Reed-Solomon decoding with infinite interleaving.

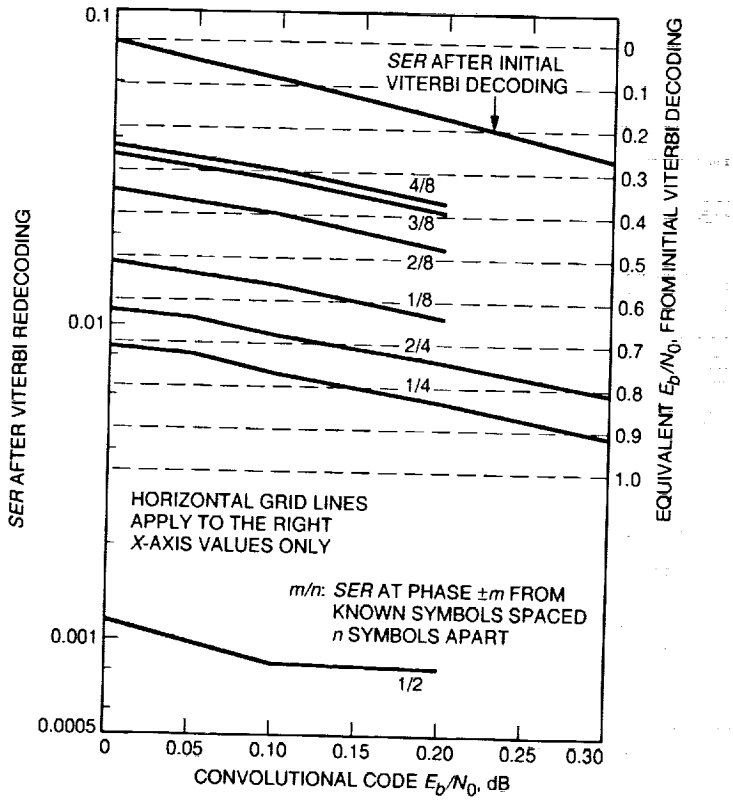


Fig. 3. Symbol error rate after Viterbi redecoding with various spacings of known symbols and relative phases of unknown symbols.

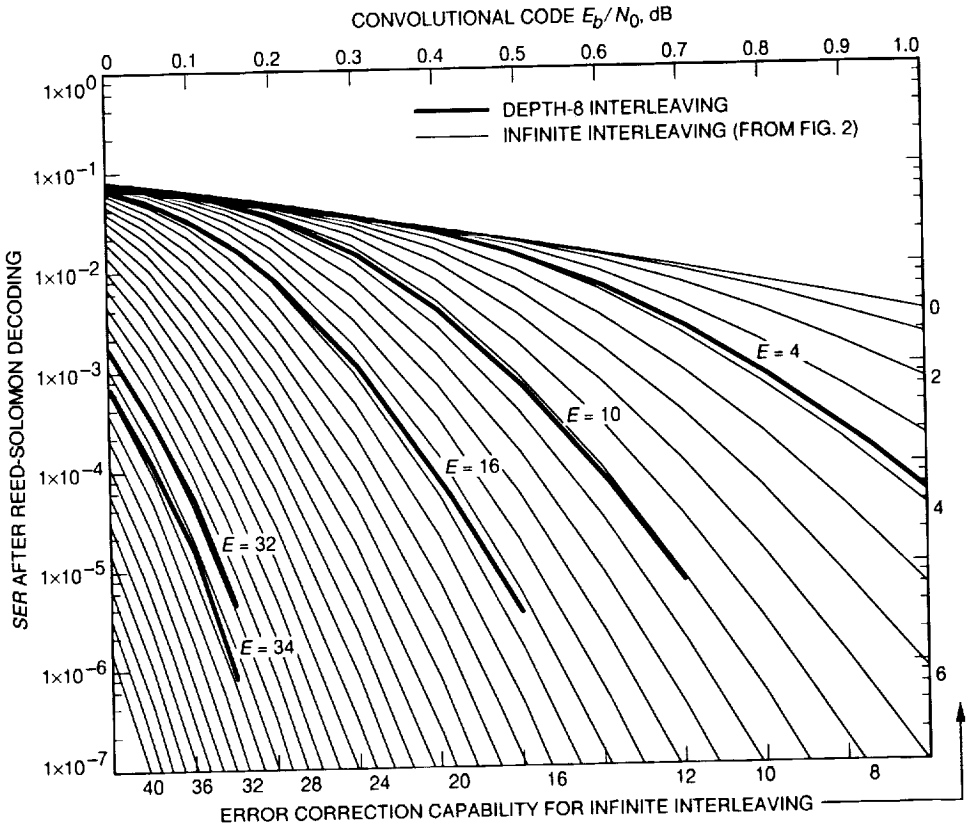


Fig. 4. Symbol error rate after Reed-Solomon decoding with depth-8 interleaving.

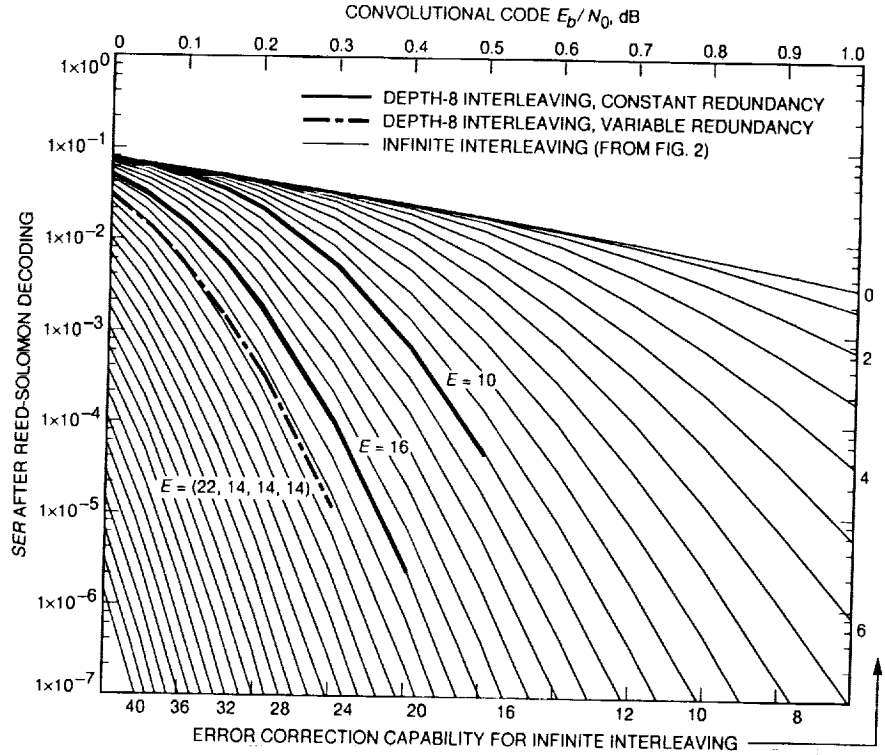


Fig. 5. Symbol error rate after Reed-Solomon decoding using erasure declarations with depth-8 interleaving.

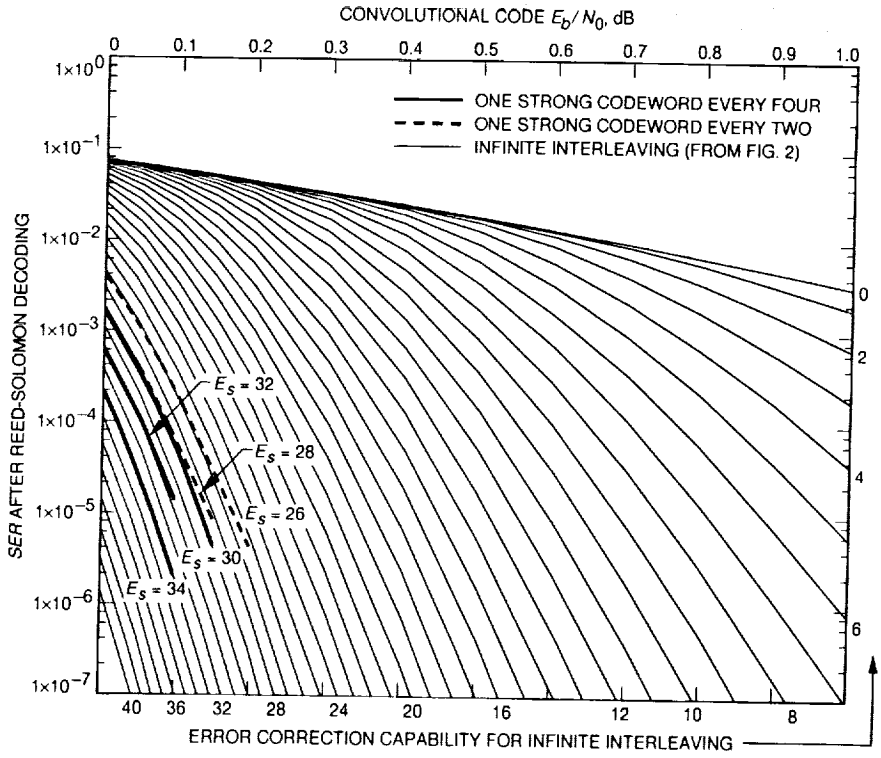


Fig. 6. Symbol error rate after Reed-Solomon decoding using erasure declarations among strong codewords only with depth-8 interleaving.

Table 1. Performance comparisons for infinite interleaving ($SER = 2 \times 10^{-7}$).

Decoding stages	1	2	4
Erasure declarations	No	No	No
Redundancy profile	(32)	(60, 18, 20, 18)	(86, 10, 26, 10, 54, 10, 26, 10)
Convolutional code E_b/N_0 (dB)	0.52	0.20	0.00
Concatenated code E_b/N_0 (dB)	1.10	0.72	0.52

Table 2. Performance comparisons for depth-8 interleaving ($SER = 2 \times 10^{-7}$).

Decoding stages	1	1	2	2
Erasure declarations	No	Yes	No	Yes
Redundancy profile	(32)	(44, 28, 28, 28)	(66, 20, 22, 20)	(64, 20, 20, 20)
Convolutional code E_b/N_0 (dB)	0.59	0.40	0.20	0.20
Concatenated code E_b/N_0 (dB)	1.17	0.98	0.78	0.76

Stability Measurements of the Radio Science System at the 34-m High-Efficiency Antennas

T. T. Pham, J. C. Breidenthal, and T. K. Peng
Telecommunications Systems Section

S. F. Abbate and S. T. Rockwell
TDA Mission Support and DSN Operations

From 1991 to 1993 the fractional frequency stability of the operational Radio Science System was measured at DSS's 15, 45, and 65. These stations are designed to have the most stable uplink and downlink equipment in the DSN. Some measurements were performed when the antenna was moving and the frequency was ramped. The stability, including contributions of all elements in the station except for the antenna and the hydrogen maser, was measured to be 0.3 to 1.3 × 10^{-15} when the frequency was fixed, and 0.6 to 6.0 × 10^{-15} when the frequency was ramped (sample interval, 1000 sec). Only one measurement out of fifteen exceeded specification. In all other cases, when previous measurements on the antenna and the hydrogen maser were added, a total system stability requirement of 5.0 × 10^{-15} was met. In addition, ambient temperature was found to cause phase variation in the measurements at a rate of 5.5 deg of phase per deg C.

I. Introduction

Good frequency stability is essential in the Radio Science System, since frequency is one of the primary observable parameters in radio science experiments. The required period of stability varies from a few hundred to several thousand seconds, depending on the nature of the experiment. For example, gravitational wave detection experiments usually require high two-way stability over a long round-trip light time between the Earth and a spacecraft. A 1000-sec duration is often used as a convenient checkpoint for the range of round-trip light time encountered in practice. On the other hand, for occultation studies, the time scale of interest is much shorter. All that is needed is stability over the time interval it takes a spacecraft to descend through the planetary atmosphere. For Mars Observer, this time scale is no more than 300 sec.

The focus of this investigation is the frequency stability performance of the open-loop Radio Science System at the 34-m high-efficiency (HEF) subnet. More interest is placed on this subnet, rather than the 70-m or 34-m standard (STD), because of the HEF's inherently greater stability. This greater stability results from a deliberate attempt during design to capitalize on the fact that phase scintillation due to the solar wind and Earth's ionosphere decreases with the inverse square of the carrier frequency. For this reason, the 34-m HEF subnet was equipped with X-band (8.4 GHz) uplink and downlink equipment of stability better than the anticipated phase scintillation. The

potential 13-fold improvement compared to S-band-only (2.3-GHz) stations, and a ninefold improvement compared to S-band uplink/X-band downlink stations, makes the 34-m HEF stations the best available for detecting gravitational radiation by spacecraft tracking.

Specifically, the objectives of this study and their significance are as follows:

(1) *Stability performance in both one-way and two-way configuration.* One-way stability refers to that of the downlink equipment only, whereas two-way includes both uplink and downlink equipment. Knowledge of system performance in both one-way and two-way configurations is important due to the difference between gravitational wave and planetary atmosphere studies. The availability of both sets of measurements also helps to reveal the stability of the uplink path, which there is currently no way to directly measure.

(2) *Stability performance under both static and dynamic frequency conditions.* Static and dynamic conditions refer to cases where the test signal frequency is fixed and ramped, respectively. Static conditions are used frequently due to their compatibility with a simple test setup. However, measurement under dynamic conditions offers more insight on how the system performs in actual experiments. The signals received from spacecraft always have an associated Doppler frequency due to the relative velocity between the spacecraft and the tracking ground antenna. As the signal traverses the bandwidth that the Radio Science System operates prior to the point of Doppler compensation, it is distorted, however slightly, by the inherently dispersive electronic components in the system. At the point of Doppler compensation, the signal is again exposed to another form of degradation, caused by the potential phase instability of a programmable local oscillator. The objective of the dynamic test is to validate that such a distortion is small, as intended by the design.

(3) *Performance improvement due to the use of a low-thermal-coefficient fiber-optic link for antenna references.* With recent improvements, the 100-MHz references used in the exciter and the downconverter in the antenna area are transported from the Signal Processing Center (SPC) via a fiber-optic link. In the past, these references were transported via microwave coaxial cables and phase-stabilized cables. It should be noted that system measurements with either fiber-optic or phase-stabilized references do not include the contribution of the reference link because of the cancellation that results from using a common cable for both the uplink and downlink. These system tests are therefore limited to validating the function of the reference link. In order to account for the contribution of the hydrogen maser clock and the fiber-optic link in total system performance, the results from these tests are combined with those from stand-alone testing on the link itself. Section VI provides more details on this process.

(4) *Impact of the environment and operation on stability performance.* Particularly, the effect of antenna motion, radio frequency interference (RFI), diurnal temperature variation, and wind conditions are examined. Due to a limitation in the test configuration, to be discussed in Section II, the effect of antenna motion is only partially observed.

The foregoing objectives are intended to make a full characterization of the system, so that scientists will know the noise characteristics of the data they receive. It is worth noting that there were other similar studies on stability performance of the DSN equipment conducted and reported in earlier issues of *The Telecommunications and Data Acquisition Progress Report.* For example, Otoshi and Franco have published a series of extensive studies regarding the stability of a research and development (R&D) 34-m antenna at Goldstone (DSS 13) [1]. In 1989, in an effort to prepare for Voyager's Neptune encounter, Ham et al. conducted a series of stability measurements of the Radio Science System at Canberra's 70-m antenna (DSS 43) [2]. The present study shows a lot of commonality with those earlier efforts in terms of test configuration and instrumentation. However, there are also differences. Compared to [1], the present measurements reflect the stability at an operational 34-m HEF station, rather than at an R&D site. The equipment under test is the X-band uplink and radio science X-band downlink system. In contrast to [2], where stability is characterized at S-band with an S-band uplink system at the 70-m antenna, this study presents measurements at X-band. Also, the link between the SPC and the antenna is a fiber-optic link, instead of another transport medium. Lastly, the presented measurements include a more realistic performance with uplink Doppler simulation.

In the following sections, more details of this study are given. Section II provides a description of the test configuration. Section III presents the data acquisition and processing. Measurements and performance analysis are discussed in Sections IV and V. A conclusion is drawn in Section VI regarding whether the DSN Radio Science System satisfies its stability requirements. The summary and conclusions are given in Section VII.

II. Test Configuration

Most of the tests were conducted at the 34-m HEF antenna at Goldstone (DSS 15). This particular deep space station was chosen as the site of study because of its proximity to JPL. The equipment design is identical at DSS 45 (Canberra) and DSS 65 (Madrid), and less extensive tests at these sites have verified the conclusions drawn from the DSS-15 tests. The test configuration is presented in Fig. 1.

A. General Configuration

In one-way measurements, a test signal of 8.4 GHz was generated from a 100-MHz reference in the antenna by a highly stable ×84 signal generator, which is essentially an ovenized ×84 frequency multiplier. The 8.4-GHz signal was injected into the downlink microwave system prior to the X-band maser at a point known as the "ranging coupler." In two-way measurements, an X-band uplink signal was generated by the exciter and was converted to a downlink frequency by the test translator.

Once injected into the downlink system, the test signal was amplified by the X-band maser and later converted to 300 MHz intermediate frequency (IF) by the very long baseline interferometry/radio science (VLBI/RS) downconverter. The 300-MHz IF signal was transported from the antenna to the signal-processing center via a hard-line cable. The Radio Science IF-video downconverter (RIV) further converted the signal to baseband frequency. This downconversion process was achieved over several stages wherein the noise bandwidth was gradually reduced by various filters. In the first stage, any simulated Doppler, if present, was taken out by a programmable local oscillator. A small frequency offset was also introduced into this oscillator to place the carrier at a chosen position in the bandwidth of the equipment about to be described. In the final stage, a set of filters allowed six different bandwidth selections, ranging from 100 Hz to 20 kHz. At this point, the signal could be either channeled to the prototype stability analyzer for real-time stability measurement or to the Deep Space Communications Complex (DSCC) Spectrum Processing Subsystem (DSP) to be digitized and recorded on magnetic tapes for later processing. All stages of frequency conversion, both on the uplink and downlink paths, are depicted in Fig. 2.

In order to minimize thermal noise effects, test signals were set at a high power level but without saturating the low-noise amplifier. Typically, a signal-to-noise spectral density ratio of 60 dB-Hz or greater was used. A bandwidth of 100 Hz or 400 Hz in the final filtering was chosen to maximize the signal-to-noise ratio (SNR). Another benefit associated with the small bandwidth is the correspondingly small data volume, which helps to minimize the data extraction effort.

B. Antenna Frequency References

Of particular interest is the fact that a fiber-optic link is used to transport the 100-MHz references to the antenna area. These references are required by the exciter, test translator, VLBI/RS downconverter, and ×84 signal generator. This fiber-optic link includes a transmitter, receiver, low thermal coefficient-of-delay (LTCD) fiber-optic cable, and a 1-to-4 distributor. A prototype unit was made available during the engineering testing period by the Frequency and Timing Subsystem (FTS) Group, and operational units were recently deployed at all 34-m HEF stations as the official configuration.

These references have a significant impact on the overall stability of the system. Large frequency multiplication applied to these references (e.g.,×65 in the exciter and ×81 in the VLBI/RS downconverter, see Fig. 2) tends to amplify small frequency and phase fluctuations caused by the link between SPC and the antenna. Care must be taken to ensure that the link itself does not significantly degrade the stability of FTS references coming from the hydrogen maser. This concern is particularly appropriate considering the long distance that the link traverses—300–550 m, depending on the station—and the relatively large daily thermal fluctuation to which it is subjected.

The advantage in using fiber optics is twofold: a better performance in both long-term frequency stability and short-term spectral purity. In the earlier DSN configuration, a combination of coaxial hard lines and phase-stabilized cables were used for reference transmission. Although each had its own merit—high stability for the phase-stabilized cables and pure spectrum for hard lines—none offered the excellent performance in both aspects that fiber optics does. In the case of phase stabilization, the use of two frequency multipliers (i.e., ×4 and ×5) to convert a 5-MHz FTS reference to the desired 100 MHz resulted in an increase in phase noise, which was unacceptable for atmospheric and ring experiments. In the case of hard lines, stability performance was compromised by the rather large sensitivity to temperature variation that these cables have. Typically, the thermal coefficient of delay (TCD) of these microwave hard-line coaxial cables is about 15–18 ppm/deg C (parts per million per degree Celsius) for a temperature range of 18–33 deg C [3]. In contrast, the fiber-optic link directly transmits the 100-MHz reference without any up- or downconversion, and its TCD varies from 0.33 ppm/deg C to 0.83 ppm/deg C over the same temperature range [3].

Unfortunately, improvement with the use of fiber-optic link is not expected to be observable in the stability measurements. As indicated in Fig. 1, both uplink and downlink references are transported via a common link. Any phase variation caused by the fiber-optic link, or by the frequency and timing source, i.e., the hydrogen maser, would appear in both the generated test signal and the local oscillator of downconversion. Since there was essentially zero time delay between the uplink and downlink, the majority, 94 percent to be exact, of phase disturbance in the references was canceled. In other words, only 6 percent of fiber-optic stability appeared in the measurement. For this reason, the stand-alone measurement on the fiber-optic link done by Calhoun and Law[1] is used to augment the presented measurements in order to determine the total system performance.

Note that such cancellation does not apply to the measurements with non-zero round-trip light time, as in a real spacecraft tracking session, since the correlation in the phase of the incoming signal and the first local oscillator is then removed. In this case, the total stability is the root-sum square of contributions of the link and the rest of the system. As indicated in Calhoun and Law,[2] the stability contribution of the fiber-optic link was measured to be at least one-and-a-half orders of magnitude smaller than the rest of the system.

C. Antenna RF Cables

In the current system design, the RF cables that run across the antenna's moving parts are ring-wrapped to accommodate the antenna motion. As the antenna rotates to stay pointed to the target, physical stress is imposed on these cables. The stress can manifest itself in the form of abrupt phase discontinuity or a gradual phase drift. To study this impact, measurements were made with both a stationary and moving antenna.

D. IF Cables

The 300-MHz IF cables between the antenna and SPC, which deliver the downconverted signal to the RIV, are also a point of concern in terms of instability. A large portion of the cables not in the SPC or the antenna housing is exposed to a thermally unregulated environment. Temperature fluctuations, particularly around the period of sunrise and sunset, may have a severe impact on stability performance. Due to the relatively long time scale involved, the expected impact, if any, is likely in the form of gradual phase drift. Particular attention was paid in the data analysis to determine if there is any correlation between the variation in the temperature and the signal phase.

E. Limitations

There are inherent limitations associated with the test setup. The measurements did not account for frequency instability of the antenna or the standard hydrogen maser. The contribution of the antenna was excluded because there were no means to radiate a test signal onto the main reflector without a serious multipath interference problem. In addition, since both the test signal and the system under study were synchronized to the same frequency standard and the round-trip light time was zero, the coherence between them resulted in a near-perfect cancellation, making any frequency instability in the hydrogen maser unobservable.

Nevertheless, data on stand-alone tests on these two subsystems are available. For instance, Calhoun and Law[3] measured the two-way stability of the fiber-optic link and the hydrogen maser. Also, Otoshi recently reported a measurement of the antenna stability at the 34-m R&D station (DSS 13) in [4]. These two sets of data are used to augment measurements, thereby estimating total system performance.

III. Data Acquisition and Processing

Three different methods are currently available for data acquisition and processing, as depicted in Fig. 3. They will be referred to as complex demodulation, digital phase-lock loop (DPLL), and time interval. Both the complex demodulation and the DPLL methods rely on digitized voltage samples from the DSCC spectrum processor (DSP) subsystem. The time interval approach, however, uses the analog output from the receiver subsystem as input. Details on the data processing, as well as the strength and weakness of each method, are described below.

A. Complex Demodulation Method

The signal frequency is first estimated by applying a Fast Fourier Transform (FFT) operation on the incoming digitized samples. This estimated frequency is used as reference for the local oscillators that downconvert the signal

[1] M. D. Calhoun and J. C. Law, Communications and Systems Research Section, unpublished results on the stability of the fiber-optic distribution assembly (internal document), Jet Propulsion Laboratory, Pasadena, California.

[2] Ibid.

[3] M. D. Calhoun and J. C. Law, Communications and Systems Research Section, unpublished results on the stability of the fiber-optic distribution assembly and of the hydrogen maser/coherent reference generator (internal document), Jet Propulsion Laboratory, Pasadena, California.

to dc frequency. The downconversion is carried out in the complex domain, with both in-phase and quadrature arms. The outputs of the mixers are averaged over 1-sec intervals in order to enhance the SNR and to eliminate the double-frequency component. By taking the arc tangent of the ratio of the two averages, one can obtain the phase of the signal.

After the phase extraction process is completed, the phase samples are used as input into an algorithm that computes the Allan deviation in the phase domain. The computed Allan deviations serve as the measurement of system stability.

The advantage of this approach lies in its simplicity. Since the phase is estimated based on a one-second segment of the digitized samples and there is no feedback involved, the algorithm can tolerate a large variation in the input phase, up to 360 deg/sec. In contrast, the DPLL method that is described next requires a relatively small phase change among the input data. In the event of a large phase jump, either because of high noise level or missing data records, the loop will temporarily lose lock and require some time to re-establish the steady lock condition.

On the disadvantage side, the complex demodulation method requires the estimated frequency to be within 1 Hz of the true frequency. This is necessary to resolve the ambiguity in the phase determination as caused by the periodic tangent function.

In the process of calibration to understand how much noise the phase extraction process contributes, it was found that the algorithm, despite its simplicity, causes some degradation, particularly when the input signal frequency is fractional. The degradation comes from the fact that the 1-second averaging does not necessarily coincide with an integer number of cycles of signal, resulting in the twice-frequency component not being fully eliminated. The problem becomes more pronounced under low sampling-rate and high SNR conditions, which are often present in the testing.

B. Digital Phase-Lock Loop Method

The DPLL algorithm in use is based on [5], with major components presented in Fig. 3(b). The voltage input is first normalized to unity in amplitude by the automatic gain control (AGC) loop. The frequency deviation is extracted from the output of the loop filter at the rate of 1 sample per second. The frequency deviation can then be fed directly to an algorithm that computes the Allan deviation based on the frequency domain. Alternatively, the deviation can be converted from frequency to phase and input to a phase-domain Allan deviation algorithm as in the first method.

Note that the proper choice of PLL bandwidth is important. Selection of an improper loop bandwidth can cause a distortion in the test result. For example, a very small bandwidth, B_ℓ, will smooth out the detected phase variation, resulting in lower Allan deviations at those integration times less than $1/B_\ell$. A larger bandwidth, on the other hand, will allow more thermal noise leakage, creating an impression of high instability in the region where thermal noise dominates, i.e., up to roughly 100 sec. For data processing, a PLL bandwidth of 1 Hz is typically chosen for phase detection since the region of interest for stability performance lies between 1 second and several thousands of a second.

C. Time-Interval Method

With the first two approaches, data quality can only be evaluated at the end of data collection. The ability to monitor the performance of the system in real time is much desired. Such a capability is currently provided by a prototype model of the radio science stability analyzer, which is currently available at Goldstone.

The test setup for this method is shown in Fig. 3(c). The near-baseband analog signal taken out from the receiver is further downconverted by the analyzer to 1 Hz with a local synthesizer. The 1-Hz sinusoidal signal then gets low-pass filtered and transformed to a 1-pps (pulse-per-second) signal by the zero crossing detector. The time of the raising transition of the 1-Hz pulse is measured relative to the 10-pps "picket-fence" reference. Time interval measurements are converted to relative phases, from which the Allan deviations are calculated. A more detailed description on the signal processing aspect of this method is provided in [6].

Some consistency among all three methods is established in Table 1. Three sets of data collected at Goldstone DSS 15 on DOY 334/1990, DOY 177/1991, and DOY 027/1993 are shown. All tests were conducted under frequency-static conditions, i.e., no ramping in the uplink or downlink. Among the three methods, the average variation was about 50 percent, although there are cases where the difference exceeded the 100-percent level. This is not necessarily bad, considering there are measurements where a large fluctuation in the Allan deviation was observed at different integration times. These troublesome measurements will be pointed out later in the data analysis section. Note also that no single method consistently

produced higher or lower Allan deviations than the other two.

This consistency ensures that the data can be processed by any of the three methods discussed above. As it turns out, most of the test results were processed with the complex demodulation and DPLL methods, although only the data from the DPLL processing are presented here. Not many results associated with the time-interval method are available. The prototype stability analyzer, despite its great benefit as real-time test equipment, is a one-of-a-kind machine and is not available at the oversea stations. In addition, the analyzer does not handle well the testing of frequency dynamic conditions. This limitation comes from the fact that the input signal to the analyzer has a small frequency drift, typically around 30 μHz/sec. This drift is the result of imperfect compensation between the uplink and downlink ramps, due to finite resolution in the frequency rate register. The time-interval samples at the beginning and at the end of the test no longer span the same period (e.g., 1 sec), thereby invalidating the calculation of the Allan deviation. In contrast, the complex demodulation and DPLL methods can remove the residual frequency drift by tuning the local oscillators accordingly.

IV. Acquired Data

The presented measurements span two-and-a-half years. The study started with the preparation to support the then-anticipated X-band Gravitational Wave Experiment with Galileo, scheduled for April 1991. It was finalized with system performance testing for the support of the recent Mars Observer Coincidence (along with Galileo and Ulysses) Gravitational Wave Experiment, in March 1993. Table 2 presents both the test configuration of each measurement and the associated stability at representative integration periods. The results are grouped according to the objective of the tests, such as two-way static, one-way static, two-way dynamic, etc. Multiple measurements within each group, when available, are shown in order to demonstrate the repeatability, or lack thereof, among test results. Measurements with known hardware problems are not included. A more detailed analysis is provided in the next section.

V. Performance Analysis

Analysis is given in the order presented in Table 2.

A. Performance Under Frequency-Static Conditions

1. Two-way link. Figure 4 presents the stability measurements, in terms of Allan deviation, in two-way frequency-static configuration at all three complexes. The respective phase profiles are shown in Fig. 5. The measurables include the stability of the exciter, test translator, microwave plumbing, X-band maser, VLBI/RS downconverter, IF/video downconverter, narrow-band occultation converter (NBOC), spectrum-processor assembly (SPA), and system cables that run between subsystems. Note that the 20-kW transmitter (TXR) was not included in the test; however, as the data from DOY 032, 1993 at DSS 15 indicates, the TXR contribution is very small. Therefore, these measurements can be used to represent all equipment in the Radio Science System, excluding the FTS and the Antenna Mechanical Subsystem (ANT). For reference purpose, the theoretical thermal contribution based on a 1-Hz PLL bandwidth and 60 dB-Hz signal level, which is the nominal level of testing, is also shown.

Overall, the measurements appeared to be consistent across different complexes. Several comments can be made:

(a) At short integration times, i.e., less than 100 sec, the drop-off in the Allan deviation somewhat followed the $1/T$ relationship, where T is the integration time. Both DOY 035 and 067 had a slower drop-off rate between 10 and 100 sec, suggesting that there was an additional degradation factor involved. In fact, as indicated in Fig. 5, there were several short-term phase glitches associated with these two data sets.

(b) DOY 105 data exhibited regular oscillation at short integration time scales. In the process of calibrating the algorithm, it was found that such oscillation can be caused by short-term phase bursts; however, the phase profile of this data set, as seen in Fig. 5, did not reveal any outstanding short bursts. Whether the cyclic nature of the phase toward the end of the test is responsible for the oscillation in the Allan deviation requires further investigation.

(c) At long integration times, i.e., 1000 sec or greater, three sets of measurement on DOY 035, DOY 067, and DOY 105 showed a consistently good stability performance at the level of 8.0×10^{-16}. This indicates that there was little noise contribution from electronic components in the system. The data from DOY 341 and DOY 027 showed a higher level of instability at these long periods. The change in the phase slope on DOY 341, which corresponded to a 300-μHz frequency jump and occurred three hours into the test, was partly responsible for the drift. The phase profile of DOY 027 also suggests that there was a frequency jump on the order of 70 μHz during the last hour of the test.

2. Stability of the transmitter. Figure 6 presents the measurements with and without the transmitter on the same day (DOY 032, 1993 at DSS 15).[4] The two plots lie practically on top of each other, indicating that the TXR contribution is insignificant. At 1000-sec integration, the measurement without the TXR was actually lower than that with the TXR; however, due to the small set of data samples, the uncertainty associated with this point was rather large.

3. One-way stability. One-way stability includes all contributions from downlink elements, ranging from the microwave to the DSP Subsystem. The result also includes the stability of the test signal source, namely the ×84 signal generator; however, its contribution was so small, e.g., 1.8×10^{-14} (1 sec), 3.0×10^{-15} (10 sec), 4.9×10^{-16} (100 sec), 1.6×10^{-16} (1000 sec), that it could be practically ignored.

Two data sets were collected on DOY 114 and DOY 087, 1992 at DSS 15. The Allan deviation and associated phase profiles are shown in Figs. 7 and 8, respectively.

The stability measurement on DOY 114 shows a hump at the integration period between 200 and 1000 sec. It is believed to be associated with a steep phase slope, i.e., frequency jump, as indicated in Fig. 8. It is unknown whether such a frequency jump occurred in the ×84 signal generator or the downlink system. Despite such a peculiar appearance, the stability performance is very good, down to the 4×10^{-16} level at 1000-sec integration.

Both stability measurements on DOY 114 and DOY 087 show a cyclic appearance. The cause is not yet identified since the phase profiles do not indicate any particular feature that may result in such oscillation.

B. Stability Performance Under Frequency-Dynamic Conditions

Measurement under frequency-dynamic conditions is limited to two-way testing. One-way measurement is not possible because the ×84 signal generator only generates a fixed signal at 8400 MHz. In these measurements, the frequency of the signal coming into the X-band maser was increased at an arbitrary rate of 0.94 Hz/sec, a dynamic that is comparable to what is received from spacecraft.

Figure 9 shows the measured stability on DOY's 019 and 068, 1991 at DSS 15, and DOY 340, 1992 at DSS 45. Among these three measurements, DOY 340 represented the best and DOY 068 the worst.

For DOY 340, the Allan deviation at 1000 sec was comparable with those measured under static conditions (see Fig. 4). At shorter integration, the DOY 340 result was somewhat higher. Oscillation in the Allan deviation was observed, but the cause is unknown.

For DOY 019, the stability was not as good. The Allan deviation was about three time larger than the static test on DOY 105, i.e., 1.985×10^{-15} instead of 7.30×10^{-16}. Several phase glitches on the order of 5 to 10 degrees were observed, as seen in Fig. 10(a).

For DOY 068, the 1000-sec Allan deviation was as high as 5.9×10^{-15}, which exceeded the system requirement. This high level of instability was believed to be caused by the changes in the signal frequency, as indicated by the three phase segments of different slopes in Fig. 10(b). A large phase glitch on the order of 90 deg was also observed at 5:56 GMT.

What could cause the frequency glitches observed on DOY 068? It is believed that the programmable synthesizer, either in the uplink or downlink path, was responsible, rather than the phase ripple in the front-end filtering. Had they been caused by the filter profiles, this phenomenon would have been observed in the other two tests on DOY 019 and DOY 340. Future testing under this configuration would be able to verify this conjecture.

One clarification is needed regarding the data processing of DOY 068. A corruption in the data extraction process made the digitized samples of the original data records not available for a period between 5:26 and 5:54 GMT. This problem forced consideration of either using only half of the collected data—before or after the gap—or joining the two segments and treating them as if no gap had occurred. The first option would have rendered the results with a higher uncertainty, due to smaller sample sets. The second option would have reduced the uncertainty associated with the computed Allan deviations at long integration times, but it required an assumption that the gap statistic was the same as the rest of the data set. The presented data were processed with the second option.

C. Effect of Frequency References at the Antenna

Figure 11 presents measurements made with three different media used to transport the frequency references to the antenna, namely hard line coaxial, phase-stabilized, and fiber-optic cables.

[4] D. Howell, Telecommunications Systems Section, unpublished test results (internal document), Jet Propulsion Laboratory, Pasadena, California.

1. Comparison with hard lines. In the earlier configuration, where the hard line cables were used to transport the antenna references, a separate link was used for the uplink and downlink references. Therefore, the link contribution was included in the system measurement, rather than canceled, as in the case of single-link fiber-optic or phase-stabilized cables. As seen in Fig. 11, the stability at 1000 sec for the hard line measurement was about 2.5 times greater than the fiber-optic measurement. The difference at short integration times is believed to be caused by different SNR settings between the two tests.

Note that a fair comparison required that a two-way contribution of the fiber-optic link be included in the fiber-optic measurement. However, since the contribution of the fiber-optic link is an order of magnitude smaller than the actual measurement for the rest of the system, the sum of the two factors, in the mean square sense, can be approximated by the actual measurement.

2. Comparison with phase-stabilized cables. In the configuration associated with fiber-optic and phase-stabilized references, a common link was used to transport the antenna references. Because of this, one would have expected both measurements to yield the same performance because the link contribution is not observable. Surprisingly, the results indicated otherwise. Fiber-optic measurement on DOY 035, 1993 was about 60 percent better than the phase-stabilized measurement on DOY 014, 1991 over the range of 1000–3600 sec of integration.

D. Impact of Antenna Motion

The primary impacts of the antenna pointing, such as surface deformation due to gravity, imperfect compensation to the loading by the subreflector, etc., are not observable in the test setup. However, some secondary effects should be observable. These include the vibration imposed on the assemblies that are housed in the antenna cone, and the stress on the RF cables that run between the antenna tilting room and the pedestal. The vibration, if large enough, should show up as periodic variation at the frequency of vibration. Cable stress, on the other hand, would be more likely to manifest itself as a fast linear phase drift. Such a drift, on the order of 70 deg over 500 sec, had been observed by Rebold[5] when the antenna was driven at maximum slew rate of 0.5 deg/sec.

In the measurements, the antenna was moved at a nominal spacecraft-tracking rate, equivalent to a sidereal rate of 0.004 deg/sec in the azimuth coordinate. As indicated in Fig. 12, comparison is done with DSS-65 (DOY 014 versus DOY 021, 1991) and DSS 15 (DOY 293, 1990). The result from DOY 293 indicates an imperceptible effect of antenna motion whereas the data from DOY 014 and 021 show some degree of degradation. However, the degradation is small enough that it can be attributed to daily variation in the measurements.

[5] T. A. Rebold, "Status of DSS 15 Radio Science System Tests for Galileo Gravity Wave Experiments," JPL Interoffice Memorandum 3933-90-82 (internal document), Jet Propulsion Laboratory, Pasadena, California, June 25, 1990.

E. Radio Frequency Interference

1. With radar. Testing on DOY 272, 1990 overlapped with the Goldstone Solar System Radar experiment at DSS 14. Figure 13 clearly shows the impact of the nearby RFI. The period during which large phase bursts occurred correlated exactly with the transmission time of the 400-kW radar transmitter at 8495 MHz. Despite the fact that the radar signal was 95 MHz away from the 8400-MHz test signal, its effect was detrimental.

2. With dual-signal reception. For earlier measurements, both one-way and two-way measurements were made at the same time. That is, signals from both the $\times 84$ signal generator and the test translator were simultaneously injected into the X-band maser. They were set at about 50 to 100 Hz apart, a separation that was considered wide enough to minimize the interference, yet close enough to have both signals inside the bandwidth of the last filter in the system. The rationales for dual-signal testing were to minimize the required test time and to enable a direct comparison between one-way and two-way measurements under the same conditions. It was later learned that there was intermodulation between the two signals that could not be eliminated by the 1-Hz lowpass filter used in data processing. Such a degradation can be seen in Figs. 14(a) and 14(b) in terms of the phase and Allan deviation, respectively, for DOY 087, 1992 testing.

F. Weather Effects

1. Impact of winds. Under strong wind conditions, antenna structural resonance may occur which would affect the equipment housed in the antenna cone. Different wind conditions were observed over the span of two-year testing. The worst case experienced was gusty winds up to 45 mph. Nevertheless, no correlation was seen between wind speed and stability degradation.

2. Impact of temperature. In order to study the impact of temperature on the 300-MHz IF cables that connect the antenna and the SPC area, an attempt was made to correlate the ambient temperature and the detected signal phase. This ambient temperature was measured by the DSCC Media Calibration (DMD) Subsystem. Although

the DMD temperature did not accurately reflect the temperature at the trench through which the cables ran, the result showed a correlation between the DMD temperature and the phase variation.

Five different data sets (DOY 035, DOY 027, DOY 114, DOY 105, and DOY 087) were used for this particular study. Any linear phase drift was first removed from the data set. This removal was necessary because any constant frequency offset caused by the inaccurate estimation of the input frequency would result in a linear phase drift. Unfortunately, by removing linear phase drift, any phase drift caused by thermal conditions that had the similar characteristic was also removed. Therefore, it was also necessary to fit the temperature with a linear function, and the residual temperature was then used in the correlation.

Figures 15(a–e) show the variation of residual temperature and phase, as a function of time, for DOYs' 035, 027, 114, 105, and 087. There was an apparent correlation between the two sets of residuals, although the effect was offset in time. Figure 16 presents a better picture of correlation as the residual phase is plotted against the residual temperature. Samples from all five data sets are used. The correlation can be observed by examining the slope of the best-fit curve, which is about 5.46 deg of phase per deg C. This value agrees with an expected value of 5.4 deg of phase per deg C for the effect of temperature on the 300-MHz IF cables of typical 400-m length, calculated from a thermal coefficient of delay of 25 ppm/deg C and velocity of propagation 70 percent of the speed of light. This derived thermal coefficient of delay (TCD) is higher than expected, but well within the observed range for coaxial cables. It is also possible that other equipment contributes to the drift. If so, the derived TCD would be proportionally lower.

VI. Performance Versus Requirement

Table 3 presents the design specifications for different subsystems in the Radio Science System. The requirement for the stability of the whole system, 4.5×10^{-15} stability at 1000-sec integration, is based on the need to support the Galileo mission.[6] Excluding the contribution of FTS and ANT, the requirement to which the measurements presented in Section V are compared is 3.0×10^{-15}. With the exception of data on DOY 068, all measurements are less than the required 3.0×10^{-15} level.

How do the specifications for the FTS and ANT measure against the actual performance? As shown in Table 4, stand-alone stability measurements for the FTS, based on Calhoun and Law's work,[7] indicated that the subsystem performs well within its specifications. For the antenna, if one assumes that the stability of the 34-m HEF antenna at X-band is similar to that of the R&D 34-m antenna at DSS 13 at K-band, then the antenna contribution is 1.4 times (i.e., $\sqrt{2}$) greater than its allocation. However, the combined FTS and ANT for the actual measurements is still less than their combined allocation. Thus, one can reasonably say that the total system performance, including both the observables and nonobservables, meets its requirements.

VII. Conclusions

Several conclusions can be drawn regarding the performance of the Radio Science System at the 34-m HEF stations.

(1) The system, under both one- and two-way configurations, performed well within its specification. However, the phase profiles indicated several small glitches and periodic variations, the cause of which was not identified. For the current requirements, these peculiar features do not cause the system to be out of specification; however, they should be investigated if future requirements will be more stringent. In the absence of known hardware failures, consistency among measurements of different dates demonstated that system performance is reliable.

(2) Frequency-dynamic testing indicated that system performance varied greatly from day to day. Out of three sets of measurements, one showed no difference in stability compared to the frequency-static results, one was more than twice as unstable, and one exceeded the specification. More data need to be collected to fully characterize the system under this mode of operation.

(3) The secondary effects of antenna motion, in terms of stretching the RF cables, and the vibration that the motion imposed on the equipment, appeared to be minimal when it occurred at the sidereal tracking rate.

(4) System stability was found to be very sensitive to RF interference. Most notable was the interference from the nearby radar 400-kW transmitter at the 70-m antenna. The presence of other signals in the

[6] Support Instrumentation Requirements Document for the 1989 Galileo Project, 625-501, Rev. A, JPL D-476 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 1988.

[7] See Footnote 2.

operating bandwidth, such as a second test signal, also degraded stability performance.

(5) Definite correlation between the signal phase and the ambient temperature was observed, indicating that the 300-MHz IF cables connecting the antenna and signal-processing center are sensitive to thermal effect. If such coaxial cables are to be replaced with fiber-optic cables in the future, stability performance is expected to improve.

Acknowledgments

The authors express their gratitude to the following individuals, without whose contribution this study would not have been possible. First, appreciation extends to Malcom Calhoun, Julius Law, and Paul Kuhnle from the Frequency and Timing Subsystem Group for providing the engineering model of the fiber-optic distribution assembly and for their tireless support in early testing. Secondly, the authors thank Diana Howell and Jim Weese for spending many working nights at Goldstone to complete the system performance tests. Percy Montoya's accommodation to urgent requests for data extraction from the original data record tapes is much appreciated. The authors also thank North Ham and Tom Rebold for their technical introduction to the Radio Science System. The measurements would not have been realizable without the great support of operations personnel at all three Deep Space Communications Complexes; thanks in particular to Len Litherland, Bob Jenkins, and Kevin Knights at DSCC 40; and Larry Bracamonte, Guy Kaufman, and Rich Overcash at DSCC 10.

Table 1. Comparison of Allan deviations among three different methods of data processing at DSS 15.

Test	Integration, sec	Allan deviation, x 10^{-15}		
		Complex demodulation	DPLL	Time interval
	1	229.0	157.0	220.4
DOY 177, 1991	10	19.5	28.16	68.52
11:13–15:07	100	8.60	10.73	22.04
	1000	2.81	5.665	6.852
	1	377.5	189.1	129.0
DOY 334, 1990	10	53.80	35.14	20.2
1:30–11:30	100	18.05	3.969	2.35
	1000	4.507	0.9240	0.795
	1	68.44	99.52	129.0
DOY 027, 1993	10	9.730	11.77	15.5
11:09–16:27	100	1.595	1.844	2.05
	1000	1.362	1.337	1.25

Table 2. Summary of measurements collected under different configurations.

Focus of study	DSS site	Test	Link configuration	Antenna configuration	Antenna reference	Frequency condition	Integration time, sec	Allan dev., x 10^{-15}
Two-way stability under static condition	65	DOY 035, 1993 04:19–13:00 8.7 hr	2-way	S[a]	F[b]	S[c]	1 10 100 1000 3600	103.6 16.35 4.159 0.8725 3.808
	45	DOY 067, 1993 00:33–07:06 6.6 hr	2-way	S	F	S	1 10 100 1000 3600	114.5 21.70 3.221 0.7944 0.5218
	45	DOY 341, 1992 11:50–21:15 9.5 hr	2-way	S	F	S	1 10 100 1000 3600	236.1 28.01 3.383 0.2626 0.6796
	15	DOY 027, 1993 11:09–16:27 5.3 hr	2-way	S	F	S	1 10 100 1000 3600	99.52 11.77 1.844 1.337 2.995
	15	DOY 105, 1992 02:50–12:50 10.0 hr	2-way	M[d]	F	S	1 10 100 1000 3600	164.4 7.750 1.081 0.7304 0.5727
Contribution of TXR	15	DOY 032, 1993 time not available 4 hr	2-way without TXR	S	F	S	1 10 100 1000	130.0 14.8 2.21 0.754
	15	DOY 032, 1993 time not available 2 hr	2-way with TXR	S	F	S	1 10 100 1000	129.0 14.8 1.61 0.421
One-way stability under static condition	15	DOY 114, 1992 01:49–10:04 8.7 hr	1-way	M	F	S	1 10 100 1000	143.9 941.31 0.9063 0.4507
	15	DOY 087, 1992 12:45–22:25 2.4 hr	1-way	M	F	S	1 10 100 1000	142.2 24.70 2.460 0.2969
Two-way stability under dynamic condition	45	DOY 340, 1992 13:00–22:25 9.4 hr	2-way	S	F	D[e]	1 10 100 1000 3600	195.9 36.98 1.943 0.5771 0.8945
	15	DOY 019, 1991 05:08–07:59 2.8 hr	2-way	M	P[f]	D	1 10 100 1000 3600	119.5 23.41 5.706 1.985 0.3898

[a] Static (antenna configuration only).
[b] Fiber-optic link.
[c] Static (fixed).
[d] Moving.
[e] Dynamic (ramped).
[f] Phase-stabilized cable.

Table 2. (cont'd).

Focus of study	DSS site	Test	Link configuration	Antenna configuration	Antenna reference	Frequency condition	Integration time, sec	Allan dev., $\times 10^{-15}$
	15	DOY 068, 1991 03:25–07:35 4.1 hr	2-way	M	P	D	1 10 100 1000 3600	296.5 48.29 14.00 5.940 11.31
Stability with other, non-fiber-optic references	15	DOY 086, 1990 19:46–21:50 2.0 hr	1-way	S	H[§]	S	1 10 100 1000	356.0 31.6 3.43 4.20
	65	DOY 014, 1991 14:22–01:28 11.1 hr	2-way	S	P	S	1 10 100 1000 3600	93.73 15.09 1.951 1.708 0.2379
Impact of antenna motion	15	DOY 293, 1990 06:06–10:36 4.5 hr	2-way	Stationary	F	S	1 10 100 1000	130.0 17.0 2.2 0.24
	15	DOY 293, 1990 11:55–13:15 1.3 hr	2-way	Moving	F	S	1 10 100 1000	130.0 17.0 3.1 0.54
	65	DOY 021, 1991 13:35–21:55 8.3 hr	2-way	Moving	F	S	1 10 100 1000 3600	94.50 60.15 5.308 0.7061 0.5633
	65	DOY 014, 1991 14:22–01:28 11.1 hr	2-way	Stationary	P	S	1 10 100 1000 3600	93.73 15.09 1.951 1.708 0.2379
Impact of RF interference								
From radar	15	DOY 272, 1990 04:10–07:45 3.6 hr	1-way	S	P	S	Not analyzed for stability. See Fig. 13.	
From exciter/ translator	15	DOY 087, 1992 Single-signal 12:45–15:05 2.4 hr	1-way	S	F	S	1 10 100 1000	142.0 24.70 2.46 0.297
	15	DOY 087, 1992 Dual-signal 10:42–12:45 3.0 hr	1-way	S	F	S	1 10 100 1000	227.0 37.3 2.79 0.343

[§] Hard-line coaxial cable.

Table 3. Design specification of components in the system.

Subsystem	Allan deviation, x 10^{-15} (1000-sec integration time)
Measurables	
Exciter	1.8
Transmitter	1.0
Receiver	2.1
Microwave (2-way)	0.7
DSCC spectrum processor	0.1
Subtotal of measurables	3.0
Nonmeasurables	
Frequency and timing (2-way)	3.0
Antenna (2-way)	1.4
Subtotal of nonmeasurables	3.3
System total (rss)	4.5

Table 4. Stand-alone measurements on the stability of the frequency and timing and antenna subsystems.[a]

Integration period, sec	Frequency and timing, Allan deviation, x 10^{-15} (1-way)	Antenna, Allan deviation, x 10^{-15} (1-way)
1	200.0	280.0
10	40.0	100.0
100	8.0	12.0
1000	1.5	1.4

[a] Based on Footnote 3 and [4].

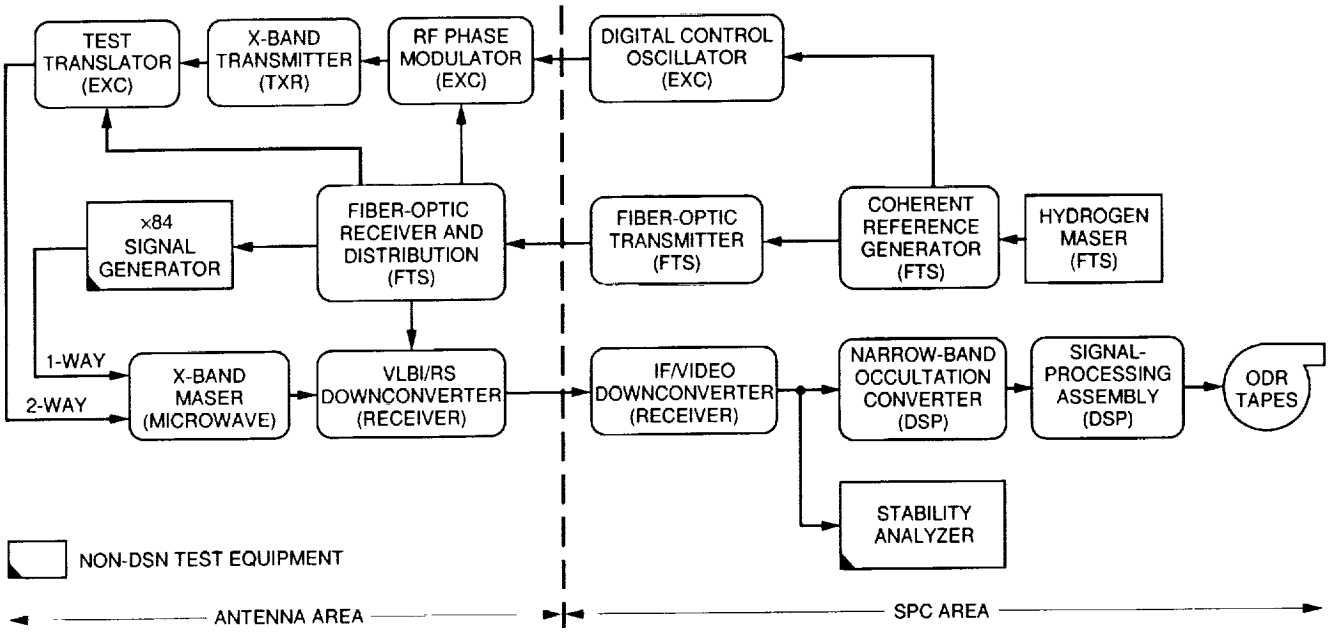


Fig. 1. Test configuration for stability measurements.

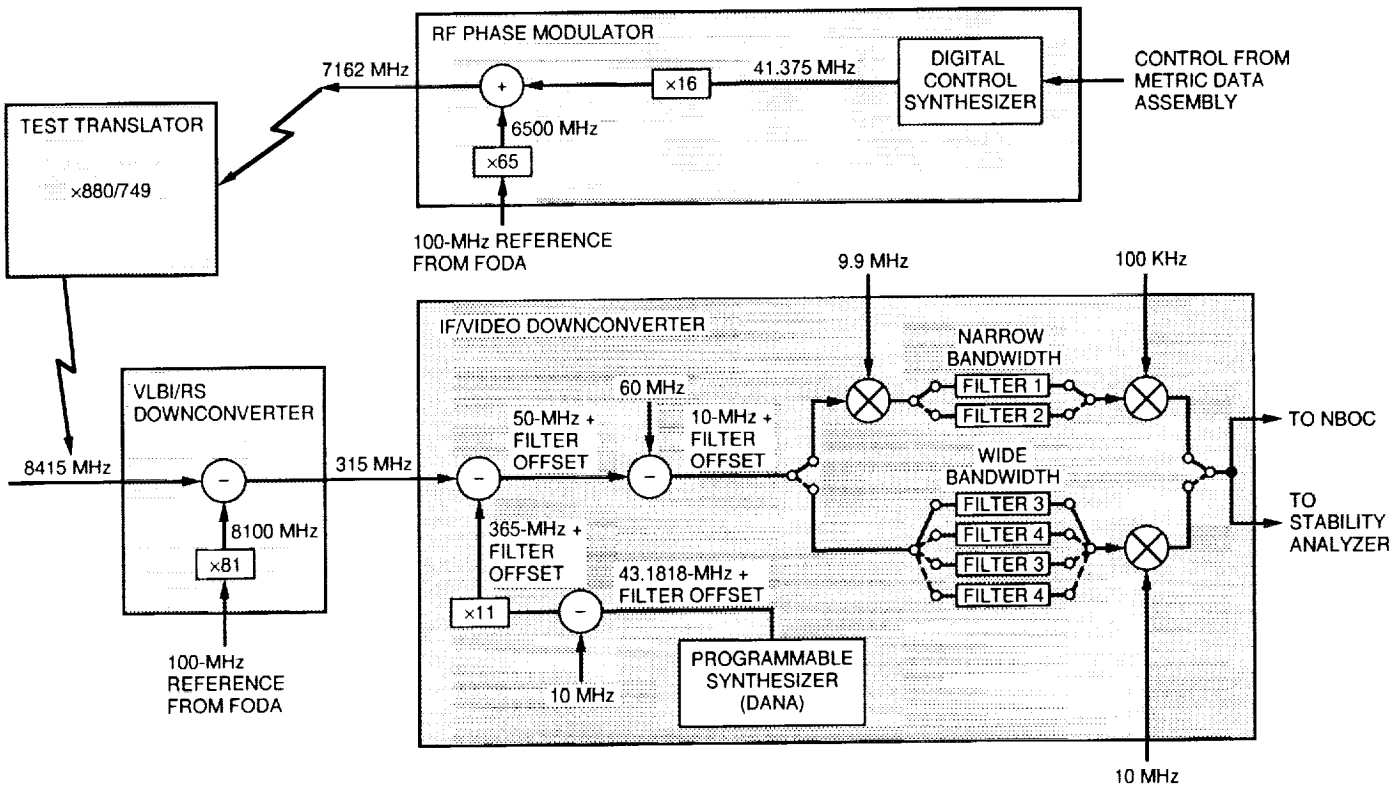


Fig. 2. Uplink and downlink frequency conversions.

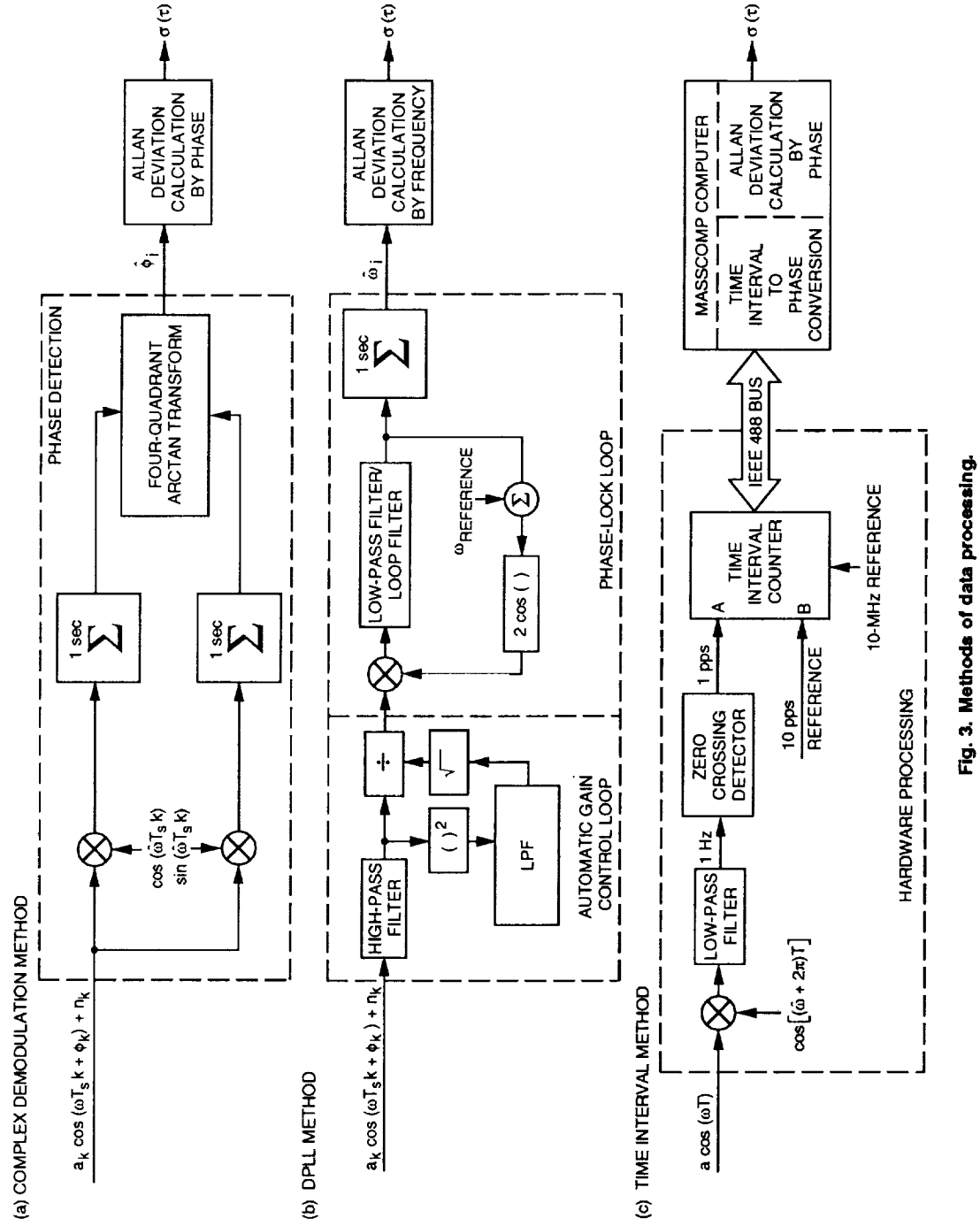


Fig. 3. Methods of data processing.

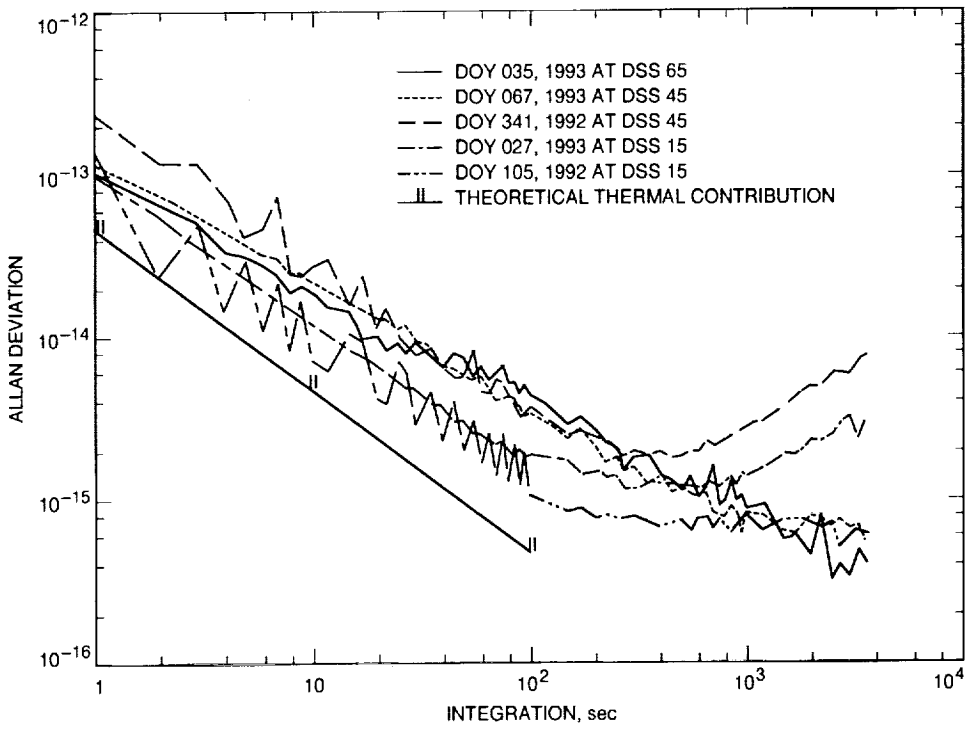


Fig. 4. Two-way stability with static frequency.

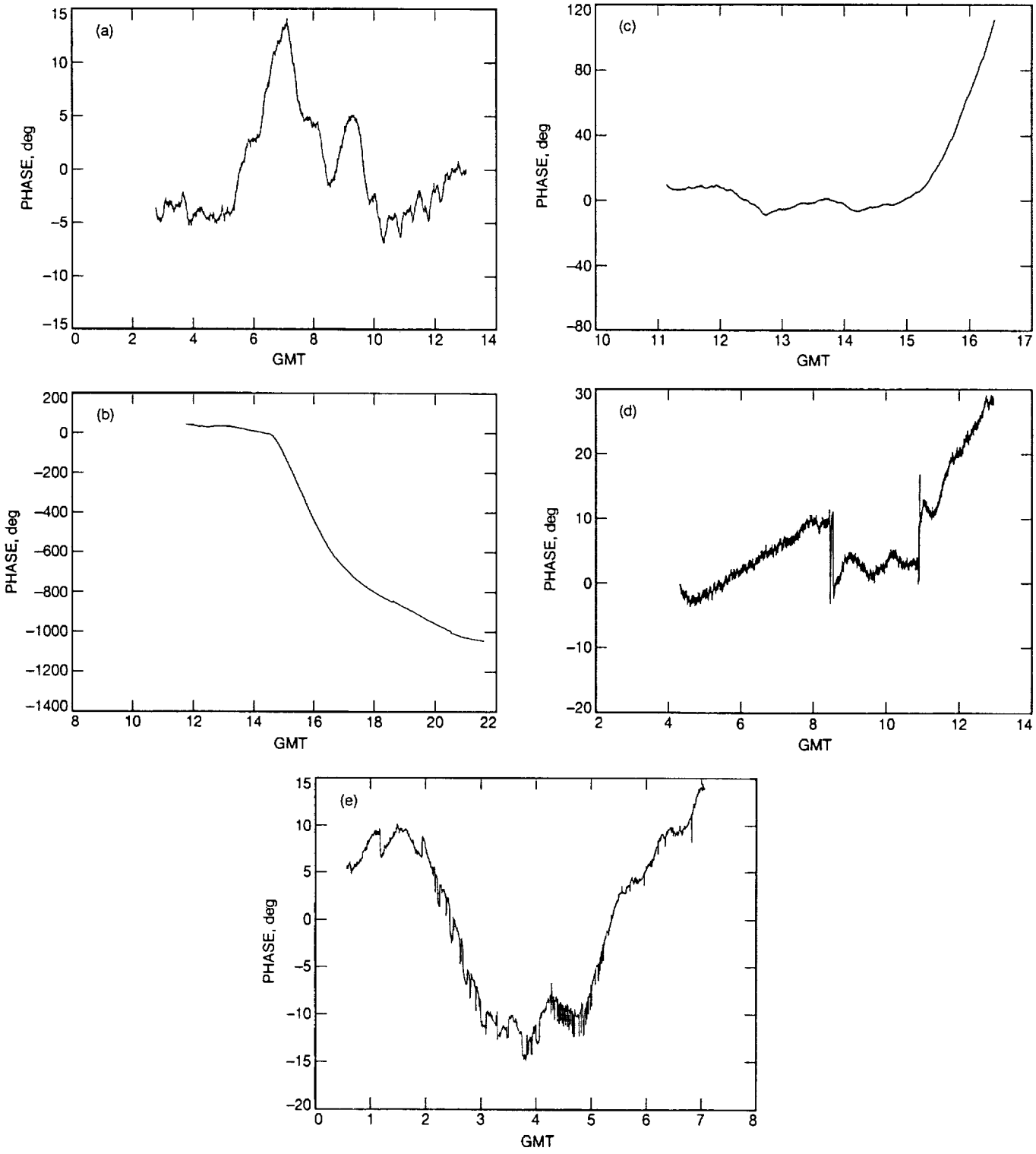


Fig. 5. Phase profiles for two-way frequency-static measurements: (a) DOY 105, 1992 at DSS 15; (b) DOY 341, 1992 at DSS 45; (c) DOY 027, 1993 at DSS 15; (d) DOY 035, 1993 at DSS 65; and (e) DOY 067, 1993 at DSS 45.

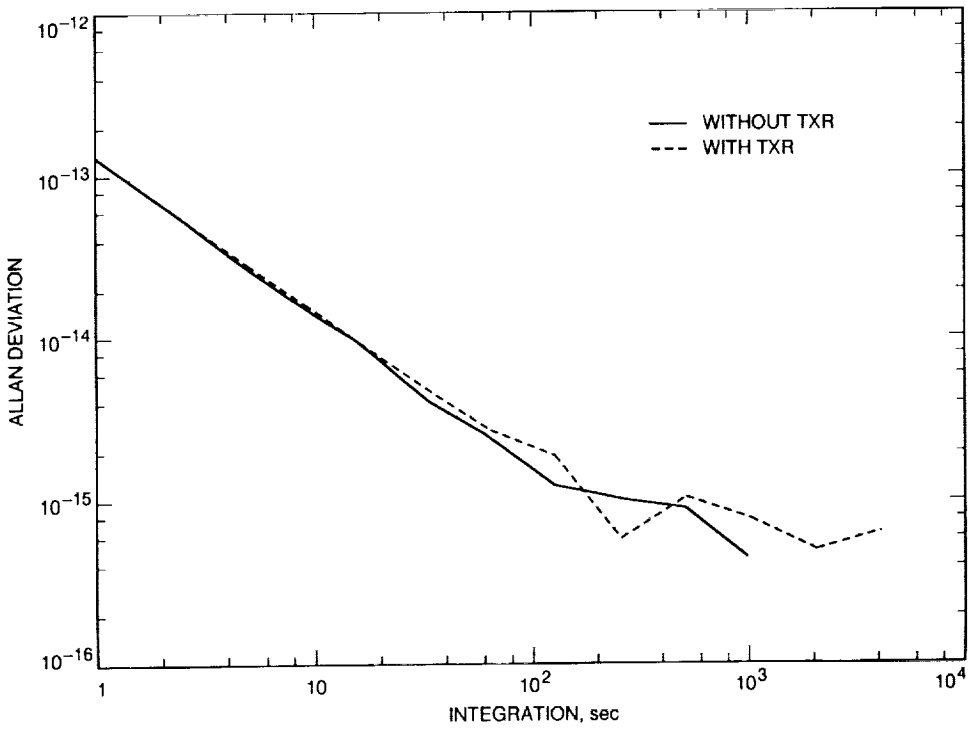


Fig. 6. Two-way stability with and without transmitter, DOY 032, 1993 at DSS 15.

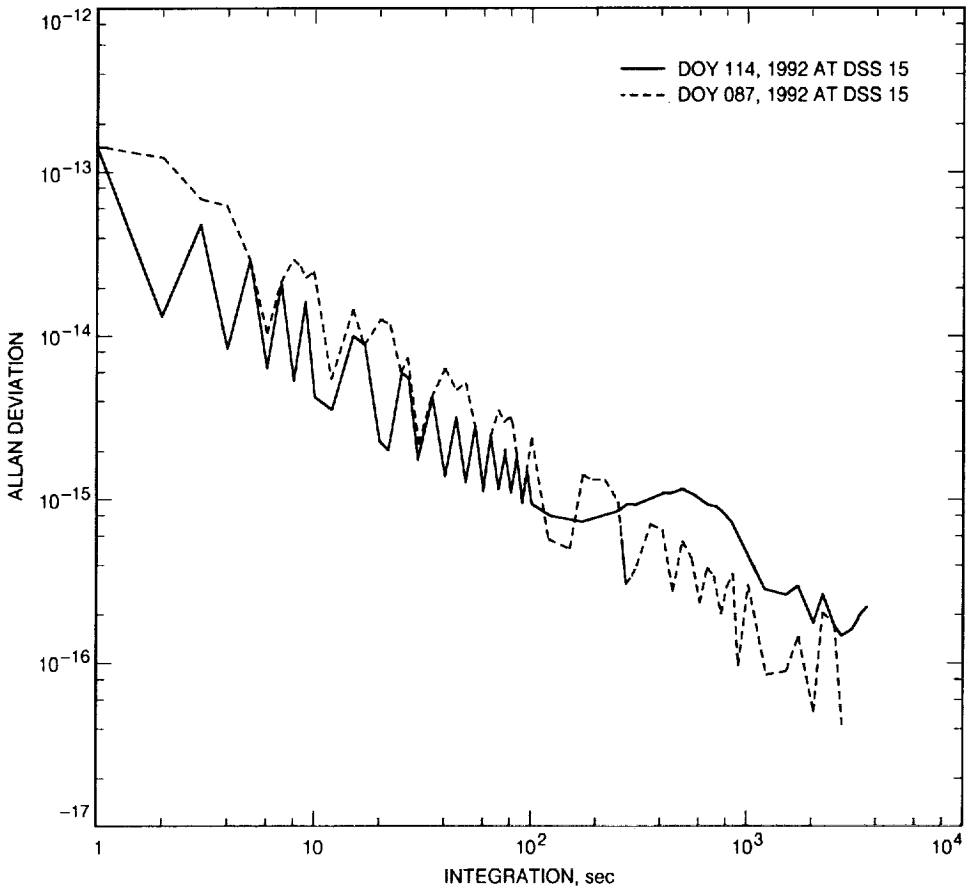


Fig. 7. One-way stability with static frequency.

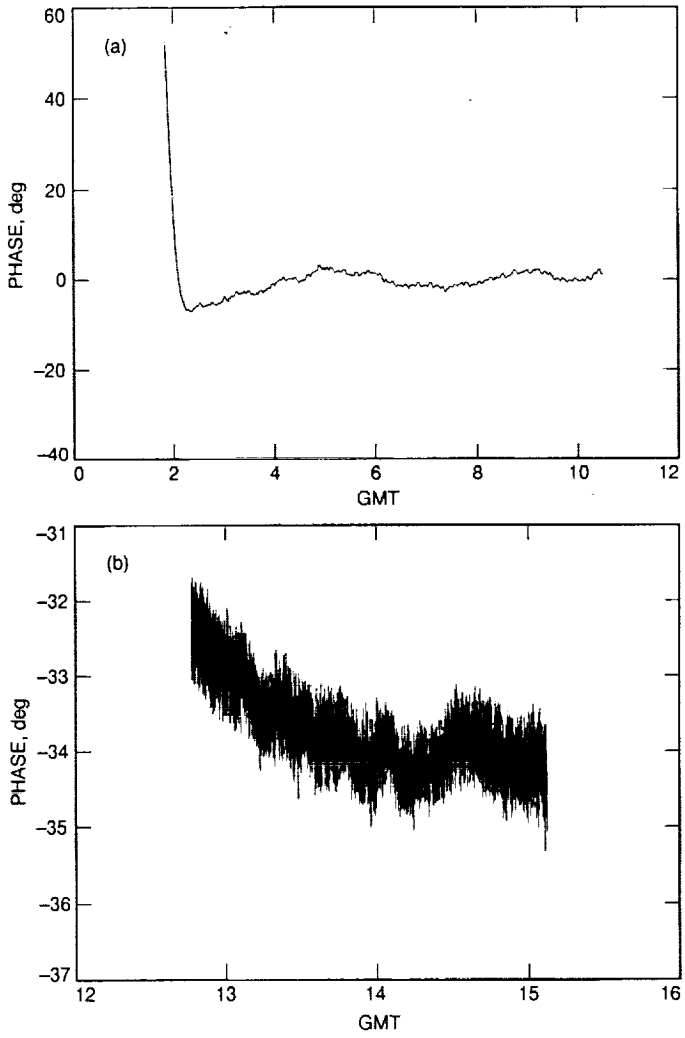


Fig. 8. Phase profiles for one-way frequency-static measurements at DSS 15: (a) DOY 114, 1991 and (b) DOY 087, 1992.

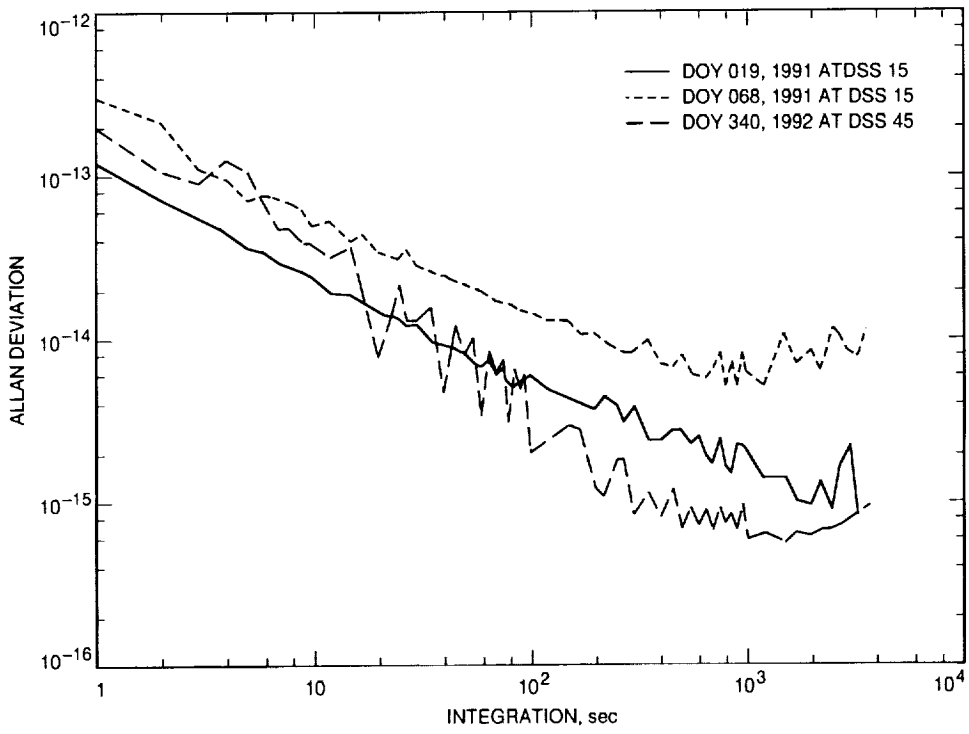


Fig. 9. Two-way stability with dynamic frequency.

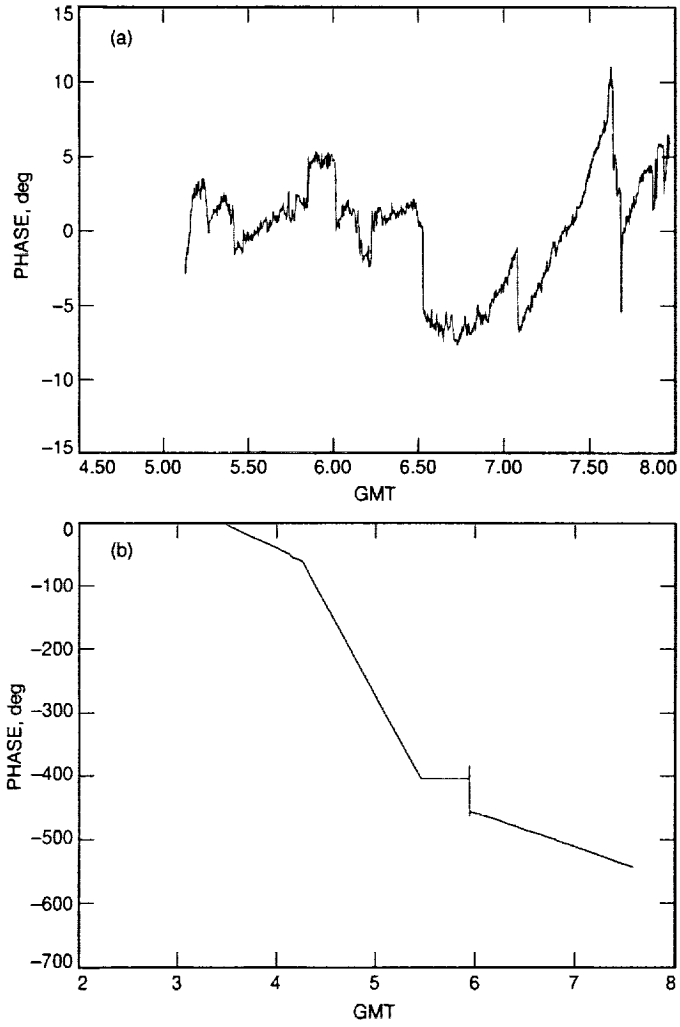


Fig. 10. Phase profiles at DSS 15: (a) DOY 019, 1991 and (b) DOY 068, 1991.

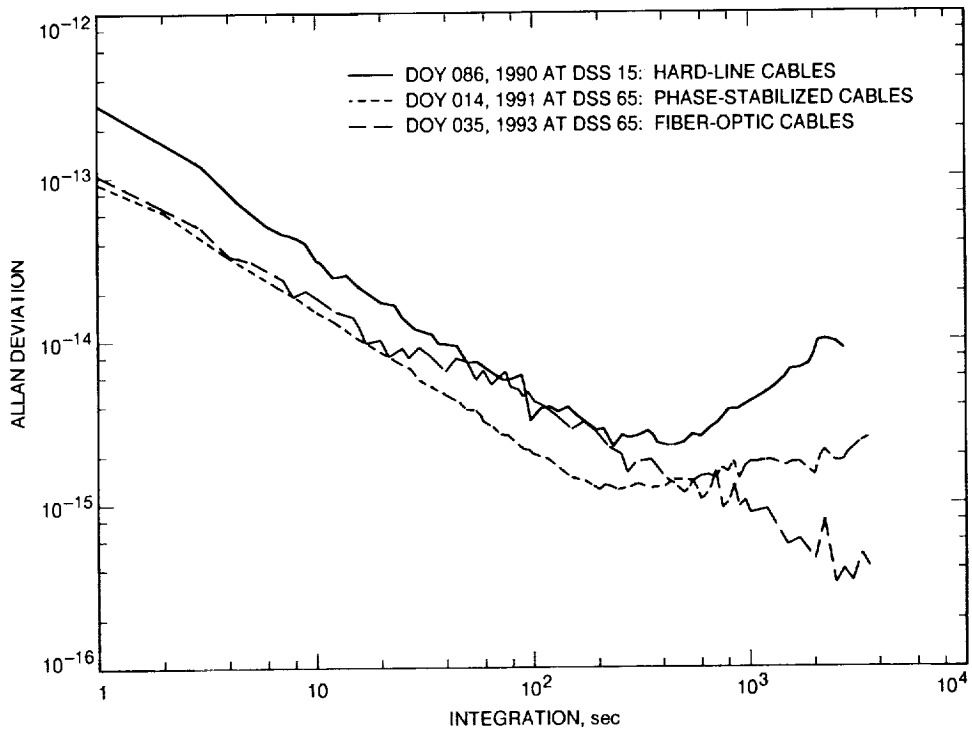


Fig. 11. Effect of antenna references.

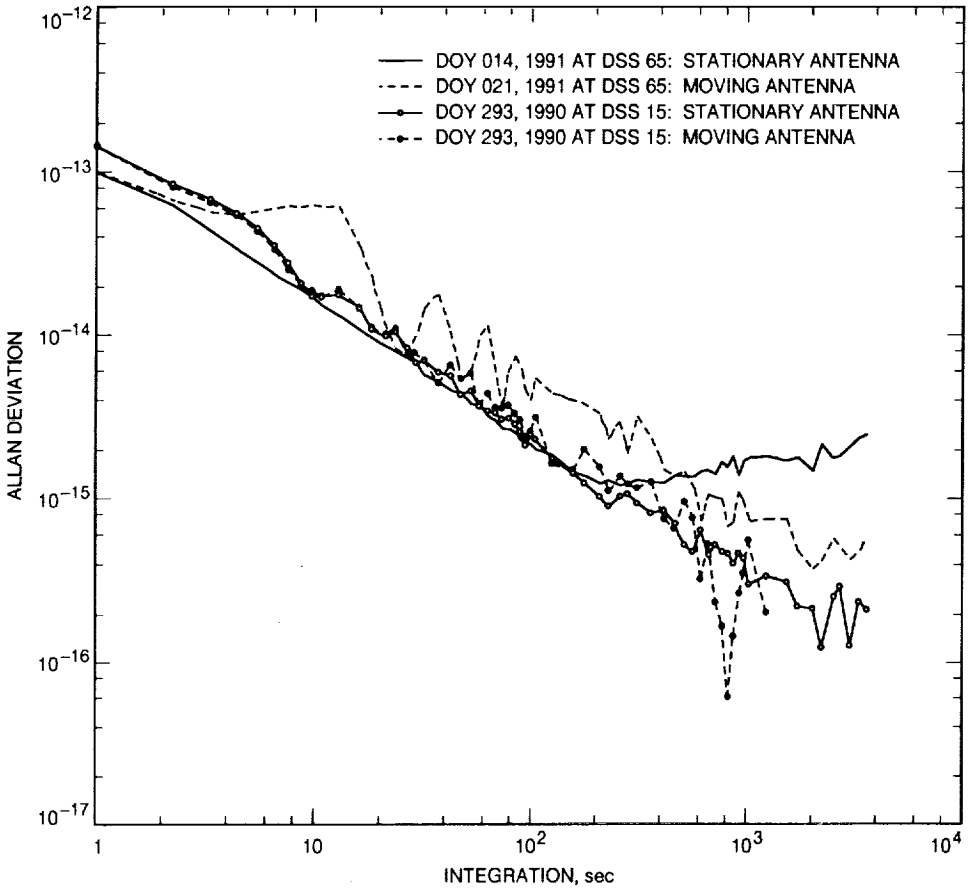


Fig. 12. Effect of antenna motion.

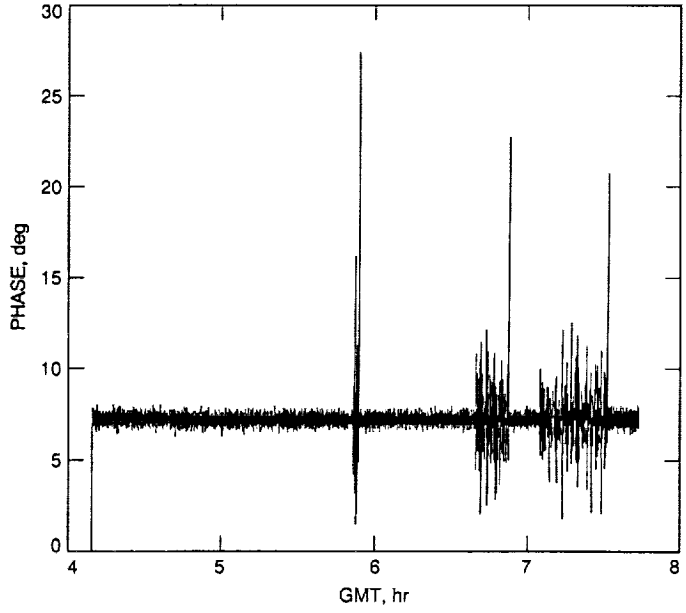


Fig. 13. Effect of RFI from radar transmission.

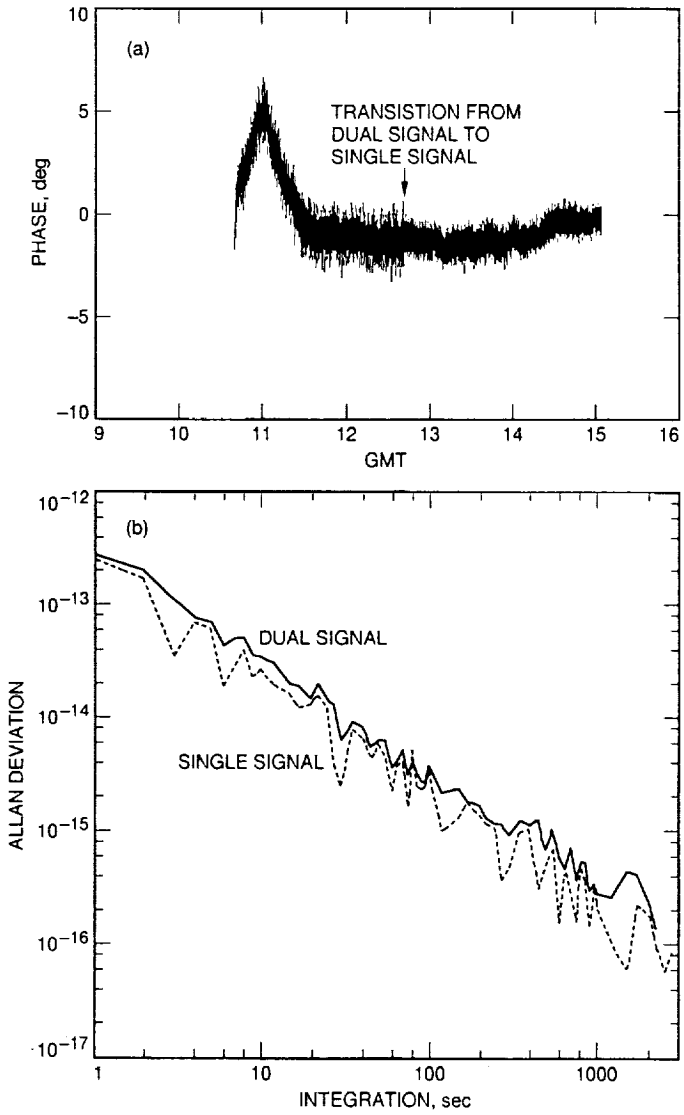


Fig. 14. Effect of RFI from dual-signal testing, DOY 087, 1992: (a) phase profile and (b) stability.

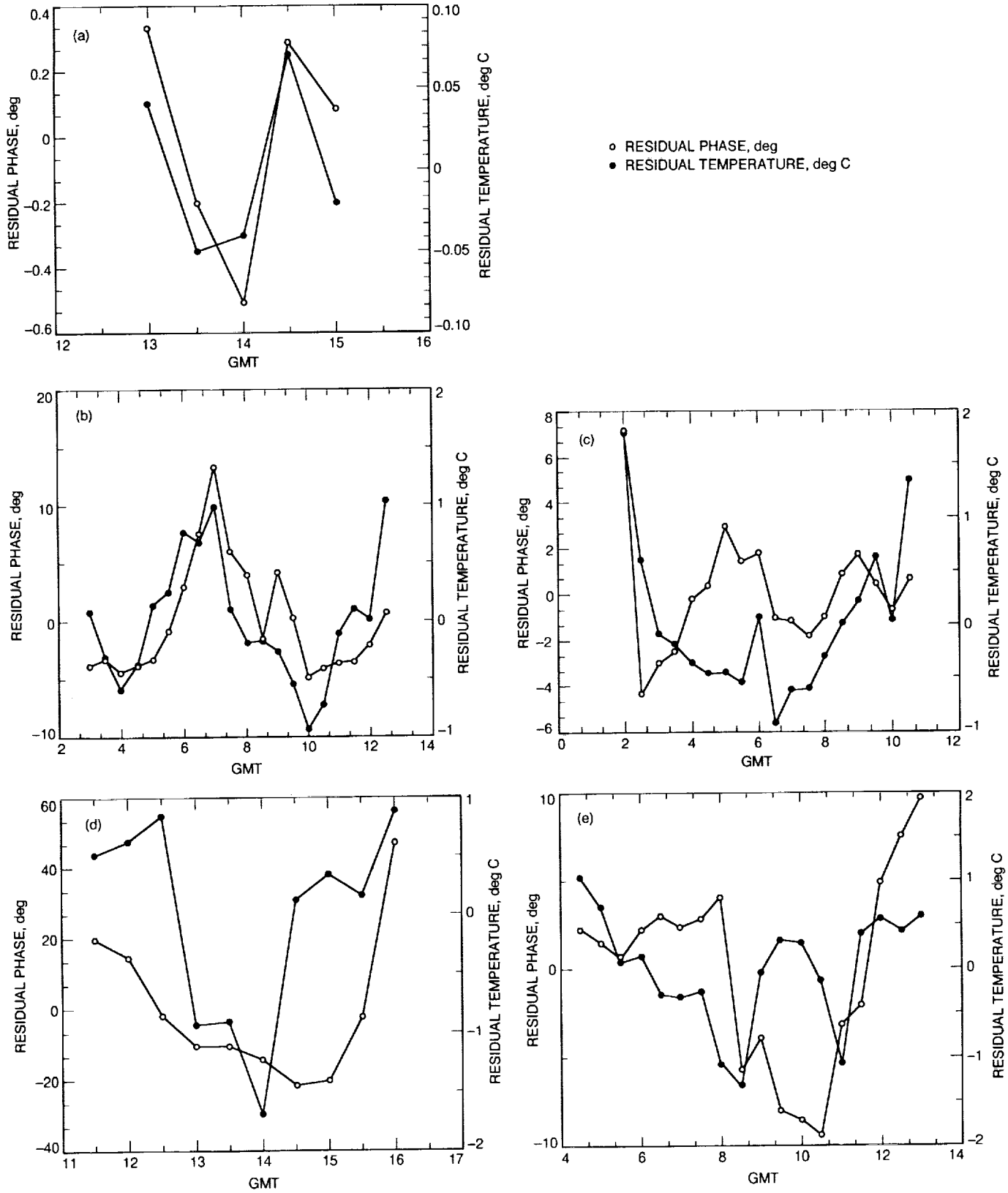


Fig. 15. Residual temperature and phase versus time: (a) DOY 087, 1992 at DSS 15; (b) DOY 105, 1992 at DSS 15; (c) DOY 114, 1992 at DSS 15; (d) DOY 027, 1993 at DSS 15; and (e) DOY 035, 1992 at DSS 65.

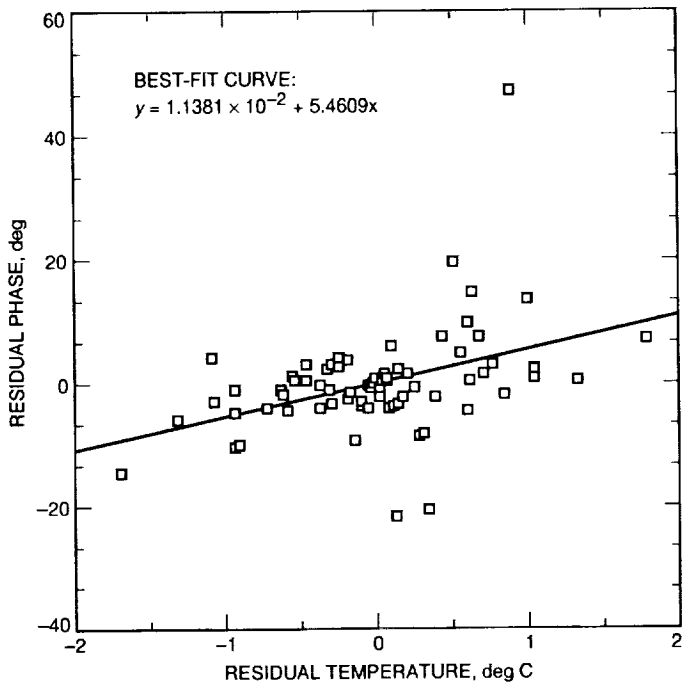


Fig. 16. Correlation of phase and temperature.

TDA Progress Report 42-114

August 15, 1993

Performance Results of a Digital Test Signal Generator

B. O. Gutierrez-Luaces
Telecommunications Systems Section

M. Marina
Radio Frequency and Microwave Subsystems Section

B. Parham[1]
Space Communications Data Systems Section

Performance results of a digital test signal-generator hardware-demonstration unit are reported. Capabilities available include baseband and intermediate frequency (IF) spectrum generation, for which test results are provided in this article. Repeatability in the setting of a given signal-to-noise ratio (SNR) when a baseband or an IF spectrum is being generated ranges from 0.01 dB at high SNR's or high data rates to 0.3 dB at low data rates or low SNR's. Baseband symbol SNR and carrier SNR (P_c/N_0) accuracies of 0.1 dB have been verified with the built-in statistics circuitry. At low SNR's that accuracy remains to be fully verified. These results have been confirmed with measurements from a demodulator synchronizer assembly for the baseband spectrum generation, and with a digital receiver (Pioneer 10 receiver) for the IF spectrum generation.

I. Introduction

A conceptual design description, the general capabilities, and expected performance of a digital test signal generator (DTSG) with immediate application in the DSN was published in [1]. A new method of generating analog test signals with accurate signal-to-noise ratios (SNR's) was described. High accuracy would be obtained by simultaneous generation of digital noise and signal spectra at a given baseband or bandpass-limited bandwidth.

A DTSG hardware system became available in September 1991. Considerable effort was put into the operational, monitor, and control aspects of its operation. At the end of the project, full capabilities for baseband generation of data modulated in a single subcarrier, biphase data, or non-return-to-zero (NRZ) data were completed by December 1991. A limited version for intermediate frequency (IF) generation up to 300 MHz was also completed.

Although some DSN-telemetry standard test cases are available to the operator, it should be understood that the potential of generating precise SNR for conditions defined by the user is an option, although each particular case probably will need special programming to attain the results expected by the experimenter. Simple DTSG operation is provided through a user interface menu. Capabil-

[1] Consultant, OAO Corporation, Altadena, California.

ities included in both versions of this demonstration unit have been summarized in [1].

A detailed description of the measured performance of a few basic functions has been found necessary in helping the user understand the system limitations.

II. DTSG Description

Figure 1 is a composite of photos of the different pieces of electronic equipment used in the demonstration unit of the DTSG. These pieces of electronic equipment are mounted in a standard DSN rack. From top to bottom, the following equipment can be found:

(1) A commercial 100-Hz to 1.5-GHz spectrum analyzer used to analyze the DTSG analog output at different frequency bands. In Fig. 1, the spectrum analyzer actually shows the DTSG analog frequency response output of a baseband 400-kHz digital filter.

(2) A commercial PC with keyboard and monitor. The monitor shows the actual histogram obtained from an unbiased-noise digital output.

(3) A commercial programmable synthesizer providing a 2- to 20-MHz system clock. This system clock may be varied in accordance with the Doppler effect desired.

(4) DTSG box, which contains all of the special hardware designed at JPL by the Radio Frequency and Microwave Subsystems Section.

Figure 2 shows the PC monitor screen with the menu provided when the DTSG is configured in the baseband mode. Figure 3 is the same monitor screen when the DTSG is configured in the IF mode. Note that all the parameters shown in Figs. 2 and 3 may be changed through the PC keyboard and the user interface menus shown in the figures. After the operator has assigned the values to use as test source parameters, the initialization (INI) command may be entered (operational commands are displayed at the bottom of the DTSG menu on the monitor screen in Figs. 2 and 3). This command directs the PC to initiate all the calculations necessary to generate all the hardware parameters to be later transmitted to the DTSG hardware box. Some of these parameters, generated in the form of ASCII output files, are a system clock to be loaded in the synthesizer, a filter coefficients file, a subcarrier frequency file, a symbol (data coded or uncoded) file, the required noise and signal attenuator factors for the symbol SNR (SSNR) defined, a noise file, and other parameters required by the statistics hardware circuitry.

Execution of the RUN command (see Figs. 2 and 3) will transmit all the required configuration files to the DTSG hardware and will automatically start the measurement and periodic update of the DTSG output statistics in the channel selected. Statistics provided are

(1) instantaneous SSNR, the mean symbol SNR corresponding to an operator-controlled fixed number of symbols

(2) instantaneous SSNR standard deviation (STD), the standard deviation of the 10 previous SSNR's

(3) instantaneous symbol error rate (SER), the SER corresponding to a fixed number of symbols (operator controlled)

(4) average SSNR (AVRGSNR), the continuous average of all the previous measurements of instantaneous SSNR

(5) average SER (AVRGSER), the continuous average of all the previous measurements of instantaneous SER.

III. General Description of Configurations To Be Tested

Figure 4 is a block diagram of the SNR generator box, which contains the high-speed hardware used to synthesize the test signals. A description of the DTSG can be found in [1]. Engineering details on software and hardware are documented.[2]

The DTSG hardware unit (refer to Fig. 4) may be configured to generate different signal spectra. The DTSG modes available are baseband spectrum generation, to generate NRZ data (coded or uncoded), biphase data, or NRZ data modulated in a subcarrier; dual-subcarrier spectrum generation, to generate two NRZ data streams modulated in two subcarriers; quadrature-phase shift key (QPSK) spectrum generation, to generate QPSK data; and IF spectrum generation, to generate an intermediate-frequency phase modulated with or without the carrier. This article will refer only to the baseband spectrum-generation mode and the IF spectrum-generation mode performance test results.

Non-return-to-zero-level (NRZ-L) uncoded data modulated in a subcarrier make up one of the types of data generated in the baseband mode. Data files are created

[2] *Digital Test Signal Generator Software Operator's Manual/User's Guide* (internal document), Jet Propulsion Laboratory, Radio Frequency and Microwave Subsystems Section, Pasadena, California, September, 1991.

in the PC by defining the different parameters available, which are given on the PC monitor (as shown in Fig. 2). The generation of these files is started after the INI command is executed. Thus, an NRZ-L data file, a subcarrier file, a noise file, a digital filter file, channel-1 attenuator and noise-channel attenuator files, and a statistics monitor file are created in the controlling PC.

The RUN command will configure the hardware with the files previously generated with the INI command. Therefore the data 64K RAM of channel 1 is loaded with the NRZ-L data file; the subcarrier 64K RAM of the same channel 1 is loaded with the subcarrier file; the digital filters are loaded with the appropriate coefficients, and the statistics-monitor hardware gain is properly scaled.

Statistical measurements of instantaneous SNR, SER, average SNR, and average SER are automatically provided and periodically updated. These statistical measurements may correspond either to the analog output or the digital output, as chosen by the operator by using the appropriate command (see Figs. 2 and 3).

Also, a real-time (20-MHz) histogram accumulator of 512 bins with a total of 10^6 samples is available in a file returned to the controlling PC. This file can be further processed in the PC or made available to any outside computer through a magnetic diskette. By this means, histograms requiring more than 512 quantized values may be produced (see Section IV).

IF spectra with residual carrier and NRZ-L data modulated in a subcarrier make up the type of data being generated in the IF mode. A limited capability for this configuration has been provided in the present software version. Only Magellan, Galileo, Voyager, and Pioneer 10 telemetry rates and characteristics may be simulated, although any modulation index or carrier suppression can be provided.

Following the same approach as described previously for the baseband mode, the PC will create the required hardware configuration files from the parameters specified in the PC monitor, as shown in Fig. 3. Therefore, different files will be generated after the INI command. The following files will be loaded in the SNR generator box after the RUN command is executed (refer to Fig. 4). The subcarrier 64K RAM of channel 1 will be loaded with a file to produce an output frequency of 5 MHz with a relative phase of 0 deg. No data file is required in channel 1. The subcarrier 64K RAM of channel 2 will be loaded with a file to produce an output frequency of 5 MHz with a relative phase of 90 deg. The data 64K RAM will be loaded with the baseband telemetry data of either Magellan, Galileo, Voyager, or Pioneer 10. The filter files will now correspond to bandpass filters centered at 5 MHz.

The capabilities of providing histograms and statistics are also available for this case with the added capability of being able to obtain statistics either from channel 1 (carrier statistics, P_c/N_0) or channel 2 (telemetry statistics, SSNR).

IV. Noise-Generation Characteristics

As discussed in [1], the noise 64K RAM of the noise channel is loaded with a discrete Gaussian distribution of data bytes (8-bit representation), to be sequentially read by a random-address generator. The DTSG has been provided with a histogram accumulator that allows acquisition of real-time histograms (20-MHz sampling clock) consisting of 10^6 samples in total.

Two histograms to cover the whole distribution of the unfiltered output of the noise memory were obtained and analyzed. The percent deviation of the discrete Gaussian distribution generated by the noise channel from a theoretical Gaussian discrete probability distribution with identical standard deviation, s, was defined as follows:

$$\text{deviation } (x_d; s) = \frac{\{[P_d \text{ actual } (x_d; s)/P_d \text{ gauss } (x_d; s)] - 1\}}{100} \quad (1)$$

where

$$x_d \in \{-127, ..., -1, 0, 1, 2, ..., 127\}, s = 32$$

and

deviation $(x_d; s)$ = percent of deviation for value x_d

P_d gauss $(x_d; s)$ = discrete probability function of a Gaussian distribution; see Eqs. (7) through (15) of [1]

P_d actual $(x_d; s)$ = measured discrete probability function of actual unfiltered noise being generated by the noise memory

Results showed the deviation to be less than 1 percent.

Figure 5 is the result of 8 histograms of the output of the memory when baseband-filtered with a 4-MHz filter

with the digital frequency response as shown in Fig. 6. Note that the set of integers now propagated through the hardware and the discrete probability function standard deviation will be

$$x_d \in \{-2047, ..., -1, 0, 1, 2, ..., 2047\}$$

$$s = 512$$

Also displayed in Fig. 5 is a fit of a theoretical Gaussian probability function of the same standard deviation. Note the excellent agreement of the fit with the actual filtered noise.

The analog responses of the digital filters used were in complete agreement with the expected digital frequency response (i.e., Fig. 6) when observed in the spectrum analyzer.

V. Repeatability

Several independent readings of the digital SSNR measurement of the baseband output were obtained as follows. The DTSG was configured to provide a given set of parameters, such as noise bandwidth (BWDTH), subcarrier frequency (Sc), symbol rate (Data), and the desired symbol signal-to-noise ratio (Nom.SSNR). The DTSG statistics monitor was obtained as the digital measurement of signal-to-noise ratio (Dig.SSNR) after a given integration time (Int.Time). The standard deviation (Std.Dev.) of a given number of measurements (Number of points) was also computed. Table 1 shows a summary of those results for different baseband configurations.

When the analog output SSNR was measured, the results were very similar to those shown in Table 1 for the digital measurement. (This similarity is further shown in Fig. 11, where SERSSNR refers to symbol error rate SSNR.)

The above procedure was repeated for the IF configuration. The DTSG was configured to generate a given set of parameters such as IF noise bandwidth at a center frequency of 5 MHz, with a carrier SNR (Nom.P_c/N_0). Results obtained in the digital output have been summarized in Table 2. (Similarities between the digital and analog measurement for the carrier SNR measurement, P_c/N_0, may be verified in Fig. 12.)

Several different configurations were used to test the setting accuracy of the given configuration of the DTSG in baseband mode. Figure 7 shows the results for a typical configuration.

Different configurations of IF filters were tested in the DTSG IF mode. A typical setting-accuracy error curve can be found in Fig. 8.

VI. Channel Isolation

Channel isolation (of in-phase and quadrature channels) was measured in the analog output of the DTSG in the IF mode, with a 5-MHz IF carrier frequency being generated in each channel with a 90 deg phase difference and P_c/N_0 = 40 dB/Hz. Results show an isolation of at least 45 dB between channels.

VII. Baseband Testing

Several baseband configurations were tested at Goldstone's Signal Processing Center (SPC 10) by feeding the output of the DTSG configured in baseband mode through the baseband patch panel to telemetry groups 4 (TG4) and 5 (TG5) (see Fig. 9). Each demodulator synchronizer assembly (DSA) was configured with a medium loop bandwidth, and the mean SSNR (<SSNR>), mean SER (<SER>), and standard deviations were obtained every 30 sec. SER refers to symbol error rate, and CER refers to carrier error rate. A summary of the results obtained is shown in Table 3.

This summary shows the capabilities of the DTSG as an SSNR source to measure telemetry-equipment performance with a high degree of accuracy. Extensive testing of the different signal-processing configurations in use at the DSN can now be performed with substantial savings in test-preparation time. Also, existing theoretical models can now be confirmed with the required accuracy.[3]

VIII. IF Testing

During acceptance testing of the Pioneer 10 receiver development unit (PN10Rcvr) at Goldstone, the DTSG 5-MHz IF output was upconverted to 300 MHz in the PN10Rcvr test equipment and fed to the PN10Rcvr (see Fig. 9). The DTSG was configured in the IF mode, providing a carrier and the corresponding telemetry modulation. The results obtained from the PN10Rcvr carrier detection

[3] *Deep Space Network Flight Project Interface Design Handbook*, JPL 810-5, Rev. D (internal document), Jet Propulsion Laboratory, Pasadena, California, July 15, 1991.

and the demodulator synchronizer assembly (DSA) symbol detection are shown in Table 4. In that table, BBA refers to baseband assembly, and L.B. refers to receiver loop bandwidth. Note that in Fig. 9 the parameters estimated throughout the system are represented by the convention <x>.

Another PN10Rcvr test was run with the DTSG in the IF mode with the carrier channel only. The telemetry channel was switched off for this test. The results have been plotted in Fig. 10 (solid line) for a Pioneer 10 one-sided receiver loop bandwidth of 1 Hz.

The expected P_c/N_0 degradation may be expressed as follows. The phase error variance may be linearly approximated by [2, p. 60]

$$\sigma_\phi{}^2 = \frac{W_L}{2P_c/N_0} \tag{2}$$

where W_L is the two-sided loop bandwidth. Assuming the phase error probability distribution $p(\phi)$ to be Gaussian with

$$p(\phi) = \frac{1}{(\sigma_\phi\sqrt{2\pi})\ e^{-(\phi/\sigma_\phi)^2/2}} \tag{3}$$

the carrier SNR P_c/N_0 degradation may be approximated by

$$P_c/N_0 \text{ degradation (dB)} = 10\ \log\left[\int_{-\infty}^{\infty} \cos^2(\phi)\, p(\phi)\, d\phi\right] \tag{4}$$

$$= 10\ \log\left[\frac{\left(1+e^{-2\sigma_\phi{}^2}\right)}{2}\right] \tag{5}$$

Results of the above equation have been plotted (the dashed line) in Fig. 10 with a bias of 0.2 dB, assumed to be due to PN10Rcvr system degradation. Also, the results of a computer simulation (asterisks) of the PN10Rcvr closed phase lock loop have been represented in Fig. 10.

IX. Accuracy

The normal (or Gaussian) distribution obtained in the noise generation is used to confirm the accuracy of the SSNR output of the DTSG. Note that the deviation between the symbol SNR (SSNR) and the symbol error rate SNR (SERSNR) for each particular case of the digital and the analog measurements will be a direct indication of the accuracy obtained in the generation of the SSNR in those particular circumstances (noise with Gaussian distribution).

DTSG accuracy was tested in the baseband mode for different baseband filters, subcarrier frequencies, symbol rates, and symbol SNR's. Fig. 11 represents a typical result. The deviation observed for each measured point is probably affected by quantization errors present throughout the high-speed hardware. Unfortunately, because of the lack of time and resources at the present time, these deviations have not been quantified. For the baseband mode, if the SSNR's to be generated are above −15 dB, a worst case of −0.15 dB in the generator accuracy can be identified. From Fig. 11, the accuracy of the SSNR generation can be stated to be better than 0.05 dB for SSNR greater than −5 dB. Note that this lower limit corresponds to most usual baseband testing conditions.

The above conclusions are also supported by the results presented in Section VII. Note that for high symbol rates, the telemetry degradation measured was on the order of 0.01 dB. These results comply with the most optimistic expectations and with the predicted results described in [1].

For the DTSG IF mode, carrier SNR measurement (P_c/N_0) is performed by integration of the carrier detection process for a period of 1 msec. (The carrier has a frequency of 5 MHz.) Also, the number of 5-MHz half-cycles in error (180 deg out of phase) is counted, converted to the equivalent SNR, and plotted in Fig. 12 as carrier-error-rate SNR.

From Fig. 12 it may be deduced that the divergence of both carrier SNR measurements of the analog output follows the same general trend as that of the measurements of the digital carrier error rate SNR, suggesting therefore a common problem. This common point of error seems to be introduced in the statistics hardware, with several contributing factors that remain to be fully explained.

From the P_c/N_0 results obtained in Section VIII and shown in Fig. 10, an approximate 0.2 dB degradation when they are measured with the PN10Rcvr is apparent. Expected PN10Rcvr degradation measurements at the lower end of P_c/N_0 are in good agreement with the actual carrier-SNR DTSG digital measurements. A more precise test using the radio science receiver remains to be

performed. These tests are expected to help identify the cause of the DTSG inaccuracies observed in the measurements of the DTSG analog output.

X. Conclusions

DTSG accuracy and repeatability at baseband (baseband mode) have been verified through different tests with the baseband assembly (BBA) part of the SPC 10 telemetry system. Results confirm the capability of providing calibrated symbol signal-to-noise ratios (SSNR's) accurate to 0.1 dB over a wide range of DSN telemetry requirements.

DTSG precise and repeatable generation of carrier and telemetry signals (IF mode) at 300 MHz has been demonstrated, although more tests are necessary to confirm the expected 0.1-dB accuracy.

Acknowledgments

R. Labelle, P. Priest, and M. Thompson are commended for their contributions to several parts of this project. Operations personnel from Goldstone (SPC 10) are also commended for their efforts in supporting some of the system test results presented. Finally, P. Ritcher and T. Peng are thanked for their constructive comments and suggestions in the process of reviewing this article.

References

[1] B. O. Gutierrez-Luaces, "Digital Test Signal Generation: An Accurate SNR Calibration Approach for the DSN," *The Telecommunications and Data Acquisition Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 161–174, February 15, 1991.

[2] J. H. Yuen, ed., *Deep Space Telecommunications Systems Engineering*, JPL Publication 82-76, Jet Propulsion Laboratory, Pasadena, California, July 1982.

Table 1. Baseband repeatability.

BWDTH, MHz	Sc, kHz	Data, symbols/sec	Nom.SSNR, dB	Dig.SSNR, dB	Std.Dev., dB	Int.Time, sec	Number of points
4	360	43,200	3	2.94	0.007	30	18
4	360	230,400	3	2.90	0.002	30	18
4	960	537,600	0	−0.37	0.002	30	10
0.4	32	32	0	−0.23	0.14	60	10
0.4	32	32	5	5.39	0.14	60	10
0.4	32	32	11	11.57	0.05	60	7

Table 2. IF repeatability.

BWDTH, MHz	IF, MHz	Nom.P_c/N_0, dB/Hz	Dig.P_c/N_0, dB/Hz	Std.Dev., dB	Int.Time, sec	Number of points
±4	5	50	49.84	0.01	60	3
±1	5	30	30.00	0.01	60	3
±1	5	17	17.36	0.12	60	8
±0.15	5	40	40.38	0.01	30	11
±0.15	5	20	19.59	0.10	60	10
±0.15	5	15	14.85	0.20	30	10
±0.15	5	8	8.14	0.28	60	10

Table 3. DSA performance verification.

DTSG configuration				SPC 10 telemetry results		
BWDTH, MHz	Sc, kHz	Data, symbols/sec	Dig.SSNR,[a] dB	TG4 SSNR, dB	TG5 SSNR, dB	Number of points
4	360	43,200	2.94	2.89 ± 0.10	2.90 ± 0.1	4
4	360	43,200	−0.06	−0.20 ± 0.15	− 0.16 ± 0.12	5
4	360	268,800	−0.08	−0.02 ± 0.05	− 0.14 ± 0.04	5
4	360	268,800	−3.12	−2.99 ± 0.06	− 3.02 ± 0.06	8
4	968	537,600	−3.36	−3.33 ± 0.08	− 3.36 ± 0.06	4
0.4	22.5	160	2.87	2.46 ± 0.05	2.63 ± 0.05	18
0.4	32.8	32	10.68	—	10.02 ± 0.26	8
0.4	32.8	32	5.78	—	4.23 ± 0.20	4
0.4	32.8	32	3.09	—	1.25 ± 0.40	7

[a] Standard deviation of DTSG output smaller than 0.1 dB.

Table 4. Pioneer 10 receiver results.

DTSG configuration					PN10Rcvr and BBA results			
BWDTH, MHz	P_c/N_0,[a] dB/Hz	Sc, kHz	Data, symbol/sec	SSNR,[a] dB	L.B., Hz	P_c/N_0, dB/Hz	SSNR, dB	Number of points
±4	40.5	360	268,800	0.86	5	40.30 ± 0.25	0.60 ± 0.01	5
±4	32.5	360	43,200	1.10	5	32.20 ± 0.40	0.78 ± 0.01	11
±0.15	40.4	32.8	32	10.0	1	40.20 ± 0.32	8.54 ± 0.1	5
±0.15	13.0	32.8	32	3.0	0.25	12.60 ± 0.20	1.88 ± 0.25	7

[a] Standard deviation of DTSG output smaller than 0.1 dB.

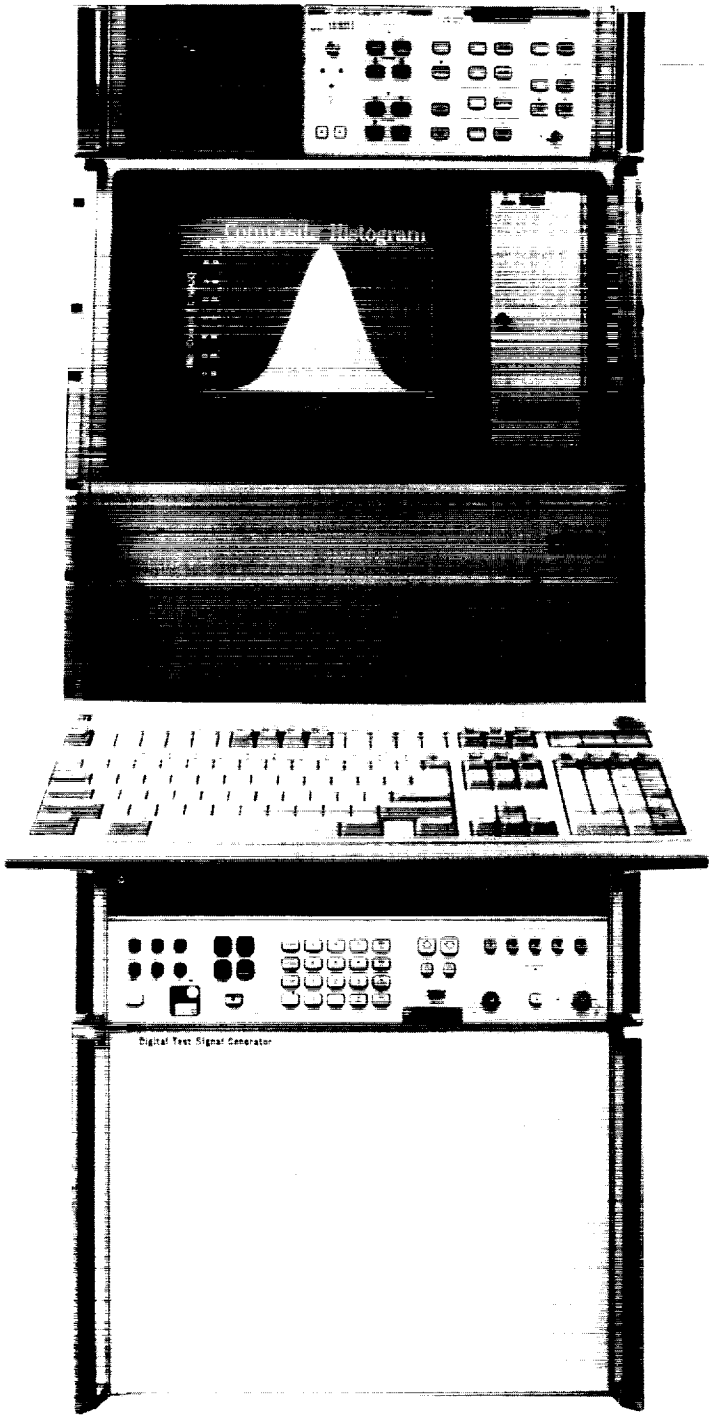


Fig. 1. DTSG rack-mounted main components.

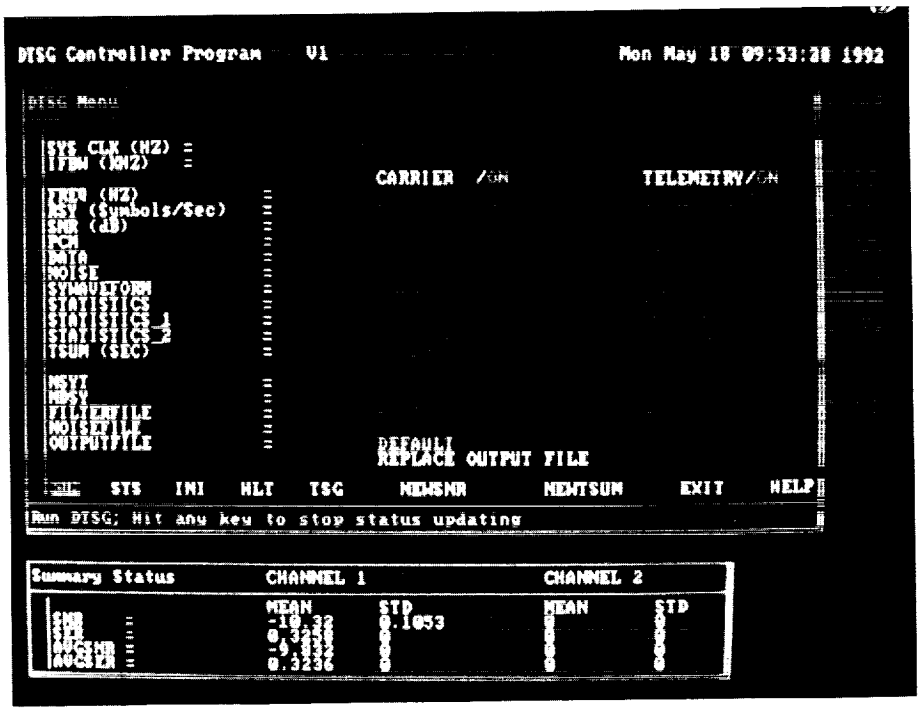


Fig. 2. User interface menu: DTSG baseband mode.

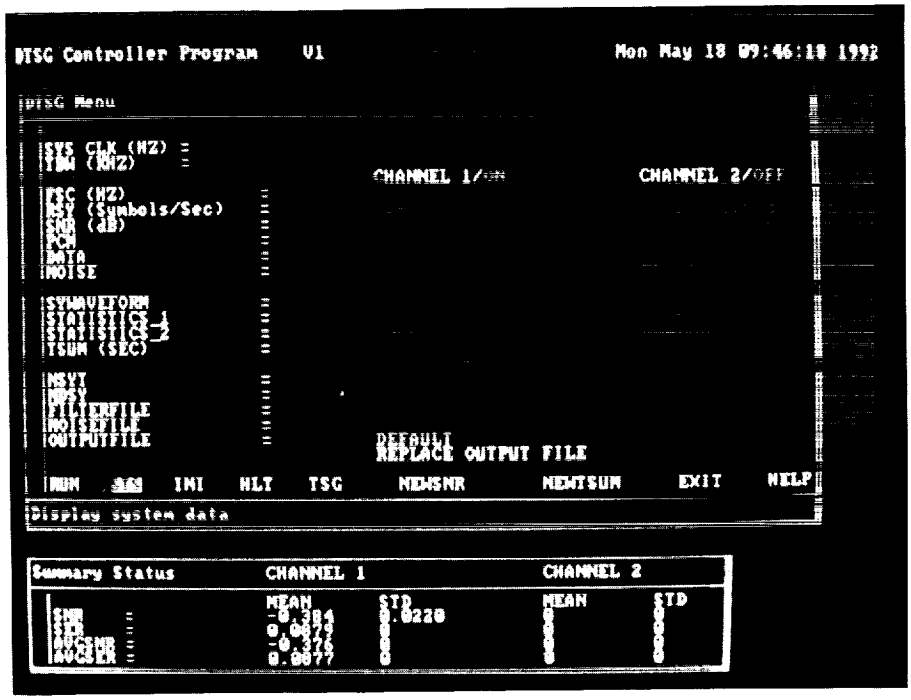


Fig. 3. User interface menu: DTSG IF mode.

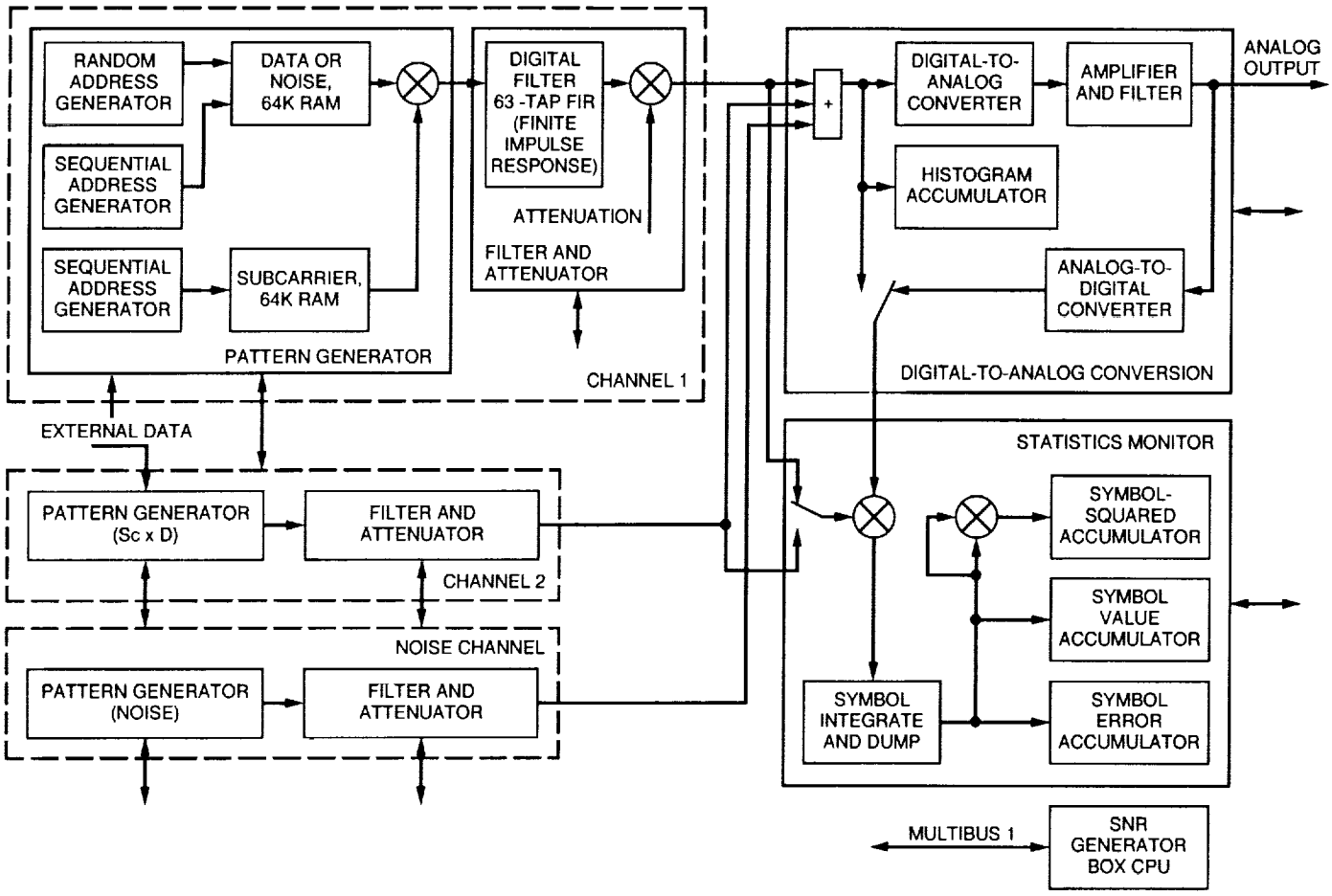


Fig. 4. DTSG hardware block diagram.

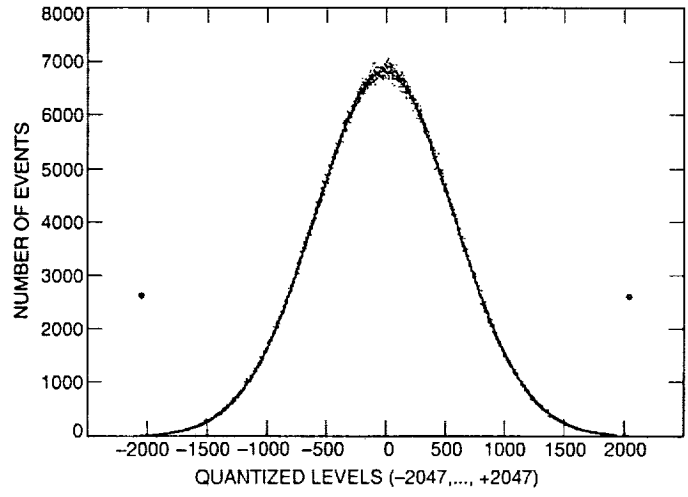


Fig. 5. A 4 M-Hz filtered discrete probability function with Gaussian fit (8×10^6 samples).

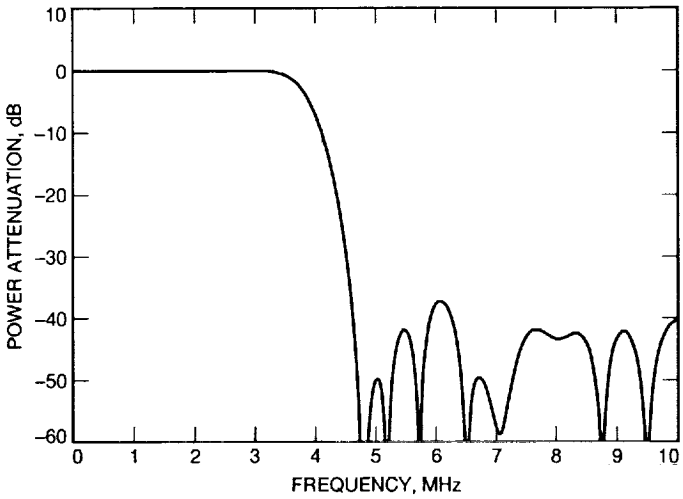


Fig. 6. Normalized frequency response of a 4-MHz digital filter.

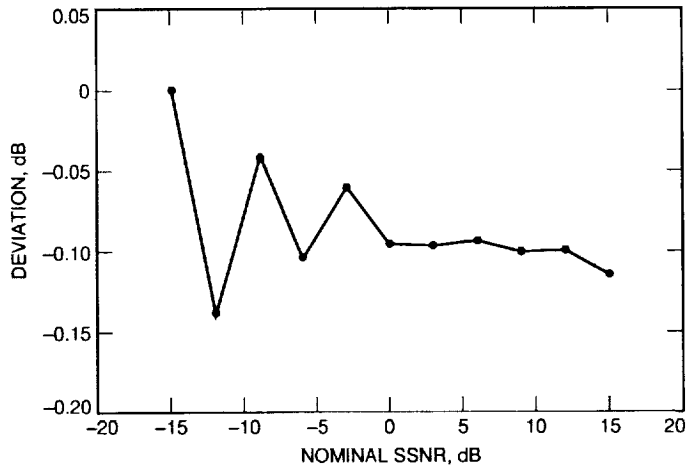


Fig. 7. Baseband SSNR setting accuracy (digital). Bandwidth = 4 MHz, subcarrier frequency = 354 kHz, symbol rate = 230.4K symbols/sec, and integration time = 30 sec.

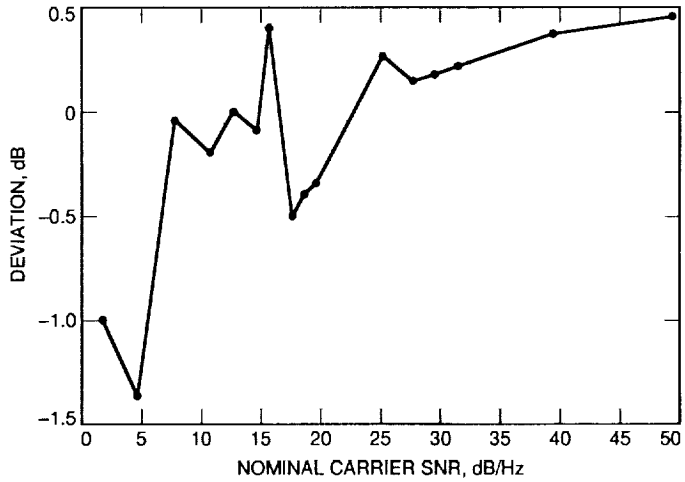


Fig. 8. IF carrier SNR (P_c/N_0) setting accuracy (digital). IF = 5 MHz, bandwidth = $\pm$150 kHz, and integration time = 0.1 msec.

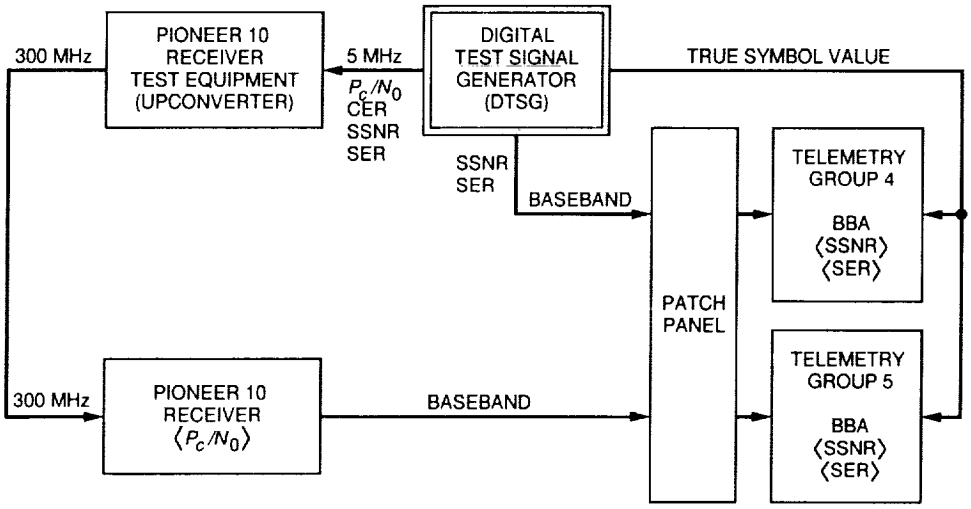


Fig. 9. SPC 10 testing configurations.

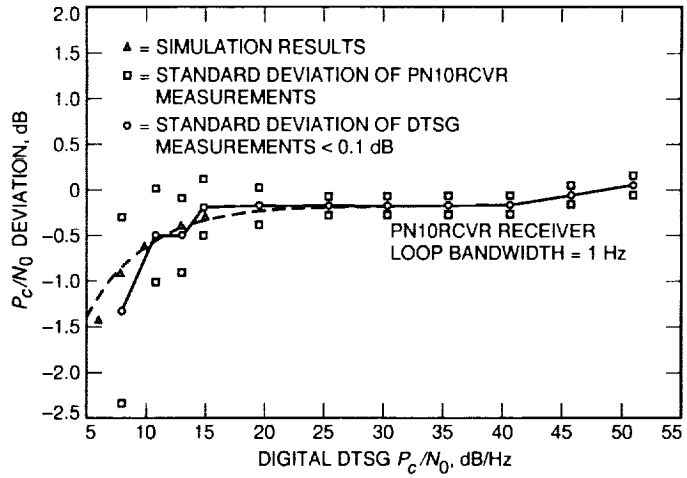


Fig. 10. Measurements of the 300-MHz PN10Rcvr carrier SNR (P_c/N_0). IF = 5 MHz and bandwidth = −150 kHz.

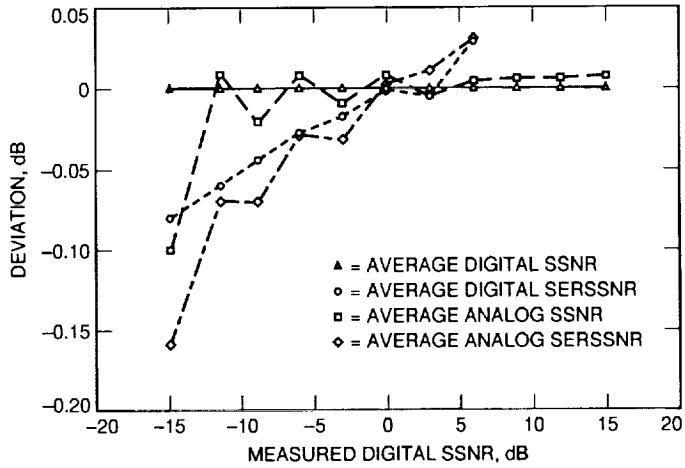


Fig. 11. Baseband SSNR measurements (σ = 0.01 dB). Bandwidth = 4 MHz, subcarrier frequency = 356 kHz, symbol rate = 43,200 symbols/sec, and integration time = 30 sec.

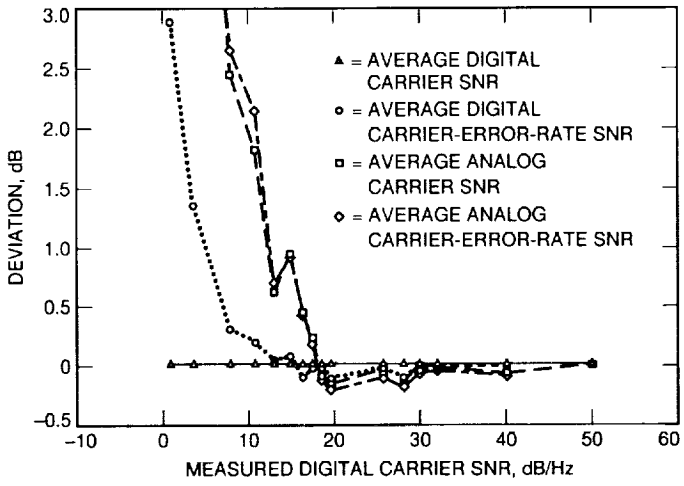


Fig. 12. IF carrier SNR (P_c/N_0) measurements (σ = 0.01 dB). IF = 5 MHz, bandwidth = ±150 kHz, and integration time = 0.1 msec.

Formal Functional Test Designs With a Test Representation Language

J. M. Hops
Radio Frequency and Microwave Subsystems Section

This article discusses the application of the category-partition method to the test design phase of hardware, software, or system test development. The method provides a formal framework for reducing the total number of possible test cases to a minimum logical subset for effective testing. An automatic tool and a formal language have been developed to implement the method and produce the specification of test cases.

I. Introduction

The focus of this article is on how the category-partition method, a method for specifying functional tests [4], can be applied to the test design phase of the testing life cycle, a required part of any DSN implementation task. Before describing the method itself, the requirements for the test design phase need to be clearly defined, as well as how they fit into the *JPL Software Management Standard*.[1] This discussion can be found below, in Section II.

Section III centers on some methods often used in the test design phase. The category-partition method is described in detail. Included in this section are the background of the method, a step-by-step description of how to implement it, and a demonstration of the method applied to a simple example.

The subsequent section, Section IV, introduces the Test Representation Language (TRL). TRL is a formal language for specifying test designs that have been created with the category-partition method and a computer tool for automatically generating test cases from the formal specification. The example from Section III is presented using the TRL format.

The conclusions in Section V provide some insight into the results that have been achieved and offer some suggestions for further study and data collection that may be necessary to assess the contribution of the TRL tool developed and the category-partition method used.

II. Problem Definition

A. Testing Life Cycle

In the field of software engineering, one is often faced with the challenge of creating an integrated, working system based on inadequate and meager requirements. The waterfall life cycle for software development, wherein requirements are systematically refined, architectural and detail design established, code written, and then the system tested, has become one of the accepted methods for dealing with the ambiguities and vagueness of the original requirements. The testing portion of this development life cycle, however, is not so clearly defined or widely accepted.

[1] *JPL Software Management Standards Package, Version 3.0*, JPL D-4000 (internal document), Jet Propulsion Laboratory, Pasadena, California, December 1988.

Figure 1 depicts the testing life cycle used in the development projects for the DSN. This life cycle is described in the *JPL Software Management Standard*[2] and is similar to the software development standard adopted by the Department of Defense, DOD-STD-2167A [2]. Table 1 defines the acronyms used in Fig. 1, along with the title of the document each acronym stands for, the phase in the testing life cycle during which the document is used or produced, and finally, whether a description of the document's contents is in the *JPL Software Management Standard*[3] and [2].

As noted in both Fig. 1 and Table 1, a key component is missing from the standards—a definition of the input, output, and purpose of the test design phase. This gap between the requirements for testing, produced in the test requirements analysis phase, and the detailed test procedures, produced in the test specification phase, is the phase during which the category-partition method can be most useful. The test design phase is explored in detail in the following section.

B. Test Design Phase

The dictionary definition of the word "design" is to conceive and to devise for a specific purpose. During the test design phase, the "specific purpose" that the test engineer is concerned with is meeting the test objectives and requirements determined in the test requirements analysis phase; what the test engineer is trying "to conceive and to devise" are the necessary and sufficient ways of validating the functional and performance requirements of the entire system. Therefore, the purpose of the test design phase is *to conceive and specify the environmental and system attributes that verify requirements and meet test objectives for each test requirement in the test plan and for each requirement in the functional and software requirements documents.*

Based on this definition of the purpose, the input to this phase is relatively simple to identify. It is

(1) The test objectives as documented in the subsystem integration and test plan and/or the software test plan

(2) The functional and performance requirements and system design as documented in the software specification documents or the functional and software requirements documents

(3) Any other pertinent design or requirements information that may be available, such as interface agreements and the preliminary software operator's manual

The output from the test design phase to the next phase, however, is not so easy to identify. The products to be developed are ways to validate requirements, which will be referred to as test designs. These designs are not expected to be test procedures specified to enough detail to be run by an operations engineer or possibly a quality assurance engineer; the test procedures written to that detail will eventually be written in the subsequent phase of the testing life cycle, the test specification phase. The test designs can have some ambiguity in the sequence of steps, the testing range of certain parameters, or the actual testing steps themselves.

Additionally, each test design should directly imply or specify a group of test cases. The test cases should have specific values for environmental and/or system parameters that have an effect on how the system under test will behave. Each of the test cases should also include the expected response or behavior of the system.

With this in mind, the output of the test design phase can be stated as follows:

(1) Test design specifications that

  (a) Are traceable to test objectives and functional and/or software requirements

  (b) Directly imply or specify a group of test cases that can be individually executed but share the same setup procedures

  (c) Identify the environmental and system features that are to be set or observed to control and determine the behavior of the system

  (d) Pass criteria for the group of test cases, and

(2) Test cases that specify

  (a) The environmental and/or system parameters and system states that should exist before the test case is executed

  (b) The test action or step to be taken to initiate the system behavior

  (c) The expected system behavior after the action has been taken

[2] Ibid.

[3] Ibid.

In the following sections, methods for determining the test designs and for automatically producing the documentation for the test cases are presented.

III. Method of Solution

A. Test Design Methods

There are many ways to create test designs that meet the needs of a certain project. Four of these methods are discussed below: the representative set method, the ad hoc method, the all-permutations method, and finally the category-partition method.

A common method for determining the number and contents of the test designs and test cases that should be transformed into test procedures is selecting a representative set of normal conditions and parameters that prove that the system works and meets requirements. On a project using this method, the emphasis will be on demonstrating that the system works rather than testing the system to detect failures, but the repeatability of the test procedures and the traceability to the requirements being tested is generally good.

On projects that are particularly short of time, money, and personnel, the test design phase is almost totally skipped. In this case, the test design method can be characterized as ad hoc. The ad hoc test case selection process is particularly prone to missing important aspects of the system behavior that could help determine where the problems are. The emphasis on a project using this method is almost always on getting the system out the door. Traceability to requirements is often poor. And most devastating of all, test repeatability is sacrificed; when a failure eventually occurs and the problem solved, it is very difficult to verify that the fix was correct because the conditions that caused the failure cannot be repeated.

Though not often seen, another method for selecting test designs and cases is a brute force method of analyzing all permutations of system parameter values. With this method, the test designs and cases are easily traced to requirements and test objectives, but it takes a lot of time and effort to analyze each permutation and decide which ones are valid and which ones are meaningless. This method allows the test engineer to find test cases that lie on the extreme boundary of the valid input space and therefore is good for error detection.

A recommended method for determining test designs is the category-partition method [4]. This method combines the benefits of choosing normal cases with the error exposing properties of the all-permutations method. Traceability can be maintained quite easily by creating a test design for each test objective in the test plan. By using an automatic tool to create the test cases based on the test design, the subsequent effort to transform the test cases into test procedures is simplified. The method allows the rapid elimination of undesired test cases from consideration and easy review of test designs by peer groups.

Section III.B discusses the category-partition method in general and is followed by Section IV, which presents the Test Representation Language (TRL) that can be used to implement the method and produce the test cases using the TRL tool.

B. Category-Partition Method

1. Background. The category-partition method was first presented by Ostrand and Balcer in 1988 [4]. A follow-on article in 1989 [1] discussed a test specification language and a tool for the automatic generation of test scripts that could be compiled and executed in the test environment that they had set up at Siemens Corporate Research. As pointed out in these two articles, the category-partition method is a way of analyzing the functional and software requirements of a system in order to determine test cases to be run. The method relies exclusively on the test engineers' reading of the requirements and design documents and their judgment of exactly which test cases should be selected for procedure development. If a formal requirements specification language is used to document the requirements and design, other methods may be more useful, such as the ones described in the article by Richardson et al. [5]. However, it is not often that the test engineer is presented with a functional requirements document or a software requirements document that is written this formally. Therefore, a structured method, such as the category-partition method, is needed to provide a systematic approach to developing test specifications from informal representations of the required system behavior.

The following sections discuss the steps in the category-partition method. The steps have been organized differently from the procedure discussed in the primary references, [4] and [1]. The organization of steps presented below has proven useful in communicating the method to the test engineers on JPL projects.

2. Steps in the Category-Partition Method. The category-partition method consists of the following four steps:

(1) Functional decomposition

(2) Category analysis

(3) Partition value analysis

(4) Partition constraint analysis

Each of these steps is discussed in the following sections.

a. Step 1: Functional Decomposition. The first step in the category-partition method is functional decomposition. The purpose of this step is to decompose the specification and/or requirements into functional units that can be tested independently. A secondary purpose of this step is to identify the parameters that affect the behavior of the system for each functional unit.

The requirement space is subdivided into subgroups, which may or may not overlap in some aspect. Each subgroup clearly identifies the requirements being tested and the input, output, and environmental parameters that affect how the system meets the requirements. The types of parameters that should be considered are user input, input from external interfaces, environmental input, output to another (observable) portion of the system, output to a user or external interface, output to the environment or state of the system, or maybe even the sequence of events. Note that there will be times when some of the parameters are not explicitly stated in the requirements specification, and therefore implicit parameters will have to be determined.

For an example, assume the following requirement specification has been decomposed from the requirement space: Sort an integer array either in ascending or descending order. The parameters mentioned explicitly in this requirements statement are *the array* and *an indication of sort order*. Implicitly, however, *the result of the sort operation* is also a parameter for this requirement.

The next step of the procedure is to further analyze the parameters identified and determine the characteristics, or categories, of the parameters that affect program or system execution.

b. Step 2: Category Analysis. The second step in the category-partition method is category analysis. The work done in the previous step, identifying functional units and explicit and implicit parameters, is carried further by determining the properties or subproperties of the parameters that would make the system behave in different ways. The test engineer should analyze the requirements and determine the features or categories of each parameter and how the system may behave if the category were to vary its value. If the parameter undergoing refinement were a data item, then categories of this data item may be any of its attributes, such as type, size, value, units, frequency of change, or source.

Choosing *the array* from the example in step 1 for further refinement, the categories that may be derived from the specification are *array size*, *the values in the array*, and, because the functional unit is a sorting function, *the arrangement of the values in the array*.

As can be seen, the original requirement statement said nothing about the valid range of *array size*. This step, along with the next one, tends to point out deficiencies in the requirements specification. The test engineer will have to work closely with the author of the requirements and the designers in order to resolve the ambiguities and uncertainties that surface from this analysis.

c. Step 3: Partition Value Analysis. After all the categories for the parameters of the functional unit have been determined, the next step is to partition each category's range space into mutually exclusive values that the category can assume. In choosing partition values, the focus should be on error-exposing values. The discussion on boundary value testing in Myers' book [3] and revealing subdomains in the article by Weyuker and Ostrand [6] should prove useful as references.

The partition values should include all possible kinds of values, especially the ones that will maximize error detection. Important values to look for are boundary values, extremal and nonextremal values, values that represent special cases or interactions, and valid and invalid values.

Returning to the example and using the category *array size* for illustration, the five partition values are

(1) 0

(2) 1

(3) 2 to the Upper bound minus 1

(4) Upper bound

(5) Greater than the Upper bound

It can be seen that *0* and *Greater than the Upper bound* represent error conditions that the sort function will have to process, while *1* and *Upper bound* represent special cases or boundary values. All the values between *2* and the *Upper bound minus 1* (inclusive) have been grouped together because the sorting function is expected to behave the same in this range; an error in processing that occurs for a particular value in this range should occur for all the

values in this range. It is left up to the test specification phase of the testing life cycle to determine the exact, or random, values that should be used to verify this partition in the test procedure.

The fact that two of the five values in this example have already been identified as being representative of error conditions gives one a head start on the next step of the category-partition method.

d. Step 4: Partition Constraint Analysis. The purpose of this final step is to refine the test design specification so that only the technically effective and economically feasible test cases are implied. There are three types of constraints defined in the category-partition method as described in [4]: errors, limits, and conditions.

An error constraint applied to a partition value is used to indicate that the partition value represents an exception state that the system under test should note and report without processing any further. Partition values of this type need to be tested in one test case but no more, due to the way exceptions are usually handled. Examples of partition values that should have error constraints are *0* and *Greater than the Upper bound* in the category of *array size.*

A limit constraint is for limiting the number of times a partition value will be used in the resulting test cases. Limit constraints can be applied to a test design in order to control the actual number of test cases implied. When economic feasibility, as in restricted time and resources, is a factor in the test execution, the limit constraint will help the test engineer to eliminate some of the test cases that seem redundant. In the above example, the test engineer may want to limit the number of times that an *array size* of *Upper bound* is used.

The remaining type of constraint is the conditional constraint. Determining these types of constraints is where the majority of the intellectual effort is spent. This part of the analysis specifies which partition values from one category can be used with the partition values of another category. Conditional constraints are specified in pairs: preconditions and postconditions. Preconditions are states or conditions that must co-occur for a particular partition value to be used in a test case; postconditions are the states or conditions that are set when a partition value is used. To illustrate their use, a slightly more involved example is discussed.

Starting with the category of *array size* and the partitions determined in the previous step, the types of conditions that are expressed by each partition value are analyzed. It can be seen that the values represent three separate conditions:

(1) "Error occurs" (for partition values of *0* and *Greater than Upper bound*)

(2) "Size is normal" (for partition values of *2 to Upper bound minus 1* and for *Upper bound*)

(3) "Size represents a degenerate array" (for an array size of *1*)

Clearly, if everything else is set appropriately, the valid partition values of the category *result* will be dependent on these conditions. Assume the following four partition values were identified in step 3 for the *result* category: *error notification*, *array unchanged*, *array in ascending order*, and *array in descending order.* A precondition for the result *error notification* is that the postcondition "Error occurs" has been set. For the values of *array in ascending order* and *array in descending order*, the postcondition of "size is normal" must have been set before these values could be used in a valid test case. The *result* of *array unchanged* could possibly be a result of many conditions, one of which is that the *array size* is *1*, where the "size represents a degenerate array."

3. Example Application of Category-Partition Method. Table 2 provides the results of the method applied to the example that has been discussed throughout the previous sections of this article.

IV. Test Representation Language (TRL)

The TRL was developed to implement the category-partition method. When used during the test design phase of the testing life cycle, the TRL files will form concise and uniform representations of the test designs for the functional testing of the system.

The TRL tool that implements the TRL language processes the ASCII formatted TRL files and produces ASCII formatted result files that document the individual test cases implied by the test design. The TRL tool documents the description, categories, and partition values to be used in each test case as they were documented in the input file. Each TRL file is created and changed with an ASCII editor and therefore can be easily modified to adapt to changes in functional specifications. The resulting test cases can be used during engineering tests of the system under test to verify preliminary procedures and functions while work continues in the test specification phase on transforming the test cases into formal detailed test procedures.

The TRL tool was written in the C programming language and can be ported to any platform; the SUN/SPARC and DOS environments are the computer platforms on which it currently runs. This tool differs from the one described in [1] in that the TRL tool is a general permutation control language that can be used in any environment; the output of the TRL tool is ASCII files that can be used for documentation rather than an executable test script, as in [1].

A. TRL Language Definition

The TRL provides a way to describe many test cases with one TRL file. The language consists of 1 comment character, 11 key words, 2 field demarcation characters, a logical AND character, and a logical NOT character. The processing rules for the key words, comments, and fields appear in the following sections, and a summary of the Test Representation Language appears in the Appendix.

1. Special Characters. There are five special characters in the TRL character set.

(1) Comment character = asterisk (*)

(2) Start field character = open bracket ([)

(3) End field character = close bracket (])

(4) Logical AND character = comma (,)

(5) Logical NOT character = exclamation point (!)

The asterisk is for initiating a comment line, which is a line defined by the comment character appearing as the first non-white-space character on a line in the TRL file. The start field character and end field character are for specifying the beginning and ending of a partition constraint field. Partition constraint fields are discussed in Section IV.A.3.

The logical AND character and the logical NOT character are for specifying a logical relation inside a partition value constraint field that is used for setting conditional constraints.

2. Line Key Words. There are two types of key words in the TRL: line key words and field key words. To be recognized as valid, the line key word should be the first word on a line. These key words are used to initiate a description of the test designs (DESCRIPTION), indicate the beginning of the categories and partitions (PARAMETERS), indicate a certain type of category (TYPE), specify the name of a category (NAME), set error message text (MESSAGE), and indicate the start of the block that describes the partition value and constraints of each category (SAMPLES). The line key words are, respectively: DESCRIPTION, PARAMETERS, TYPE, NAME, MESSAGE, and SAMPLES.

3. Field Key Words. The field key words are used in the partition value constraint fields to either describe a partition value (LABEL) or to specify the constraints determined during step 4 of the category-partition method. The field key words for setting labels and constraints are: SET, IF, LIMIT, ERROR, and LABEL.

A partition value constraint field is associated with a particular partition value by its physical location in the partition value block. A line in this block consists of the partition value text followed by zero or more constraint fields. The constraint fields can extend beyond the physical line of the TRL file, but the partition value text cannot. Partition value constraint fields are started by the start field character ([) and ended by the end field character (]).

As previously mentioned, partition value text cannot start with the comment character (*) and cannot contain any start or end field characters, ([) or (]).

B. Example Application of Category-Partition Method with TRL

In this section, the same example from Section III.B.2 will be discussed, but this time TRL will be used. To avoid confusion, the procedures for creating a test design using TRL are referred to as stages, and the procedures for implementing the category-partition method are referred to as steps. These stages will be performed for each functional unit and/or test objective in the system under test.

1. TRL Stage 1: Unconstrained Representation. The first stage in the TRL procedure is to create an unconstrained representation of the test design. This is accomplished by performing the first three steps in the category-partition method.

(1) Step 1: functional decomposition (Section III.B.2.a)

(2) Step 2: category analysis (Section III.B.2.b)

(3) Step 3: partition value analysis (Section III.B.2.c)

As for creating a TRL file, the following TRL key words and information should be created:

(1) DESCRIPTION key word and the description block. Create a description block that contains the requirements to be tested, the pass criteria to be used, and any other information pertinent to the test design.

(2) PARAMETERS key word. Start the parameter specification block.

(3) TYPE key words and NAME key words. For each type of parameter and category identified in step 2 of the category-partition method, create a TYPE and NAME specification in the TRL file.

(4) SAMPLES key words and the partition values. For each category, add in the unconstrained partition values that the category can assume during a test.

Note that the example in Fig. 2 with the unconstrained representation would produce 1440 test cases.

2. TRL Stage 2: Error Constrained Representation. The second stage of this process is to add in the error indicators and the message descriptions. This corresponds to a portion of the fourth step, partition constraint analysis, in the category-partition method.

The following key words and information should be added to the TRL file:

(1) ERROR field key words. For each partition value that should raise an exception during testing, create an [ERROR] field and add it to the test design.

(2) MESSAGE key word and error message list block. For each ERROR field, make sure there is a corresponding error message in a message list block.

See the example for TRL stage 4 for an illustration. When the error indicators are added to the three partition values as indicated below, 651 test cases result (Table 3).

3. TRL Stage 3: Condition Constrained Representation. The third stage of test design creation using TRL is probably the most difficult and time consuming. Adding in the conditional statements to make sure that only the technically feasible combinations of partition values get produced in the resulting test cases often takes many iterations. Investigating exactly which combinations are valid when used together, and what the expected output of the system should be, can expose many inconsistencies and undocumented requirements.

This stage, similar to the previous one, corresponds to the fourth step in the category-partition method. The purpose of this stage is to determine the precondition and postcondition pairs that describe the behavior of the system under test.

To modify the existing TRL file so that the conditions are expressed, the SET and IF field key words must be added. There will be some occasions where the addition of "don't care" partition values, or even the addition of repeat partition values with different conditional fields attached, will be necessary in order to produce the optimum set of resulting test cases.

The following key words and information should be added to the TRL file:

(1) SET field key words and postconditions. For each partition value that should cause a postcondition to exist if it is used in a test case, create a postcondition value and append it to the inside of the [SET] field. Use a logical AND character (,) to separate multiple postconditions being set for the same partition value.

(2) IF field key words and preconditions. For each partition value that is valid only when combined with a particular partition value in another category, append the condition value to the inside of the [IF] field. Use a logical AND character (,) to separate multiple preconditions to be applied to the same partition value. A logical NOT character (!) in front of a condition expresses that a condition should NOT exist in order for the particular partition value to be used in a resulting test case.

Again, the reader should refer to the stage 4 discussion in Section IV.B.4 for an example that has preconditions and postconditions. Before the LIMIT fields are added to the TRL file in stage 4, the TRL results file contains 32 test cases, which together represent the complete functionality of the requirement being tested in this functional unit. The purpose of the fourth stage is to reduce the number of test cases even further so that testing of this functional unit takes less resources.

4. TRL Stage 4: Limit Constraint Representation. This final stage of TRL file development produces the limit constrained representation of the test design. The purpose of the LIMIT field is to specify how many times a partition value can be used in the resulting set of test cases. Setting these limit values corresponds to the last step, or substep, of the category-partition method, where the remaining partition value constraints are determined.

Also included in this stage is the labeling of the partition values. The purpose of the labels is to provide the test engineer, who is performing the tests or transforming the test cases into detailed procedures, as much information about the test case as possible. The labels recommended are ones that describe the partition value in terms of its range, such as "normal," "low boundary," "high out-of-bounds," etc.

Therefore, the following key words and information should be added to the test design:

(1) LIMIT field key words. For each partition value that should only be used a certain number of times, n, in the resulting test cases, create a [LIMIT n] field. Note that partition values with an [ERROR] field are automatically limited to one test case.

(2) LABEL key words and label text. For some or all of the partition values in the TRL file, add a [LABEL label_text] field such that the "label_text" provides a description of the partition value that will be useful to the other test engineers.

The example given in Fig. 3 produces 24 test cases when processed by the TRL tool. Figure 4 gives an excerpt of the first two test cases from the resulting test cases produced by the TRL tool from the TRL test design documented in Fig. 3.

V. Conclusion

The purpose of the test design phase is to determine a set of technically feasible and resource-frugal test cases that meet the test objectives of the test plans and that verify the functional requirements of the system under test. The category-partition method can be used to determine test designs that meet this goal.

The Test Representation Language (TRL) and the TRL computer tool, used to process files written in the language, have proven very useful and efficient in implementing the category-partition method. For one task in particular at JPL, the Block V Receiver Task, the test cases that result from the TRL tool are being used to verify the system requirements in the engineering testing stage. Detailed test procedures are being developed based on the output of the tool. The TRL tool was also used on the Microwave Generic Controller Task to help develop the system and software acceptance test procedures.

As of yet, no objective data have been collected that can be used to compare the results of the testing process changes introduced by the use of the TRL tool. However, the qualitative feedback received from both test engineers and software designers is that the category-partition method and the TRL tool help them engineer tests rather than just perform tests. The effects of the method and the tool may be hard to quantify on an ongoing project. A way could be found to determine these effects if a small, controlled case study were to be initiated where two groups perform the same job—one using TRL and the category-partition method and the other using neither.

Work is continuing on enhancing the TRL tool to meet the needs of the test engineers using it. Some key words are being added to allow some very fine-tuned control over which test cases get included in the results.

In summary, the purpose and requirements of the test design phase of the testing life cycle have been explored and defined. The category-partition method and the TRL tool are efficient ways to produce the test designs and resulting test cases needed as input to the following phase of the testing life cycle. The Test Representation Language and the TRL tool can be of use to the test engineer or programmer no matter what level of testing is being performed. More effort in gathering the necessary metrics would be useful to be able to quantify the benefits received from implementing this process. If qualitative results are enough, however, most organizations could profit from an implementation similar to the TRL tool and the category-partition method for bridging the gap between test requirements and test specifications.

References

[1] M. J. Balcer, W. M. Hasling, and T. J. Ostrand, "Automatic Generation of Test Scripts From Formal Test Specifications," *SIGSOFT Software Engineering Notes*, vol. 14, no. 8, pp. 210–218, December 1989.

[2] *Military Standard Defense System Software Development*, DOD-STD-2167A, Washington, DC: U.S. Government Printing Office, February 1988.

[3] G. J. Myers, *The Art of Software Testing*, Wiley Series in Business and Data Processing, New York: John Wiley and Sons, 1979.

[4] T. J. Ostrand and M. J. Balcer, "The Category-Partition Method for Specifying and Generating Functional Tests," *Communications of the ACM*, vol. 31, no. 6, pp. 676–686, June 1988.

[5] D. J. Richardson, O. O'Malley, and C. Tittle, "Approaches to Specification-Based Testing," *SIGSOFT Software Engineering Notes*, vol. 14, no. 8, pp. 86–96, December 1989.

[6] E. J. Weyuker and T. J. Ostrand, "Theories of Program Testing and the Application of Revealing Subdomains," *IEEE Transactions on Software Engineering*, vol. SE-6, no. 3, pp. 236–246, May 1980.

Appendix

Test Representation Language (TRL) Summary

Character or Key Word	Purpose and/or Usage
*	Indicates a comment line.
DESCRIPTION	Indicates the start of a description block that will be included in test cases.
PARAMETERS	Indicates the beginning of parameter specifications.
NAME	Specifies the name of a parameter or category.
TYPE	Indicates the type of category.
SAMPLES	Indicates the beginning of a samples block defining the partition values and constraints. Comma (,) is used for logical AND; exclamation point (!) for logical NOT.
[	Beginning of sample value constraint field.
]	End of sample value constraint field.
IF	Field identifier indicating that postcondition constraints are listed in the current field. Comma (,) is used for logical AND; exclamation point (!) for logical NOT.
LIMIT *m*	Field identifier indicating that the number of test cases involving this partition value should be limited to *m*. If *m* is unspecified, the limit is one test case.
LABEL	Field identifier indicating that the specified label should be listed for this partition value.
ERROR *n*	Field identifier indicating that the sample value is an error exit. The error can be specified using the optional *n*.
MESSAGE *n*	Indicates that a message block follows corresponding to the errors in the partition values. The message number can be specified using the optional *n*.
Command line options	For performing "count only" (-c), writing results into separate files (-s), including preconditions/postconditions in output (-p), and including the partition label text in the output (-l).

Table 1. Testing life cycle input and output.

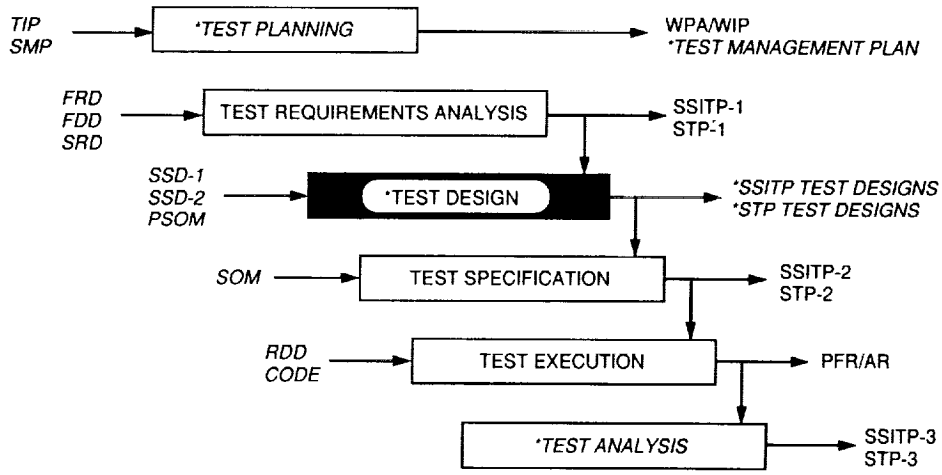
Acronym	Document title	Testing life cycle phase	Sufficiently described in standards?
D-4000	JPL Software Management Standard	All	Yes
TIP	Task implementation plan	Test planning	Yes
SMP	Software management plan	Test planning	Yes
WPA/WIP	Work package agreement/work implementation plan	Test planning	Yes
Test management	Test management plan for defining procedures of complete testing cycle	Test planning	No
FRD	Functional requirements document	Test requirements analysis	Yes
FDD	Functional design document	Test requirements analysis	Yes
SRD	Software requirements document	Test requirements analysis	Yes
SSITP-1	Subsystem integration and test plan-1, requirements	Test requirements analysis	Yes
STP-1	Software test plan-1, requirements	Test requirements analysis	Yes
SSD-1	Software specification document-1, architecture	Test design	Yes
SSD-2	Software specification document-2, detail design	Test design	Yes
PSOM	Preliminary software operator's manual	Test design	Yes
SSITP test designs	Subsystem integration and test plan—test designs	Test design	No
STP test designs	Software test plan—test designs	Test design	No
SOM	Software operator's manual	Test specification	Yes
SSITP-2	Subsystem integration and test plan-2, procedures	Test specification	Yes
STP-2	Software test plan-2, procedures	Test specification	Yes
RDD	Release description document	Test execution	Yes
PFR/AR	Problem failure report/ anomaly report	Test execution	Yes
SSITP-3	Subsystem integration and test plan-3, report	Test analysis	No (phase) Yes (report)
STP-3	Software test plan-3, report	Test analysis	No (phase) Yes (report)

Table 2. Example of formal functional test design: application of category-partition method (functional unit: sort an integer array either in ascending or descending order).

Categories	Partition values	Partition constraints: postconditions	Partition constraints: preconditions
Array size	0 (array unspecified)	SET "error occurs"	
	1 (degenerate array)	SET "size represents a degenerate array"	
	2 to upper bound minus 1	SET "size is normal"	
	Upper bound	SET "size is normal"	
	Greater than upper bound	SET "error occurs"	
Array values	All zero	SET "values identical"	IF "size is normal"
	All the same but nonzero	SET "values identical"	IF "size is normal"
	All negative values	SET "not identical"	IF "size is normal"
	All positive values	SET "not identical"	IF "size is normal"
	Mixed positive, negative, and zeros	SET "not identical"	IF "size is normal"
	Don't care	SET "values identical"	IF "error occurs" or IF "size represents a degenerate array"
Value arrangement	Minimum value before maximum value		IF "not identical"
	Maximum value before minimum value		IF "not identical"
	Don't care		IF "values identical" or IF "error occurs"
Sort order	Unspecified	SET "error occurs"	IF "size is normal"
	Ascending order	SET "ascending order"	IF "not identical"
	Descending order	SET "descending order"	IF "not identical"
	Don't care		IF "values identical" or IF "error occurs"
Result	Error notification		IF "error occurs"
	Array unchanged		IF "values identical"
	Array in ascending order		IF "ascending order"
	Array in descending order		IF "descending order"

Table 3. Stage 2 example of error indicators added to partition values.

Category	Partition value	Fields
Array size	0	[ERROR] ...
Array size	Greater than Upper bound	[ERROR] ...
Sort order	Unspecified	[ERROR] ...



LEGEND: *D-4000 INPUT* D-4000 PHASE OR OUTPUT **PHASE OR OUTPUT NOT DEFINED IN D-4000*

Fig. 1. Testing life cycle.

```
Example:
 *
 DESCRIPTION
 *
 Functional Unit:  Sort an integer array either in ascending or
                   descending order.
 *
 PARAMETERS
 *
 TYPE     Input-Categories for Parameter: Array

    NAME array size
    SAMPLES
         0
         1
         2 to Upper Bound minus 1
         Upper Bound
         greater than Upper Bound

    NAME array values
    SAMPLES
         all 0's
         all the same but not 0
         all negative
         all positive
         mixed +/-/0
         don't care

    NAME value arrangement
    SAMPLES
         minimum before maximum
         maximum before minimum
         don't care
 *
 TYPE     Input-Parameter: Sort Order

    NAME sort order
    SAMPLES
         ascending
         descending
         unspecified
         don't care
 *
 TYPE     Output to program or change in state

    NAME result
    SAMPLES
         error notification
         array unchanged
         array in ascending order
         array in descending order
 *
 * end of file
 *
```

Fig. 2. Stage 1 example of an unconstrained representation of a test design.

```
*************************************************************************
*        STAGE 4.  ADD [LIMIT ] AND [LABEL ] FIELDS to the TRL File     *
*************************************************************************
*
DESCRIPTION
*
        Test Representation for SORT requirement.
        File Name:       SORT.TRL
        Version:          1.5  Errors/Messages/Conditions/Limits/Labels
        Last Modified:   9/4/91
        Modified By:     J. Hops
*
PARAMETERS
        TYPE     Input-Categories for Parameter: Array
                NAME     array size
                SAMPLES
                *
                *  5 partitions
                *
                    0 (array unspecified)        [ERROR 1]
                                                  [ SET error, dont_care]
                                                  [ LABEL error condition ]
                    1 (degenerate array)          [ SET size_1, dont_care ]
                                                  [ LABEL degenerate array]
                    2 to Upper Bound minus 1      [ SET size_ok  ]
                                                  [ LABEL valid]
                    Upper Bound                   [ SET size_ok  ]
                                                  [ LABEL valid upper bound]
                    greater than Upper Bound      [ERROR 2]
                                                  [ SET error, dont_care]
                                                  [ LABEL invalid array size]
                *
                MESSAGE 1
                    Array size of 0 is invalid or array size is unspecified.
                    Array size is greater than the Upper Bound of sizes
                *
                NAME     array values
                SAMPLES
                *
                *  5 partitions, 1 don't care
                *
                   all 0's                        [IF size_ok]
                                                   [SET all_same, dont_care]
                    all the same but not 0        [IF size_ok]
                                                   [SET all_same, dont_care]
                    all negative                   [IF size_ok] [SET not_identical]
                                                  [LIMIT 4]
                    all positive                   [IF size_ok] [SET not_identical]
                                                  [LIMIT 4]
                    mixed +/-/0                    [IF size_ok] [SET not_identical]
                   don't care                     [IF !size_ok]
                *
                NAME      value arrangement
                SAMPLES
                *
                *      2 partitions, 1 don't care
                *
                     minimum before max    [IF    size_ok, not_identical]
                     maximum before min    [IF    size_ok, not_identical]
                     don't care            [IF    !not_identical]
                *
        TYPE     Input-Parameter: Sort Order
                NAME      sort order
                SAMPLES
                *
                *      3 partitions, 1 don't care
                     ascending                [IF size_ok, not_identical] [ SET ascend]
                     descending               [IF size_ok, not_identical] [ SET descend]
                     unspecified      [ERROR 3][IF size_ok]
                                               [SET error, dont_care]
                    don't care               [IF  dont_care]
                *
                MESSAGE
                    Sort order is not specified
                *
        TYPE     Output to program or change in state
                NAME     result
                SAMPLES
                *     4 partition values
                      error notification         [IF error]
                      array unchanged            [IF dont_care, !not_identical]
                      array in ascending order   [IF ascend, not_identical]
                      array in descending order  [IF descend, not_identical]
                *
```

Fig. 3. Stage 4 example of a TRL test design.

```
Description:

        Test Representation for SORT requirement.
        File Name:       SORT.TRL
        Version:          1.5  Errors/Messages/Conditions/Limits/Labels
        Last Modified:   9/4/91
        Modified By:     J. Hops


**********************************
Case #        1
Label:        1.6.3.4.1

PARAMETERS:

Type:   Input-Categories for Parameter: Array

        Category Name: array size
            Partition Value:      0 (array unspecified)
            Partition Label:      error condition
            Iteration number: 1

        Category Name: array values
            Partition Value:      don't care
            Partition Label:      instance value needed to pass error

        Category Name: value arrangement
            Partition Value:      don't care
            Partition Label:      instance value needed to pass error

Type:   Input-Parameter: Sort Order

        Category Name: sort order
            Partition Value:      don't care
            Partition Label:      instance value needed to pass error

Type:   Output to program or change in state

        Category Name: result
            Partition Value:      error notification
            Partition Label:      instance value needed to pass error

Error #1: Array size of 0 is invalid or array size is unspecified.


**********************************
Case #        2
Label:        2.6.3.4.2

PARAMETERS:

Type:   Input-Categories for Parameter: Array
        Category Name: array size
            Partition Value:      1 (degenerate array)
            Partition Label:      degenerate array

        Category Name: array values
            Partition Value:      don't care
            Partition Label:      valid

        Category Name: value arrangement
            Partition Value:      don't care
            Partition Label:      valid

Type:   Input-Parameter: Sort Order

        Category Name: sort order
            Partition Value:      don't care
            Partition Label:      valid

Type:   Output to program or change in state

        Category Name: result
            Partition Value:      array unchanged
            Partition Label:      valid

No error conditions exist.

**********************************
```

Fig. 4. Test case results of a stage 4 example of a TRL test design.

C-3

TDA Progress Report 42-114

August 15, 1993

PC4CAST—A Tool for DSN Load Forecasting and Capacity Planning

S. J. Loyola
Systems Analysis Section

Effectively planning the use and evolution of the DSN is a complex problem involving many parameters. This article discusses the tool that models many of these complexities, yet requires simple structured inputs and provides concise easy-to-understand metrics to aid in the planning process. The tool, PC4CAST, is used for both load forecasting (predicting how well planned that DSN resources meet expected demand) and as a decision support tool in the capacity-planning process (determining the relative benefits of capacity expansion options). It is now in use in the TDA Planning Office, has been used in numerous studies, and is also being used by the JPL Multimission Operations System Office (MOSO) as an integral part of Resource Allocation Team activities. Experience using the tool has helped to identify additional requirements that will further improve the planning process, which can be met by future PC4CAST versions.

I. Introduction

The DSN is a set of resources in high demand that requires careful planning to ensure its success in providing telecommunications and radio astronomy services to its community of users. Design and construction of additional DSN antennas take considerable time and money. For the DSN to remain responsive to its users, future resource capacity and capability (e.g., transmitter/receiver frequency) to satisfy expected user demand must be determined well in advance. When there is mismatch between resources and user requests, analysis is required to determine what additional resources are required and/or how the user demand can be modified [1]. Because of funding and schedule constraints, long-range planning typically involves a cost-benefit comparison of numerous capacity-expansion options.

Analysis to support this planning can be visualized along a spectrum. At one end is the amount of user demand that can be satisfied by a static view of DSN resources (i.e., the current set of antennas and any planned additions, upgrades, and removals). This approach is most applicable for evaluating the fit between the DSN evolution plan and the current set of mission requirements. At the other end is fixed user demand and the resource set (i.e., number, type, location, etc.) necessary to satisfy it, excluding issues of available funding and construction schedule constraints. The systems view, at the intersection of

these two views, considers both ends of the spectrum and provides trade-offs between user support and resource augmentation. For cost-effective planning of TDA resources, a planning tool must provide analysis appropriate to the problem.

The PC4CAST tool supports this spectrum of planning analysis in an integrated, consistent, accurate, and timely manner. The structure of analytical inputs (e.g., user requests, resource availability, and geometric constraints between user target and resources) is consistent for all types of analyses, while the type of output metrics may be chosen to fit the analysis being performed. The tool statistically models how user requirements would be accommodated without having to generate detailed schedules. PC4CAST is used both to forecast the DSN's ability to support any given mission set and to identify the DSN resources required for varying levels of mission support. A major capability of the tool is analysis of the marginal impact of new missions and new capacity.

Different approaches are required for planning the 34-m/70-m ground resources (which primarily support deep-space missions) than those required for 26-m and smaller resources (which typically support Earth-orbiting missions). The operational version of PC4CAST is designed to facilitate planning of the 34-m/70-m resources. A prototype PC4CAST is currently under development, which will model stochastic demand for 26-m and smaller resources. The balance of this article will focus on the operational capabilities of, and products from, PC4CAST for 34-m/70-m ground resources.

II. System Architecture

The PC4CAST tool consists of a spreadsheet-based user interface and a forecasting engine (see Fig. 1). Microsoft Excel version 4.0 is used for both input of problem-dependent parameters and graphical display of forecasting metrics. A forecasting engine, written in the C++ language, uses information entered into Excel, along with problem-independent information stored in databases, and provides the raw forecasting results. The Microsoft Windows environment seamlessly integrates Excel and the forecasting engine, so the separation of input/output functions from processing is invisible to the user and does not affect tool operation. This architecture takes advantage of the capabilities of standard commercial off-the-shelf products, allowing more development effort to be spent on design of the forecasting and capacity/capability planning algorithms.

Inputs and outputs are maintained in a type of Excel file called a workbook (a file which may contain multiple spreadsheets and charts) [2]. Each forecasting workbook represents one calendar year of a study. It contains spreadsheets for forecasting inputs, and both spreadsheets and charts for forecasting outputs. Inputs are entered by week, while results are output by week and month. The spreadsheet interface allows output formats to easily evolve with demand for different ways to view forecasting results. In addition, Excel users are able to design their own output formats to suit the needs of a particular study.

III. Inputs

Inputs to the PC4CAST tool can be grouped into three categories: user requirements, resource-capacity description, and view periods. User requirements are input in the same format that the JPL Multimission Operations Systems Office (MOSO) gathers and maintains them. This format is concise, yet of sufficient detail to allow accurate modeling of the user demand. Resource-capacity descriptions include each antenna's planned availability and any options for future capacity changes. View periods (times when a particular object is in view at a particular antenna) represent the intersection of the user and ground resource domains. The tool requires all mission view periods (typically one view period per station per day for deep-space objects) for the time period to be forecast.

User requirements are entered and stored in a spreadsheet within a forecasting workbook (Fig. 2). Requirements for all users of the resources must be input including those from spacecraft, ground-based scientific users such as Goldstone Solar System Radar (GSSR), High Resolution Microwave Survey (HRMS), and Radio Astronomy and Special Activities (RASA), and those for the DSN's own use (e.g., antenna calibration, maintenance). User requirements include the number of tracks or resource usages required each week and the following parameters: view period object (e.g., Cassini), user description (e.g., GSSR Mercury), usable resources, average and minimum durations for tracks in each week of this requirement, and pre- and post-calibration time, which includes antenna setup and tear-down overhead before and after each track. The set of resources that are able to satisfy the requirement is represented by the usable-resources parameter, described below.

The PC4CAST tool allows a requirement's usable resources to be stated in a variety of ways. Usable resources may be specified by antenna (e.g., DSS 14 or 14 for the 70-m antenna at Goldstone) or subnet (e.g., 70M, 34S,

34H). Antennas and/or subnets may be combined to represent antenna arrays, the simultaneous use of two or more antennas, or resource equivalents, when two or more resources can satisfy a requirement. For example, "70M, 34S/34H" represents a requirement that can use either a 70-m antenna or an array of 34S and 34H antennas.

Available resources are described by resource definitions, scenarios (evolution paths), and downtimes. Resource definitions simply define each resource by its code, location (i.e., which Deep Space Communications Complex), and subnet. This tool also provides network augmentation studies by maintaining multiple resource scenarios, each including the initial on-line and final off-line dates for each current and planned resource. The PC4CAST operator chooses the appropriate scenario based on the study to be performed. Resource downtimes represent extended periods of time when a particular resource is taken down. Regular antenna maintenance is treated as a user requirement, since this best models the scheduling environment.

IV. Outputs

The PC4CAST system has several different output metrics, each providing a different slice through detailed antenna-usage information. The calculation of *expected-usage profiles* is discussed in detail in Section V, but a quick description is necessary here. The software uses the input user requirements, resource scenario, and view periods to calculate a detailed view of the expected demand on each antenna throughout the study time period. Figure 3 is an example of the expected demand on one antenna for one week. The expected-usage profile provides the raw data from which many concise, useful metrics may be derived.

One primary metric used in forecasting studies is the amount of user community requirements that cannot be satisfied given the resource capacity defined by the input resource scenario and downtimes. In PC4CAST, this unserved user demand is labeled *lost time*. It is calculated as the hours that cannot be supported, and is usually output as a percentage of the total hours of user requirements on a per-subnet basis (Fig. 4). This is a convenient level of aggregation since each subnet usually has its own community of users. Lost-time information can be presented at varying levels of aggregation (e.g., the whole network) as required by the study being performed. Besides showing how the current resource implementation plan can support current user requirements, lost-time information may be used to compare the relative merits between two resource scenarios or the impact of adding, removing, or modifying a user's requirements.

Whereas the lost-time metric is from the users' point of view, the *load-duration curve* presents user demand from the ground resources' perspective. This metric has been successfully used in the electric power utilities industry for management decision making on cost-effective capacity evolution [3]. It shows the percentage of time that a resource can expect various levels of demand (load). The load-duration curve for the 70-m subnet for 1995 (Fig. 5) reveals that demand exceeds capacity for half of the year. Moreover, there is demand exceeding 200 percent of capacity for roughly 8 percent of the year (about one month).

The load-duration curve displays the full range of user demand relevant to capacity planning over the time period of interest. The y-axis intercept shows the value of peak expected usage, and the shape of the curve displays the proportion of time that demand is above, at, or below capacity. When the user load is above capacity (greater than 1.0), opportunities exist for adding capacity from DSN and non-DSN sources and/or managing user loads. When user loads are under 1.0, the ground system has the potential for accommodating other missions, if there is a favorable view period, or non-view period-constrained missions such as HRMS. Opportunities for, and benefits of, load management can also be quantified. Postponing a user's period of high demand so that it does not coincide with another user's can have a leveling effect by shifting load from the left side of the curve to the right. Reducing a user's required minimum track duration can have a similar effect on the shape of the curve. Reducing pre- and post-calibration times will usually reduce the height of the curve at all points. The impact of these and other assumptions, such as resource scenarios and mission sets, may be examined by comparing the shape of a load-duration curve generated with these inputs to a baseline curve.

Another useful metric, which is shaped somewhat like the load-duration curve, is the *resources-versus-lost-time graph*. This graph shows how the percentage of unsatisfiable requests varies with changing resource capacity. Conversely, it can be viewed as showing what resource capacity is necessary to satisfy a certain percentage of user requirements. This metric should be used with some caution relative to nonintegral subnet values since the location of antenna sites impacts the ability to satisfy user loads. For example, Fig. 6 suggests that there would be 31 percent lost time if only one subnet were available and 3 percent if there were two 70-m subnets. But, the implication that 1-2/3 subnets (i.e., 5 antennas) would result in 8 percent lost time is highly dependent upon antenna location and

the declinations of the spacecraft involved. In cases like this, the metric provides direction for further analysis of additional antenna-placement strategies.

The *trade space graph* displays the information contained in multiple resources-versus-lost-time graphs as the number of resources required for a few selected lost-time levels over multiple years. Figure 7 shows the number of 34-m subnets that is required to provide four different levels of lost time (0, 5, 10, and 20 percent). Since the absolute lost-time value is dependent on many factors, including how much filtering has been performed on the user requirements and how many missions are in prime phase, no one level of lost time can always be considered acceptable. Nevertheless, this graph does help to define the trade space where the cost of additional resources must be weighed against the cost of losing scientific data due to unsatisfied user requirements.

The PC4CAST system's library of metrics can be used to analyze problems in both load forecasting and capacity planning. The current status of user support can be measured with the lost-time metric. Load-duration curves can be used to assess the marginal benefits of adding resource capacity and/or capability. Cost-effective capacity evolution paths may be identified using resource-versus-lost-time and trade space graphs. Together, these metrics provide an internally consistent package to aid DSN management decision making.

V. Algorithms

The PC4CAST system uses a statistical approach that provides quick run time and models how user requirements would be scheduled, without actually generating schedules. Figure 8 is a process flow chart that gives a graphic presentation of the PC4CAST forecasting engine algorithm. This algorithm can be divided into two main phases. The first is the calculation of expected-usage profiles from forecasting inputs (user requirements, resource-capacity descriptions, and view periods). The second is the derivation of higher level metrics from expected-usage information. Presently, each week of a study year is calculated separately. Forecasting metrics for one year of inputs are calculated in around five minutes, running on a 486-based computer. The system design allows for relatively easy modification to other time granularity, should the need arise in the future.

A. Calculation of Expected-Usage Profiles

The calculation of expected-usage profiles is a three-stage process that incorporates user requirements, resource capacity, and view period information. As stated earlier, the goal of this process is to determine the expected level of demand for all time periods within the study time frame. The first stage is to calculate an *expected-usage value* for each requirement. This value represents a time-weighted distribution of the requirements over the view period length. Second, this expected usage value is used with the view periods and the required tracking overhead to generate individual expected-usage profiles. Each of these profiles represents the demand from that requirement for each point in time for each antenna. Third, all individual expected-usage profiles for each antenna are summed, resulting in one expected-usage profile for each antenna. These expected-usage profiles are the product of the first phase of the algorithm. Each of these three stages will be now discussed in detail.

The following describes the calculation of the expected-usage value for one requirement in one week; each requirement and each week are processed in a similar manner. First, the total amount of requested time is calculated from the requirement's average duration, pre- and post-calibration times, and the number of tracks requested in the week, as follows:

$$\text{requested time} = \text{number of tracks} \times (\text{average duration} + \text{precal.} + \text{postcal.})$$

Calibration times are included because they represent demand on the resources just as the actual track does. The next step is to find all *usable view periods* for the resources specified in the usable resource field that are long enough to support this requirement. In other words, the individual view period duration must not be less than the requirement's minimum duration. Then, *request slots* are defined as usable view periods with pre- and post-calibration appended to them. Request slots are defined this way because the time before and after each view period represents time when the calibrations could take place and still have the track within the view period. The *request slot time* is then calculated as the sum of all request slot durations.

$$\text{request slot time} = \sum_{\text{all request slots}} \text{duration (request slot)}$$

Finally, the expected-usage value is calculated as

$$\text{expected usage} = \frac{\text{requested time}}{\text{request slot time}}$$

The expected-usage value therefore represents the percentage expectation that when the requirement can use an antenna (constrained by the minimum duration constraint), it would use it or be scheduled to use it.

If the expected usage is greater than 1.0, the requirement cannot be supported by the resources it specified. This is caused by physical constraints such as positions of antennas and users, required view period mask, and overly restrictive requirements; it is not due to competition with other users. The system informs the analyst of any *infeasible requests* at the end of the attempted run. It is then the analyst's responsibility to modify the requirement until it is feasible. This can be accomplished by one or more of the following means: adding to the usable resources field; or reducing the minimum duration, average duration, precalibration, postcalibration, or number of tracks. Forecasting metrics will not be generated until all requirements are feasible.

The next stage is to generate individual expected-usage profiles for each requirement on each antenna requested. These are defined as stepwise constant functions, which can be visualized as the expected-usage value assigned to each request slot.

The final stage involves the creation of an expected-usage profile for each antenna from all individual expected-usage profiles for that antenna. Each individual profile is summed to a total expected demand for each antenna over the study period. Figure 9 graphically shows the generation of an expected-usage profile for a hypothetical mission set. This sample profile is used in the next section to help describe how PC4CAST high-level metrics are derived from expected-usage profiles.

B. High-Level Metric Derivation

Once the expected-usage profiles are calculated for all requested antennas, many high-level metrics may be derived from them. The derivation of the most frequently used metrics—lost time, load-duration curves, resources-versus-lost-time graphs, and the trade space graph—are discussed below.

Within PC4CAST, lost time is defined as the area of the expected-usage profile that is above one expected-usage unit (Fig. 10). Since the antennas that are being modeled can support only one user at a time, any expected usage over 1.0 represents time that cannot be supported. This is insupportable user demand or lost time. While lost time is calculated by antenna, the values for all antennas in a subnet are usually summed to give the lost time for the subnet. This subnet lost time is then charted as a percentage of the requested time on that subnet. Lost time is calculated on a per-week basis, but may be aggregated to any longer time periods, such as monthly or quarterly. Parenthetically, it should be noted that subnet requested time is simply the sum of the areas of expected-usage profiles for each antenna in the subnet. The lost-time metric captures the essential information contained in the expected-usage profiles and provides a concise measure of the impact of the input resource scenario upon user requirements.

Load-duration curves include all expected usage-level information, but aggregate the time-of-day information to a level more useful for capacity planning. An easy way to visualize a load-duration curve is as an expected-usage profile with the highest peaks sorted to the left and the lowest usage levels to the right (Fig. 11). The load-duration curve is generated for each antenna, but may be aggregated to the subnet level.

The resources-versus-lost-time graph displays lost time for a range of hypothetical resource capacities. The lost-time metric is usually calculated based on the fact that an antenna can support only one user simultaneously. But, for this graph, it is calculated with an assumed resource capacity ranging from zero to the highest expected-usage level (Fig. 12). As with the load-duration curve, these graphs can be used at the antenna or subnet level. This results in a graph that contains some immediately obvious data points. For instance, a resource capacity of zero results in 100 percent lost time; likewise, capacity equal to the highest expected usage results in no lost time (but very low capacity utilization). Also, the lost-time value for a capacity of one resource corresponds to the usual lost-time metric. The value provided by the resources-versus-lost-time graph is the identification of how addition or removal of ground resources will impact the users.

The trade space graph (Fig. 7) is derived from the resources-versus-lost-time graph for a few selected lost-time values. Each line on the graph corresponds to the resources required to provide for a lost time less than or equal to the selected level. The required-resources value is calculated by reading, on the resources-versus-lost-time graph, the number of subnets required for the selected lost-time level. This value is then rounded up to the nearest whole antenna (e.g., 1.5 subnets becomes 1-2/3 subnets, or 5 antennas).

VI. Future Potential

To date, PC4CAST has proven to be helpful in load forecasting and capacity-planning analysis. Experience

with the tool has identified requirements for further algorithmic enhancements which would provide improved modeling of the DSN system, more accurate metrics, greater insight into the distribution of lost time, and measures of augmentation cost effectiveness. The definition of a user requirement's usable resources could be extended to include required antenna size and additional equipment parameters such as transmitter and receiver frequencies. Lost time would be redefined to include not only requested time that is unsupportable because of competition, but also requests made infeasible by an antenna being down. Also, definition and implementation of a priority scheme would allow the lost-time and load-duration curve metrics to be disaggregated to the level of prioritized mission sets or even individual users. Resource augmentation costs and the value of user support would also be included to better support the decision-making process. These features would provide for improved analysis of the marginal benefits of resource evolution paths.

Earth-orbiting missions have not been modeled in this implementation of PC4CAST. Accurate calculation of the required view period knowledge for all users and time frames of interest is usually not a problem for the deep-space missions that normally request time on the 70- and 34-m antennas. The Earth orbiters that typically request the 26-m and smaller antennas have far less predictable view periods. The Systems Analysis Section has been developing probabilistic methods to model such users where view periods cannot be calculated very far into the future. This work has followed the general PC4CAST paradigm so the two approaches could be readily integrated to provide a comprehensive DSN load-forecasting and capacity-planning tool.

Acknowledgments

The initial PC4CAST prototype was funded by MOSO. The author thanks Gene Burke, Robertson Stevens, and J. R. Hall for providing PC4CAST metric development motivation and feedback at various stages; Fred McLaughlin for years of nearly continuous operational testing and feedback; David Werntz for his efforts in the development of the PC4CAST algorithms; Sil Zendejas for design and implementation of the underlying database structures; and Chet Borden for review of this project and the notion to apply load-duration curves to this domain.

References

[1] E. C. Posner, "Expected Antenna Utilization and Overload," *The Telecommunications and Data Acquisition Progress Report 42-107*, vol. July–September 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 104–112, November 15, 1991.

[2] *Microsoft Excel User's Guide (Version 4.0)*, Microsoft Corporation, Redmond, Washington, 1992.

[3] Decision Focus, Incorporated, *Costs and Benefits of Over/Under Capacity in Electric Power System Planning*, Report no. EA-927, Electric Power Research Institute Palo Alto, California, October 1978.

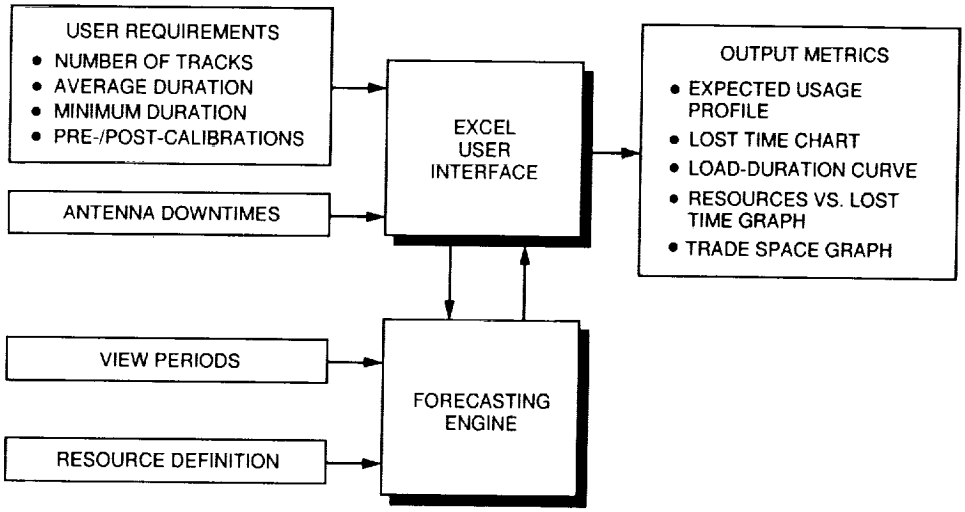


Fig. 1. Overview of the PC4CAST forecasting system.

USER LOADING PROFILE - 1997

VP			Durations		Calibration		January					February				March				April					May				June				July					August				September					October				November				December				
Object	User	Resource	Ave	Min	Pre	Post	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53
GLL	Galileo	63,14	7.0	7.0	1.50	0.25	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14				
GLL	Galileo Array	43/42/45/34	11.7	11.5	1.50	0.25	7	7	7	7	7	7	7	7	7	7	7	7	7	6	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7				
GLL	Galileo DOR	14/63,14/43	1.0	1.0	1.00	0.25	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2				
STRN	GSSR	14	8.0	8.0	1.50	0.50																																							2	4	4	5	4	4	2								
MERC	GSSR	14	5.9	5.9	1.50	0.50	3	2																2	3	2														2	3	2																	
91VK	GSSR Ast 1991VK	14	8.0	8.0	1.50	0.50		3																																																			
NONE	HRMS	43	8.0	8.0	0.75	0.25								1	1	1	1	1	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7																							
ICE	ICE	70M	4.0	4.0	0.75	0.25	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
DSS	Maintenance	70M	8.0	8.0			6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	
MO	Mars Observer	70M	8.0	8.0	1.25	0.25																																									10	14	14	14	14	14	14	14	14	14	14	2	
PATH	MESUR Pathfinder	70M	8.0	8.0	0.75	0.25																											1	7	7	7	7	3																					
PN10	Pioneer 10	70M	8.0	8.0	0.75	0.25	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	
NONE	Pioneer 6/7/8	70M	4.0	4.0	0.75	0.25											3													3													3												3			
NONE	Radio Astronomy	70M	8.0	8.0	1.50	0.50	2	3	3	4	3	5	4	3	6	3	3	3	3	3	3	4	5	4	3	3	6	3	3	3	2	3	3	4	3	5	4	3	4	6	3	3	3	3	3	3	4	5	4	6	3	4	3	3	3	3	3	3	
NONE	Radio Astronomy	43/42	8.0	8.0	1.50	0.50				1				1					1				1				1					1				1				1					1				1				1				1		
AGLL	Space VLBI	70M	6.5	6.5	1.00	0.50	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	
NONE	VLBI Clock Sync	14/63,14/43	3.0	3.0	1.50	0.50	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
VGR1	Voyager 1	70M	4.0	4.0	0.75	0.25	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	
VGR2	Voyager 2	70M	4.0	4.0	0.75	0.25	2	2	2	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	3	3	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2	
DSS	Maintenance	34S	6.0	6.0			3	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
ULYS	Ulysses	61	10.0	8.0	1.50	0.25	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
VGR1	Voyager 1	12,61	8.0	8.0	0.75	0.25	14	14	14	14																																													
VGR1	Voyager 1	61	8.0	8.0	0.75	0.25					7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
CAS	Cassini	34H	9.0	7.8	1.50	0.25																																									21	21	21	21	21	21	21	19	7	1	1	1	
DSS	Maintenance	34H	6.0	6.0			3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
MO	Mars Observer	34H	8.0	7.5	1.25	0.25	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	12	10	9	9	10	9	9	10	9	9	10	1	
PATH	MESUR Pathfinder	34H	8.0	8.0	0.75	0.25	21	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	3	3	3	3	3	21	7	7	7	7	3																					
NONE	Radio Astronomy	45	8.0	8.0	0.50	0.25					3													3													3													3									
NONE	VLBI CAT M & E	15/65,15/45	12.0	12.0	1.50	0.50					2	2					2	2					2	2					2	2					2	2					2	2					2	2					2	2					
VGR2	Voyager 2	34H	15.5	15.5	0.75	0.25	7	7	7	7	7	7	7	7	7	7	7	7	7	1																																							
DSS	Maintenance	34B2	6.0	6.0			2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
VGR1	Voyager 1	25	8.0	8.0	0.50	0.25					7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
NONE	HRMS	24	8.0	8.0	0.75	0.25	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	
NONE	Infrared Space Obs.	24	4.0	4.0	0.75	0.25	7																			7	7	7																															
ISO	Infrared Space Obs.	24	8.0	7.0	0.75	0.25				7	7	7	7	7	7	7	7	7	7	7	7	7	7								7	7	7	7	7	7	7	7	7	7	7	7	5																
ISO	Infrared Space Obs.	24	5.0	3.0	0.75	0.25		7	7															7	7				7	7																													
DSS	Maintenance	34B1	6.0	6.0			1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
VGR2	Voyager 2	34B1	15.5	15.5	0.75	0.25														6	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
5/28/93 11:29 AM							1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53

Fig. 2. User requirements input spreadsheet for 1997.

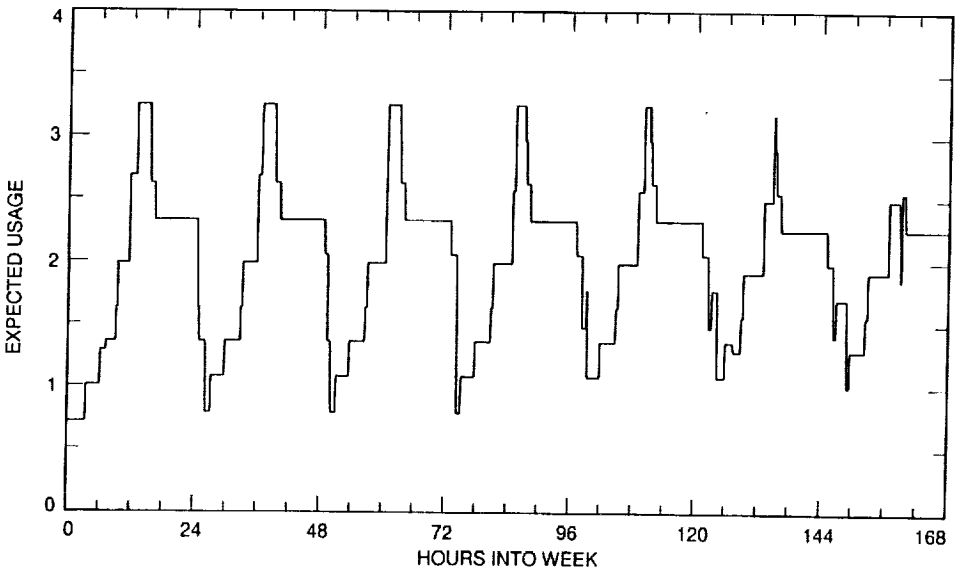


Fig. 3. DSS 12 expected-usage profile for week 18 in 1995.

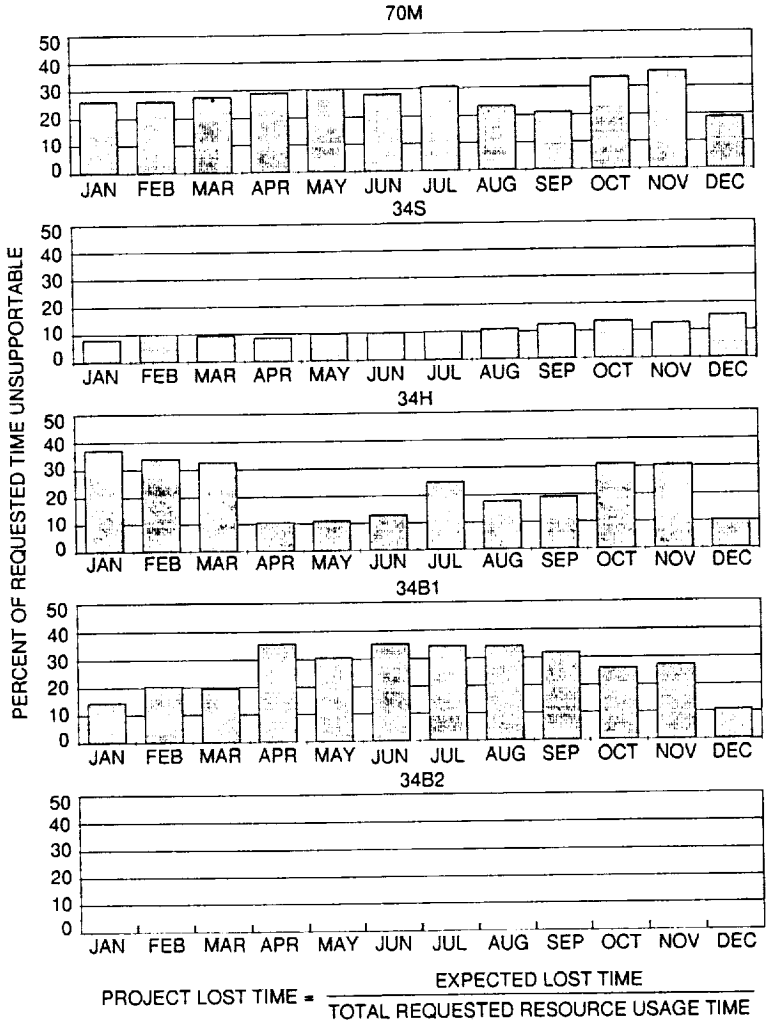


Fig. 4. Monthly lost-time chart for 1997.

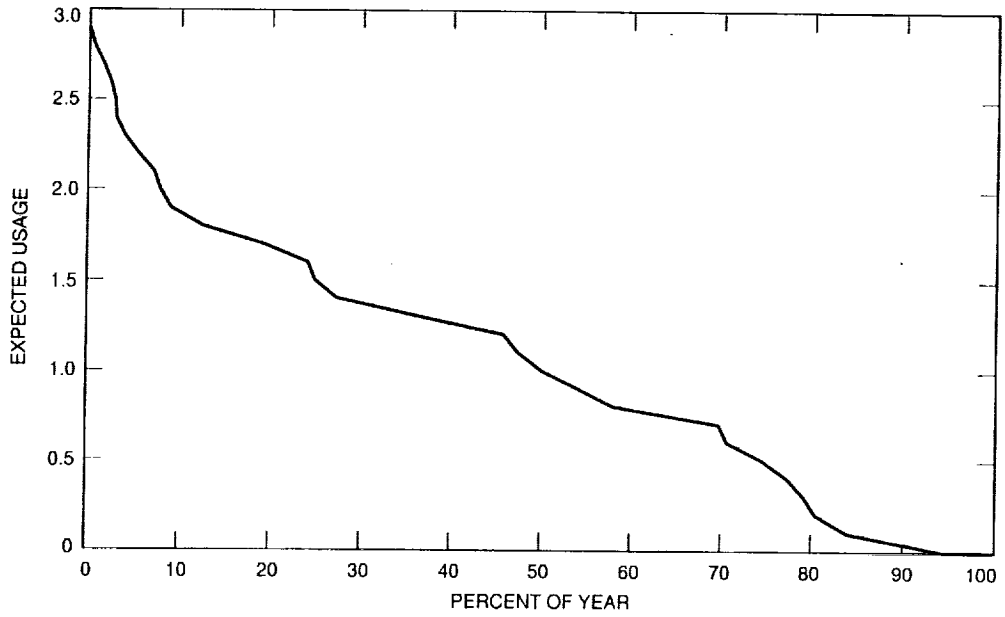


Fig. 5. Load-duration curve for the 70-m subnet in 1995.

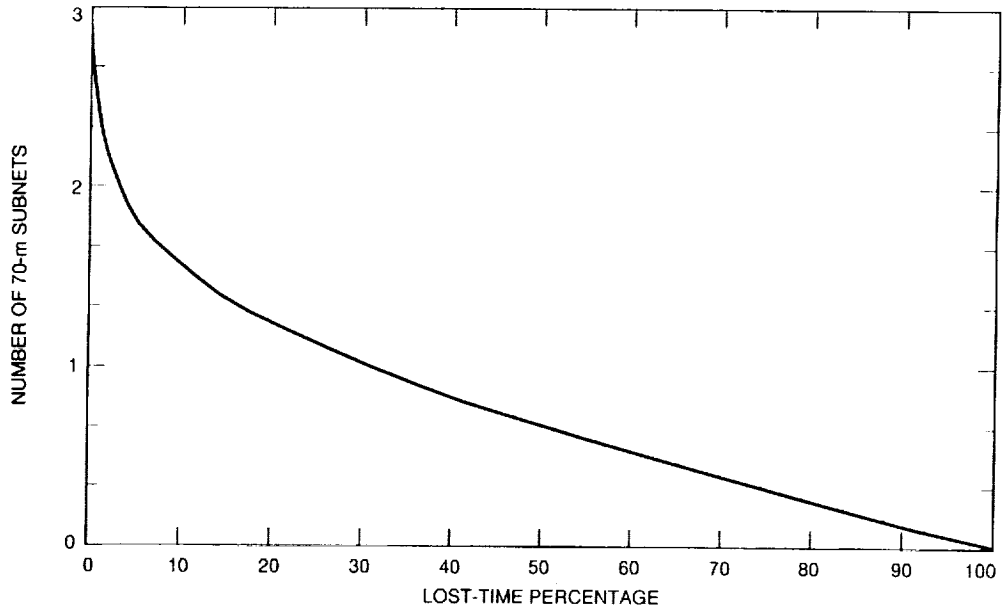


Fig. 6. Resources-versus-lost-time graph for the 70-m subnet in 1995.

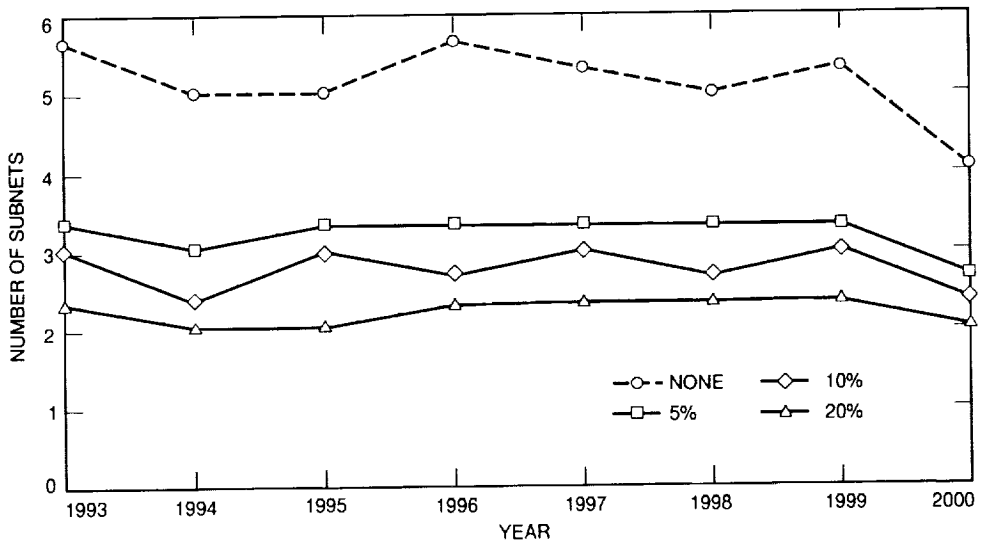


Fig. 7. Trade space graph for 34-m subnets for 1993 to 2000.

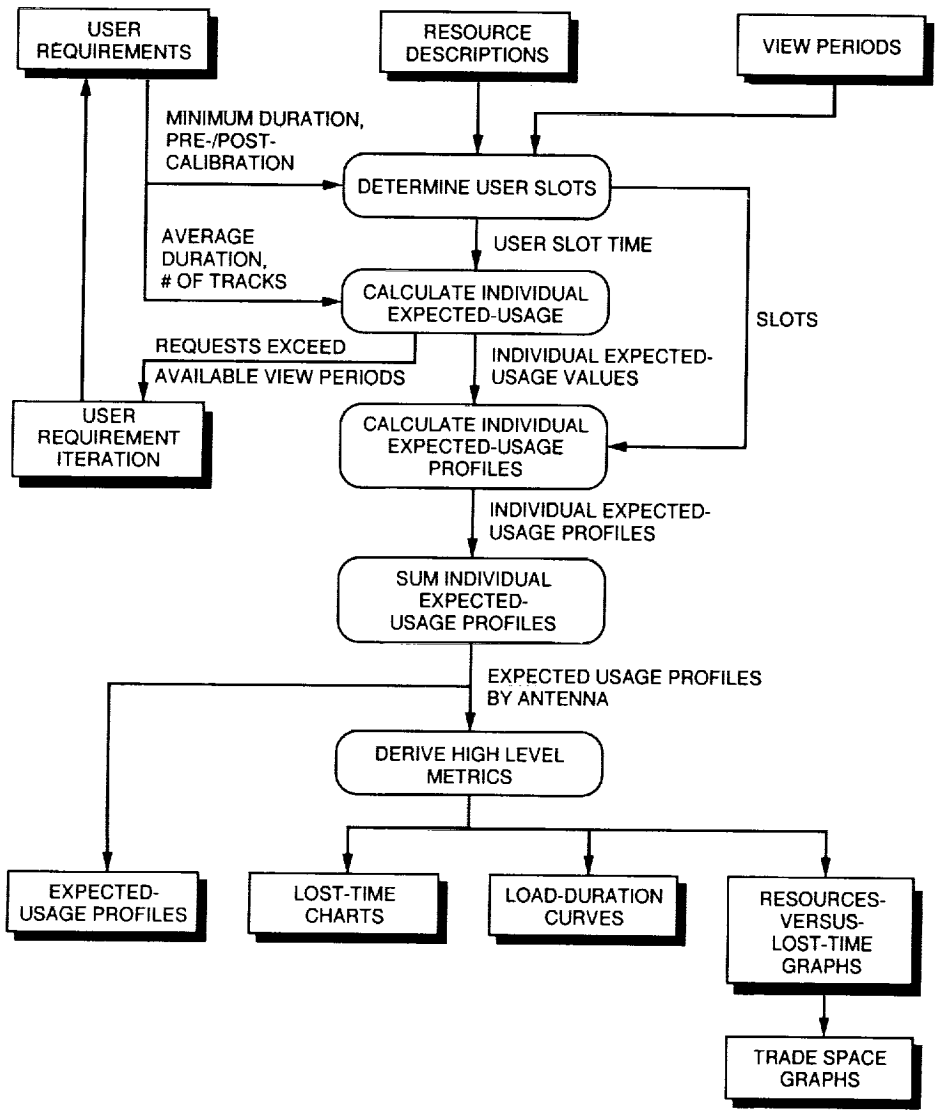


Fig. 8. PC4CAST forecasting engine process flow chart.

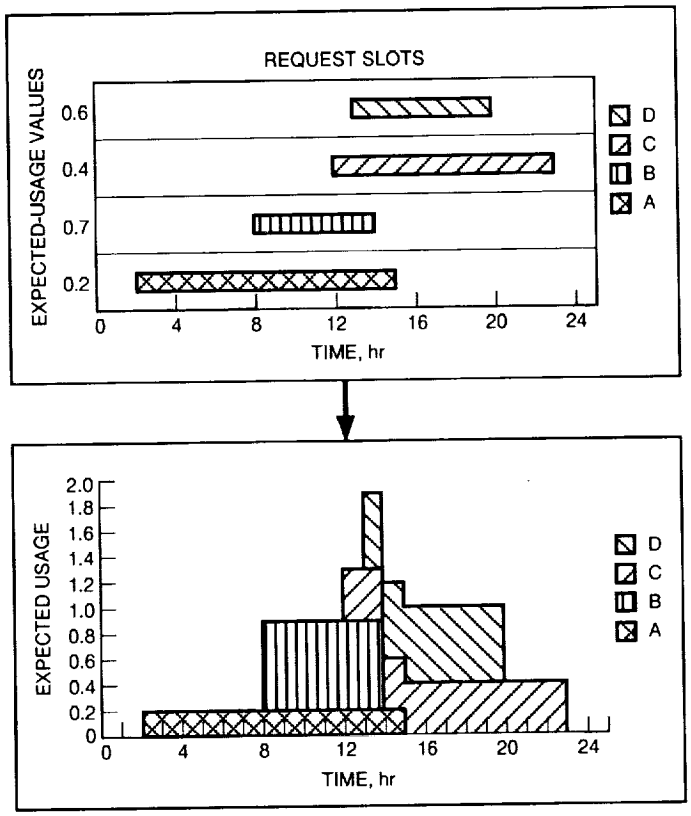


Fig. 9. Generation of an expected-usage profile using expected usage values and request slots for a hypothetical mission set (A, B, C, and D).

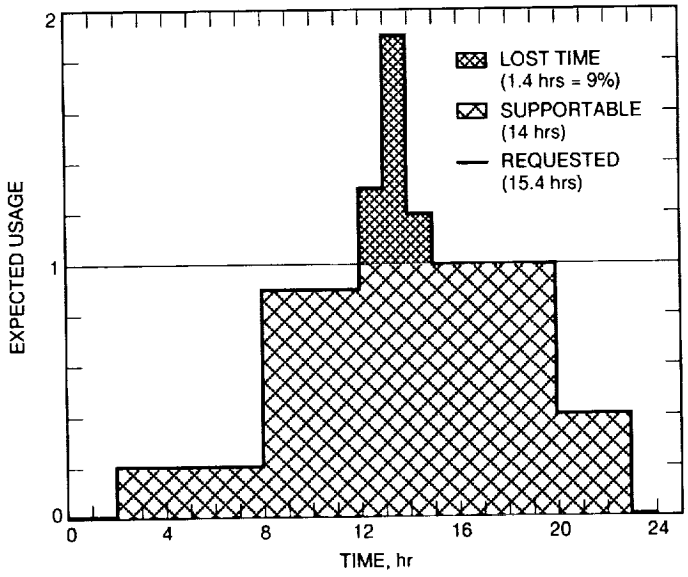


Fig. 10. Calculation of lost time from a hypothetical expected-usage profile.

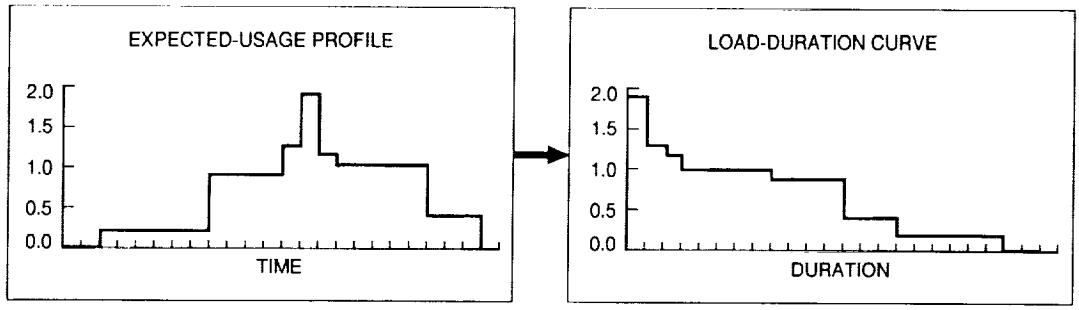


Fig. 11. Derivation of a load-duration curve from a hypothetical expected-usage profile.

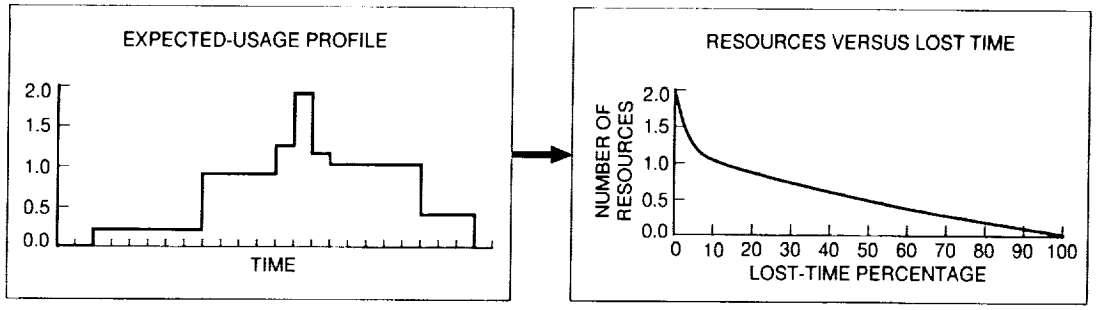


Fig. 12. Derivation of a resources-versus-lost time chart from a hypothetical expected-usage profile.

Asynchronous Transfer Mode Link Performance Over Ground Networks

E. T. Chow and R. W. Markley
Advanced Information Systems Section

This article reports the results of an experiment to determine the feasibility of using asynchronous transfer mode (ATM) technology to support advanced spacecraft missions that require high-rate ground communications and, in particular, full-motion video. Potential nodes in such a ground network include Deep Space Network antenna stations, the Jet Propulsion Laboratory, and a set of national and international end users.

The experiment simulated a lunar microrover, lunar lander, the DSN ground communications system, and distributed science users. The users were equipped with video-capable workstations. A key feature was an optical fiber link between two high-performance workstations equipped with ATM interfaces. Video was also transmitted through JPL's institutional network to a user 8 km from the experiment. Variations in video depending on the networks and computers were observed, and this article reports the results.

I. Introduction

The traditional role of the DSN has been to support deep-space robotic missions with relatively low data requirements (40 bps to 300 kbps). NASA is considering future missions that require high-rate (>1 Mbps) space communications and ground support from the DSN. These missions include a large number of low-cost Earth orbiters (0.1 to 10 Mbps), lunar missions with rovers and human exploration (10 to 100 Mbps), and Mars missions (1 to 10 Mbps). In order to support higher rate missions, the telemetry, command, and ground communications architecture must be significantly upgraded. This article reports on one potential upgrade to the ground communication system.

The most recent upgrades to the DSN include 2.4-Mbps telemetry processing and 865-kbps real-time ground communications, somewhat short of the requirements for future missions. The command uplink is limited to 2 kbps. The focus of this experiment was to evaluate an advanced ground communications architecture to support a lunar rover mission (Fig. 1).

The present ground communications architecture is designed to (1) transmit data directly to the user in real time, (2) buffer the data and meter it back slowly to the user (near-real time), and/or (3) store the data on transportable media for shipping (nonreal time). As an example, small amounts of data may be transmitted to users

in real time to support mission planning. Most telemetry data, such as data from the recently launched Mars Observer, are delivered in near-real time. Very-high-volume data, such as data generated by the Magellan Venus radar mapping mission, are handled in nonreal time in order to reduce communications costs.

In order to support a lunar rover mission, the DSN must support devices with radically different communications requirements than those for deep-space probes. A rover can be expected to simultaneously downlink video and telemetry, and receive commands from the rover operator on an uplink channel. It is highly likely that the operator, presumably a principal investigator, will want to be able to control the rover in real time, tolerating only a 1- to 2-sec propagation delay due to the distance of the Moon. The data rates for the video and telemetry could range from 1 to 10 Mbps, depending on the degree of video quality required. The command stream would be low (2 to 10 kbps), but with an additional requirement to be virtually real time with no ground communications delays.

One of the keys to success from the user's perspective in such a mission is an architecture that provides telepresence. Telepresence may require multiple displays to enable the rover operator to sense—at least visually—the immediate environment and take advantage of the rover's maneuverability in a scientifically interesting location.

A colony of microrovers would require several video channels for each microrover. Therefore, the scientists' workstations should be capable of managing multiple video sessions on one screen. In addition, real-time video conferencing among scientists would enable cooperative experiments.

The communications architecture should also support other services in addition to high-speed downlinks and real-time connectivity. For example, video, telemetry, and command data should be received with no detected errors. To reduce costs, the video channel should be compressed and decompressed with no loss of data (nonlossy). Greater circuit efficiency should be possible by dynamically allocating bandwidth to data and video in order to take advantage of lapses in transmission.

New network technologies enable high-speed data communications among local users and widely distributed users. For example, commercial fiber local area networks (LAN's) are capable of speeds of up to 100 Mbps. Traditional commercial services for wide-area networks are available at speeds up to 45 Mbps. Switched-based architectures based on the International Telegraph and Telephone Consultative Committee (CCITT)-standard asynchronous transfer mode (ATM) protocol will extend these rates into the range of gigabits per second. This switching technology, coupled with virtually unlimited bandwidth in optical fiber media, justifies consideration of a new approach to a ground network infrastructure to support space missions.

II. Technology

In 1990, the CCITT chose ATM as the standard switching technique for transport within the Broadband Integrated Services Digital Network (BISDN). BISDN is an emerging digital network planned by common carriers for integrating high-speed data, voice, and video services at transmission speeds starting at 155 Mbps. The following is a summary of ATM; a more complete, highly readable description may also be found in [1].

At the core of the ATM switching technique is the concept of cell switching. Cell switching refers to multiplexing of multiple logical data flows on a single physical interface, where the information from each source is in the form of fixed-size packets called cells. CCITT members, including both European and North American standards-setting groups, agreed on a standard ATM cell size of 53 eight-bit bytes. Five bytes are used for header information and 48 bytes carry the user information.

The ATM standard comprises three layers in the BISDN protocol stack. The BISDN stack is slightly different from the traditional Open System Interconnection (OSI) stack, but uses the same layering concepts. In the BISDN model, the three lower layers are occupied by ATM (Fig. 2). They correspond roughly to the bottom two layers of the OSI model (physical and data link layers). The three lower BISDN layers are (1) the physical layer, (2) the ATM layer, and (3) the ATM adaptation layer (AAL). The physical layer defines physical interfaces and framing protocols for the network. Physical connections in ATM are based on synchronous digital hierarchy (SDH) principles, but are not restricted to SDH framing protocols, and are thus very flexible. Because the functions are layered, the physical interface may be changed without impacting the higher layers. For example, the physical layer may be changed from Synchronous Optical Network (SONET) to T1 or T3 with no impact on the overlying ATM layer.

The ATM layer defines the information within each cell, and how cells flow over the logical connections in an

ATM network. Figure 3 shows the structure of those cells. Cells belonging to different virtual paths are distinguished by their virtual path identifier (VPI) and virtual channel identifier (VCI). Virtual paths can contain multiple virtual circuits. Each virtual circuit can be either permanently established or set up dynamically. Other fields in the 5-byte header of the ATM cell include a generic flow control (GFC) field, a payload type (PT), a cell loss priority (CLP) bit, and header error control (HEC) field. As mentioned earlier, the 48-byte information field carries user information.

While the AAL layer is a part of the overall communication process, it is not a network process, but is a process performed by the user's equipment. By keeping AAL functions separate, the ATM network becomes primarily concerned about delivering 48-byte payloads based on the information contained in the header. The adaptation layer plays a key role in making it possible for multiple types of services to use the ATM network. There are five types of services:

(1) Constant bit rate (CBR) services for isochronous traffic, such as voice and video.

(2) Variable bit rate (VBR) timing-sensitive services for features such as compressed video.

(3) Connection-oriented VBR Data Transfer for intermittent, large, long-period data file transfers over preestablished ATM connections. AAL 3 provides error detection.

(4) Connectionless VBR data transfer for short, bursty messages over non-preestablished ATM connections where call setup time could be longer than the transfer time.

(5) Simple and efficient adaptation layer (SEAL), for services similar to AAL 3 but with the assumption that the user's higher-layer processes provide error control.

The fundamental parts of the ATM standard have been defined by CCITT. The user network interface (UNI) (which incorporates the AAL layer) is being developed by the ATM Forum, a consortium of interested users and developers. The ATM Forum will also address switch-to-switch interfaces, and signaling specifications across the user interface, and establish network management specifications.

III. Experiment Configuration

The experiment simulated a lunar rover exploring the surface of the Moon (Fig. 1), with a downlink to science operations on the Earth. The major links in a full simulation would have included a lunar radio link between the rovers and a lander, a radio link from the Moon to the Earth's deep-space stations, overseas communications to JPL, and a link to science operations. The focus of this experiment was very limited; it was to determine the feasibility of using ATM networks on the Earth for transmitting data from the antenna stations to science operations. Therefore, design and prototyping of the space segment were postponed for future work.

Figure 4 illustrates the laboratory configuration. The rover and simulated lander were in a laboratory equipped with a rock-strewn floor. The principal investigator (PI) environment was in an adjacent room. Another investigator was in an office 8 km away. The experiment was run using the resources of the DSN Information Systems Engineering (ISE) Laboratory.

A. Microrover

The microrover was equipped with a miniature television camera, a radio transmitter, and a 2.4-kbps radio modem for transmitting telemetry and receiving commands.

B. Computers and Networks

The computer in the lunar environment (A, Fig. 4) was a Sun Sparcstation 10/41 rated at 96 million instructions per second (MIPS). The PI environment computer (B) was a Sparcstation 2 (24 MIPS). Each workstation had an Ethernet interface, ATM interface, and a video digitizer. The computers had the following capabilities:

(1) Video processing and compression were accomplished with Parallax video digitizers. Compression was based on the Joint Photographic Experts Group (JPEG) standard. Each board was capable of processing two video signals simultaneously; this enabled the science users to simultaneously oversee the rover and its immediate terrain.

(2) The Internet suite Transmission Control Protocol/Internet Protocol (TCP/IP) was the common protocol suite among both systems.

(3) The ATM interfaces were Fore Systems SBA-100 S-Bus ATM computer interfaces. The ATM physical layer was a 140-Mbps optical fiber interface. The user-level application interface was a Berkeley socket interface. Underlying software in the drivers utilized AAL 4 protocols, providing cell-level error detection (but no correction). If a bit error was detected by the driver, the cell was dropped and TCP performed error recovery.

(4) Computer A had an RS-232 interface to a radio modem for sending commands to and receiving instrument telemetry from the rover.

(5) The rover video was also transmitted from workstation A to workstation C, another Sparc 2, over an institutional network. Simple teleconferencing software provided by Parallax made this capability possible, and the software was integrated with other X Windows software to enable remote control of the rover. The path to C was a series of networks of varying types, most notably Ethernet and serial channels, which yielded low-rate, intermittent stop-action video.

(6) A voice/video teleconferencing demonstration was performed among users at workstations B and C using a commercial videoconferencing package called Communiqué. This was also low-rate video.

IV. Results

A. ATM Network Performance

The raw TCP/IP throughput across the ATM network was approximately 14.9 Mbps (11 percent of the bandwidth). This illustrates the limitations of the TCP/IP software implementation in the ATM board and Sun operating system. When video was generated by the Parallax board, the throughput dropped to 3.5 to 4.5 Mbps, depending on the complexity of the images.

Comparable data across a dedicated Ethernet connection was 8.5 Mbps (85 percent of the Ethernet bandwidth) for raw TCP/IP and 1.4 to 1.5 Mbps for video data.

Visually, the SPARC 10 transmitted 8 to 10 frames/sec (fair quality video) over the ATM network.

Two suggestions to improve the video and use the complete bandwidth of ATM and the associated optical fiber networks were (1) an alternate implementation to TCP/IP and (2) special ATM onboard processing hardware to reduce the central processing unit overhead in the workstations.

B. Compressed Video

JPEG compression can produce a good quality picture at a 20:1 to 30:1 compression ratio. At above 30:1 compression, blocking effects start to show up. Improved compression software may also make an improvement in the quality of the video.

C. ATM Versus Ethernet

This experiment illustrated the advantage of ATM over Ethernet for multimedia communication. The ATM full-motion video was relatively smooth and continuous. When video was sent from workstation A to C, some frames were delayed because the Ethernet was being shared with other users. This video was of poor quality—only 1 to 2 frames/sec. This type of performance has been noted elsewhere [2]. The isochronous nature of an ATM channel is very important for multimedia communications.

D. Teleoperation

It was demonstrated that potential teleoperation of NASA rovers by the general public is feasible. Numerous individuals, ranging from children to higher management personnel, were each able to learn how to drive the rover in less than 3 min. With the help of multimedia networking, it is very easy for anyone to do teleoperation.

E. Video Teleconferencing

Teleconferencing was attempted between a PI in the DSN ISE Laboratory and the remote investigator using commercial software. The video frame rate was low, typically 1 to 2 frames/sec, because of intervening institutional networks. For many teleconferences, such a low rate may be good enough. One technical problem was voice feedback. Two techniques were tried to improve the voice quality: (1) use of an echo canceller (this alone did not do the job) and (2) use of a unidirectional microphone. By combining (1) and (2), acceptable, but not really good, voice quality was achieved.

Acknowledgments

The authors wish to express their appreciation to Charles Goodhart and Julia George, who designed and built the software for this experiment. They also express sincere thanks to Rajiv Desai, who designed and built the microrover used in this demonstration; without his help and cooperation the demonstration would not have been possible.

References

[1] *Asynchronous Transfer Mode: Bandwidth for the Future*, Norwood, Massachusetts: Telco Systems, Inc., 1992.

[2] B. Cole, "The Technology Framework," *IEEE Spectrum*, vol. 30, no. 3, pp. 32–39, March 1993.

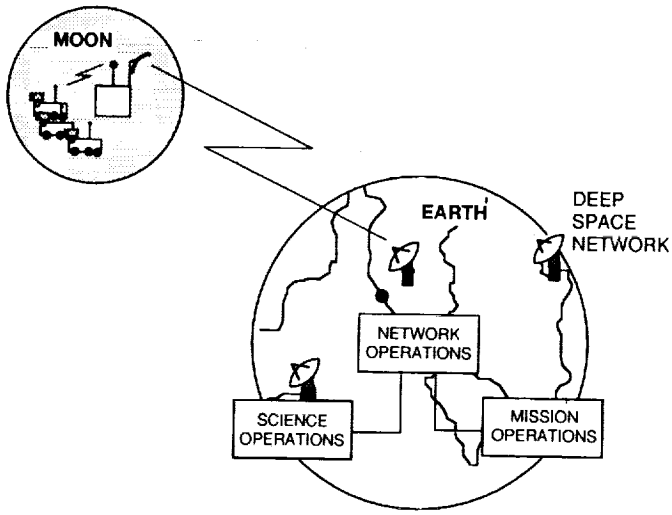


Fig. 1. Lunar microrover exploration.

APPLICATION LAYER	
HIGHER LAYERS	
CONVERGENCE SUBLAYER	ATM ADAPTATION LAYER
SEGMENTATION AND REASSEMBLY SUBLAYER	
ATM LAYER	
TRANSMISSION CONVERGENCE SUBLAYER	PHYSICAL LAYER
PHYSICAL MEDIUM DEPENDENT SUBLAYER	

Fig. 2. BISDN protocol stack.

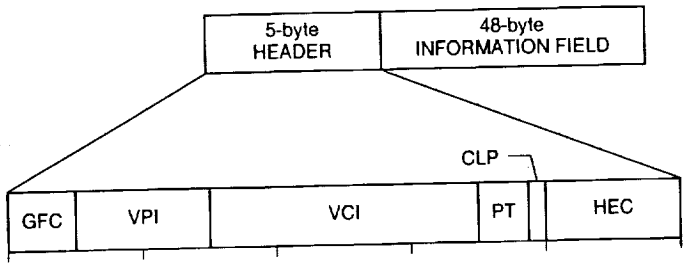


Fig. 3. ATM cell structure.

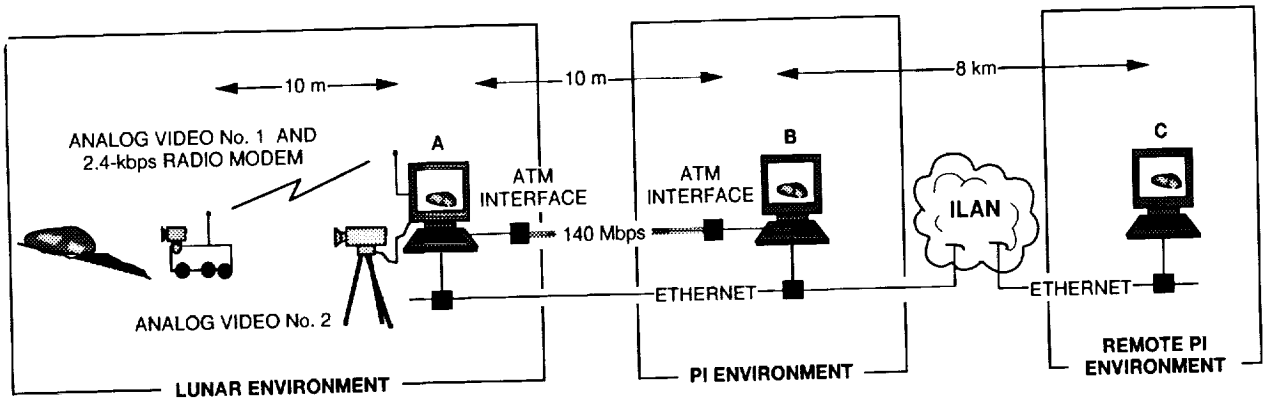


Fig. 4. Experiment configuration.

TDA Progress Report 42-114

August 15, 1993

An Overview of the Galileo Optical Experiment (GOPEX)

K. E. Wilson and J. R. Lesh
Communications Systems Research Section

Uplink optical communication to a deep-space vehicle has been demonstrated. In the Galileo Optical Experiment (GOPEX), optical transmissions were beamed to the Galileo spacecraft by Earth-based transmitters at the Table Mountain Facility (TMF), California, and Starfire Optical Range (SOR), New Mexico. The demonstration took place over an eight-day period (December 9 through December 16, 1992) as Galileo receded from Earth on its way to Jupiter, and covered ranges from 1–6 million km. At 6 million km (15 times the Earth–Moon distance), the laser beam transmitted from TMF eight days after Earth flyby covered the longest known range for transmission and detection.

I. Introduction

Optical communications technology is considered a viable contender for future space link applications. This is especially true for mini- and micro-spacecraft applications, where weight and size are the primary driving considerations. System studies, technology development, and deployment planning have been under way for the past 13 years and have resulted in a plan for both the spacecraft and Earth-reception ends [1]. Although there has been significant progress on component technologies, the need for systems-level demonstrations was becoming apparent. The Galileo spacecraft's second flyby of Earth, part of the Venus-Earth-Earth Gravity Assist (VEEGA) trajectory [2], afforded a unique opportunity to perform a deep-space optical uplink with the spacecraft as it receded from Earth on its way to Jupiter. The Galileo Optical Experiment (GOPEX) was conducted over the period December 9–16 from transmitter sites at Table Mountain Facility (TMF), California (see Fig. 1) and at the Starfire Optical Range (SOR), New Mexico (see Fig. 2). The spacecraft's Solid-State Imaging (SSI) camera was used as the optical communications uplink receiver. The experiment had three principal objectives, namely:

(1) Demonstrate laser beam transmission to a spacecraft at deep-space distances.

(2) Verify laser-beam pointing strategies applicable to an optical uplink based solely on spacecraft ephemeris predicts.

(3) Validate the models developed to predict the performance of the optical link.

Galileo's phase angle after its second Earth flyby was approximately 90 deg. Thus, as the spacecraft receded from Earth, it looked back at a half-illuminated Earth image. This geometry allowed laser beam transmission against a dark-Earth background. GOPEX was therefore conducted between 3:00 a.m. and 6:00 a.m. PST. This nighttime uplink had two distinct advantages:

(1) It allowed the uplink to be performed at the frequency doubled neodymium:yttrium-aluminum-garnet (Nd:YAG) laser wavelength of 532 nm, where the responsivity of the SSI camera is high.

(2) Long-exposure camera frames could be taken, which facilitated the identification of the detected laser transmissions. Pre-experiment analysis of the stray-light intensity in the focal plane of the camera showed that the camera shutter could remain open for up to 800 msec before the scattered light from the bright Earth saturated the pixels that detected the laser uplink.[1]

By scanning the camera across the Earth, parallel to the Earth's terminator, during each exposure, the laser signal was readily distinguished from spurious noise counts in the camera frame. With this strategy, the laser uplink appeared as a series of evenly spaced bright dots within the camera frame, quite distinct from other features in the frame.

GOPEX was a very successful experiment. Frames of laser uplink data were received on each of the seven days of the experiment. (The demonstration covered a period of eight days, but other spacecraft activities precluded laser transmission on Day 5.) The laser uplink was unequivocally detected on 48 of the 159 GOPEX frames taken. Because of an unanticipated pointing bias in the scan platform direction, pulses were not detected on any of the frames with exposure times less than 400 ms. Inclement weather, aborted transmissions, and restrictions imposed by regulatory agencies and by the Galileo Project team accounted for the loss of data on the remaining frames.

This article, an overview of the GOPEX demonstration, describes the experiment and presents a summary of the experimental results. An overview of the laser transmitters from the two sites is presented in Section II. A description of the optical receiver is given in Section III, followed by a brief discussion of the telescope pointing strategy in Section IV. The GOPEX results are presented in Section V, and the conclusion is presented in Section VI.

Companion articles included in this issue describe various aspects of the GOPEX demonstration in more detail. The article on telescope pointing [3] describes the process used to develop the Galileo telescope pointing files for the TMF and SOR transmitters and the strategy implemented in pointing the TMF telescope to the spacecraft. The article on the TMF optical train [4] describes the optical train design, laser beam characterization, characterization of the transmitter by using retroreflecting Earth-orbiting satellites in precursor experiments, and the TMF GOPEX uplink procedures. A. Biswas' article [5] describes in detail the characterization of the optical detection system that was used in the precursor experiments. This optical system was used to detect the laser signal retroreflected from the target satellites.

It was important to control the timing of the laser emission to ensure that the pulses arrived at the spacecraft while the camera shutter was open. Pulse-by-pulse monitoring of the laser output during GOPEX, stipulated in the GOPEX concurrence negotiated with the Galileo Project Office, not only satisfied the concurrence requirements but also provided the necessary data for comparing the received signal strengths with theoretical predictions of the optical channel. The article "GOPEX Laser Transmission and Monitoring System" [6] describes the configuration of the transmission and monitoring system at TMF and presents the laser output data recorded at this facility taken during GOPEX. These data were used to improve the estimates of the predicted signal strengths at Galileo. Good agreement was found between the experimental results and the theoretical predictions. The article by Levine, Shaik, and Yan [7] describes the analysis of the GOPEX data.

Ground-based deep-space optical communications of the future will employ a strategy of site diversity in ground transceiver deployment to ensure good line-of-sight visibility between the spacecraft and the transceiver station. A study of the December weather statistics for TMF over a four-year period, 1988 to 1991, showed that for 77 percent of the time the cloud cover at this facility lasted for more than two days.[2] The threat of inclement weather at TMF during GOPEX posed an unacceptably high risk to the demonstration. A dual-site transmitter strategy was therefore implemented to mitigate this risk. The selection of SOR as a second uplink site was based on a set of selection criteria with geographical location, (i.e., proximity to the TMF site) and the capabilities of site personnel as the principal considerations. The final article in this series [8] describes the SOR optical system along with the GOPEX and the precursor experiment operations as they were conducted from SOR.

[1] K. E. Wilson, "GOPEX Second Transmission Site" briefing to GOPEX Advisory Board (internal document), Jet Propulsion Laboratory, Pasadena, California, May 19, 1992.

[2] Ibid.

II. GOPEX Laser Transmitters

The laser transmitters at both sites consisted of a frequency-doubled Nd:YAG laser (532 nm) coupled to a Cassegrain telescope through a coudé mount arrangement. The transmitter characteristics are given in Table 1.

The TMF telescope used was the 0.6-m equatorial-mount astronomical telescope that had been used in 1968 to perform the laser transmission to the Surveyor 7 spacecraft on the Moon. The telescope is f/36 at the coudé focus, and the appropriate beam-forming lens set was inserted into the optical train, Fig. 3, to achieve the required laser beam divergence. The divergences used are given in Table 1. In the optical-train design, the laser beam reflected off the 0.2-m secondary mirror and illuminated a 15-cm subaperture on the telescope primary. The advantage of this subaperture illumination over full aperture was that it eliminated the large loss in transmitted signal strength that would have resulted from occultation by the 0.2-m secondary.

The two beam-forming lens sets, one for the 110-mrad divergence and a second for the 60-mrad divergence, were designed so that the laser beam was brought to a focus at a distance of 1.3 km when the telescope was focused at infinity. Light from the reference stars that was used to point the telescope to Galileo was collected across the full 0.6-m collecting aperture of the instrument.

The SOR telescope used for GOPEX was the 1.5-m system that is used for adaptive-optics experiments at this facility. A thin-film-plate polarizer served as the aperture-sharing element and coupled the laser output to the telescope optical train while allowing reference stars to be observed by the charge-coupled device (CCD) camera positioned in the orthogonal leg of the optical train. The required laser beam divergence was achieved by focusing the outgoing laser beam at ranges of 38km and 19km in the atmosphere. This corresponded to 40-μrad and 80-μrad beam divergences, respectively.

The output from the SOR laser was transmitted through the full 1.5-m aperture of the telescope. The effects of occultation by the 10-cm secondary were mitigated by reconfiguring the laser resonator so that it generated a flat-top intensity profile across the beam. With this design, occultation by the secondary accounted for less than 1 percent of the transmission loss in the optical train.

III. GOPEX Receiver

The GOPEX receiver was the Galileo SSI camera. The camera was mounted on the scan platform located on the despun section of the spacecraft. It consisted of a 12.19 x 12.19-mm CCD array composed of 800 x 800 silicon pixels and was located at the focal plane of a 1500-mm focal length f/8.5 Cassegrain telescope [9]. The angular resolution per pixel was 10.6 μrad, and the full well capacity per pixel was 100,000 electrons. The dark current was less than 10 electrons per pixel with 8-electrons readout noise. The camera has four gain states that are used to scale the video analog data to the 8-bit ADC. Over the eight-day experiment window, two gain states were used for GOPEX. Gain state 2, in which a data number (dn) of 255 corresponded to 100,000 detected photoelectrons, was used on the first two days; gain state 3, where the maximum dn corresponded to 40,000 detected photoelectrons, was used on the subsequent days.

Field correction elements, an eight-position filter wheel, and a two-blade shutter are positioned in the intervening space between the telescope primary and the focal plane of the SSI camera. The latter two elements were inherited from the Voyager program. The filter wheel contained six 20-nm bandpass color filters, one clear filter, and one infrared transmitting filter that were rotated into the optical train as required to enable the color reconstruction of an imaged scene. For GOPEX, the green color-filter with 50-percent transmission at 532 nm (the peak transmission was 90 percent at 560 nm) was used to reduce the effects of Earth-shine.

The shutter was operable in any one of 28 exposure times, ranging from 4.16 msec to 51.2 sec. The selection of exposure times was based on the estimate of the best balance between the conflicting requirements of short duration, to reduce stray light effects, and long duration, to ensure that enough pulses were detected across the image to confirm the laser uplink. Data taken at the Gaspra encounter were used to improve the original estimate of the scattered light intensity. Estimates of the scattered light rates for the GOPEX transmitter sites ranged from a high of 110 e-/msec on the end of the first day for the SOR location to 32 e-/msec on day 8 for both sites.[3] Based on these estimates, the GOPEX imaging sequence was designed with shutter times ranging from 133 to 800 msec. Actual scattered light rates measured during GOPEX ranged from 8 to 10 e-/msec [7]. These low scatter levels would have allowed longer camera exposure times and the accumulation of more data.

[3] K. Klaasen, "Scattered Light Predictions for GOPEX Observations," JPL Interoffice Memorandum (internal document), Jet Propulsion Laboratory, Pasadena, California, July 15, 1992.

IV. GOPEX Telescope-Pointing Strategy

The telescope-pointing files for both transmitter sites were generated from updates of the spacecraft ephemeris file that were provided to the GOPEX team on December 8 and December 11. The strategy was to off-point the telescope from reference stars that were located within 0.5 deg of the spacecraft's position. Over the eight-day period, six guide stars of magnitudes 6 to 10 were used to point the TMF telescope at Galileo.

Transmission to Galileo was accomplished by using a "point and shoot" approach. In this technique, the telescope was set to track the reference star in the intervals between the three-second bursts of laser transmissions. Two and one-half minutes before laser transmission, the reference star was positioned in the center of the field of view of the focal plane aperture at coudé and the telescope was calibrated. Ten seconds prior to transmission, the telescope was pointed to Galileo's predicted location and set to track the spacecraft for the next thirteen seconds. This procedure was repeated during the three- to six-minute intervals between the laser transmissions. Because the telescope calibration was performed just before transmission, the pointing errors introduced by mount sag were reduced significantly. In addition, the high elevation of the spacecraft during the uplink—the experiment was conducted when Galileo's elevation from TMF was greater than 30 deg—and the proximity of the reference stars obviated the need to implement atmospheric refraction compensation techniques while pointing to the spacecraft.

To test the accuracy of the telescope-pointing predicts, SOR dithered the laser beam in a circle of 85-mrad radius about the predicted position of the spacecraft, while the TMF transmitter pointed directly to the predicted spacecraft position. This was done on the first day of GOPEX for several of the long-duration frames (frames with exposure times greater than 400 msec). The results are shown in Fig. 4. Nine pulses can be clearly discerned in the figure; seven are from the 15-Hz TMF transmitter, and two are from the 10-Hz SOR transmitter. Without beam scanning, a total of four pulses would have been detected from the SOR transmitter. The presence of only two pulses from SOR and seven from TMF clearly demonstrates that the error in the telescope pointing predicts was significantly less than 85 mrad. This was further confirmed by the successful use of a 60-mrad beam from TMF for laser transmissions on the last three days of GOPEX.

V. GOPEX Results

A summary of the detected GOPEX laser transmissions over the duration of the experiment is given in Table 2. Over the eight-day period, transmissions to the spacecraft were made over a range beginning at 600,000 km on the morning of December 9 and ending at 6,000,000 km on the morning of December 16. Signals were successfully detected on each of the experiment days, although not on all frames within a given day. There were several reasons for the lack of detection on all frames. These included unfavorable weather (which caused outages), regulatory agency restrictions on transmissions, temporary signal-to-noise anomalies on the downlink, and an unexpected camera-pointing bias error. Final results show that the laser uplink was successfully detected on 50 camera images during the experiment window. Two representative images showing the detected laser pulses are shown in Figs. 4 and 5.

Although transmitted laser pulses were detected on each of the seven days of the uplink, adverse weather at the sites and not telescope pointing was the most severe impediment to successfully detecting the laser transmissions. Winter storms at TMF and SOR brought snow, heavy clouds, and ground fog to these facilities. Transmission from TMF was most affected on the first and fourth days of the experiment. The last seven frames obtained on the first day were taken with TMF completely overcast and SOR in daylight. On the fourth day, falling snow at TMF precluded transmission from this facility; also on that day, during only one of the ten transmissions there was clear sky between the SOR transmitter and the spacecraft. Falling snow and heavy cloud cover prevented transmission from SOR on the last three days.

Restrictions from regulatory agencies were also responsible for data outages. Transmission of the GOPEX laser beam into space required the concurrence of the U.S. Space Defense Operations Center (SPADOC). On the first day, SPADOC restrictions prevented TMF from transmitting during four frames. An additional frame was lost because the ground receiving station (at Goldstone, California) momentarily lost lock on the Galileo spacecraft downlink signal. Owing to the loss of downlink signal, the orientation of the spacecraft could not be confirmed, and since one of the GOPEX concurrence conditions was that laser uplink would proceed only if the spacecraft orientation was known, no laser transmissions were sent during this data outage.

During the first two days of GOPEX, the spacecraft orientation resulted in the low-gain antenna being pointed away from Earth. This resulted in a low signal-to-noise ratio (SNR) of the spacecraft downlink and was evidenced by the numerous burst errors in the downlinked data. See also [7]. The GOPEX images for these days showed numerous

streaks across the frames, which made it difficult to discern successful laser transmissions on the images. Just after the GOPEX uplink on day 2, a planned spacecraft maneuver was executed. This maneuver increased the SNR of the radio frequency downlink and resulted in clearer GOPEX images for the remainder of the demonstration.

The GOPEX demonstration required that the SSI camera be operated in a mode for which it was not designed (that is, slewing the camera during imaging). To get the GOPEX transmitter sites in the field of view during the slew, the camera was initially pointed to a position above or below the targeted direction and the shutter was opened at a prescribed time after the start of the slew. Uncertainties in the stray-light intensity in the focal plane of the SSI camera dictated the shutter times used for GOPEX. The times chosen ranged from 133 to 800 msec, and these were loaded into the spacecraft sequence of events prior to GOPEX. As the experiment progressed, it was observed that laser transmissions were consistently detected only on frames with greater than 400-msec exposure times; on days 4 and 7, laser pulses were detected in one of the 200-msec exposure frames.[4] The consistent absence of detections on the shorter duration frames was traced to a pointing error caused by the scan platform acceleration being slower than predicted. As a result, no clear evidence of laser transmission was observed on 88 of the 90 frames taken with exposure times less than 400 msec.

[4] B. M. Levine, "GOPEX Data Products Summary," JPL Interoffice Memorandum (internal document), Jet Propulsion Laboratory, Pasadena, California, December 22, 1992.

The remaining 69 frames were analyzed to determine the detected pulse energy statistics and to compare those statistics with theoretical predictions. Figure 6 shows a typical comparison. The histogram of the detected pulse energies is plotted along with the theoretical distribution using atmospheric turbulence data and the appropriate turbulence model. A log-normal intensity distribution based on statistics of the measured data is also shown. As the figure shows, there is good agreement between the measured and theoretically predicted distributions.

VI. Conclusion

The results of the first deep-space optical transmission to a spacecraft in flight have been presented. The transmission was performed from transmitters located at TMF, California, and SOR, New Mexico. The laser uplink was detected on every day of the experiment—out to a range of 6,000,000 km for the TMF transmission. The camera images returned from Galileo and the analysis of the data show that the distribution of the detected signal strengths is consistent with theoretical predictions. All experiment objectives were achieved.

Acknowledgments

The authors would like to express their gratitude to all the members of the GOPEX team for their invaluable contributions. The authors thank the Management of Divisions 33, 38, the TDA Program Office, the Galileo Project Office, Phillips Laboratory, and the Management of Sections 331, 381, 382, 385, and 320 for their unwavering support and commitment to the success of GOPEX. The authors thank the GOPEX Advisory Board for its insightful guidance and the Readiness Review Boards for their critical assessment of the operations teams' readiness to perform GOPEX, and especially for their dedication and support. Finally the authors thank the SOR and TMF Operations teams for a superb job preparing for and executing GOPEX.

References

[1] J. R. Lesh, L. J. Deutsch, and W. J. Weber "A Plan for the Development and Demonstration of Optical Communications for Deep Space," *The Telecommunications and Data Acquisition Progress Report 42-103, vol. July-September 1990*, Jet Propulsion Laboratory, Pasadena, California, pp. 97–109, November 15, 1990.

[2] K. E. Wilson, J. Schwartz, and J. R. Lesh, "GOPEX: A deep space optical communications experiment with the Galileo spacecraft," *SPIE Proceedings*, vol. 1417, pp. 22–26, January 1991.

[3] W. M. Owen, Jr., "Telescope Pointing for GOPEX," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 230–235, August 15, 1993.

[4] J. Yu and M. Shao, "Galileo Optical Experiment (GOPEX) Optical Train: Design and Validation of the Table Mountain Facility," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June, 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 236–247, August 15, 1993.

[5] A. Biswas, "Calibration of the Receiver Channel for the GOPEX Precursor Experiments," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 205–212, August 15, 1993.

[6] G. Okamoto and K. Masters, "GOPEX Laser Transmission and Monitoring Systems," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 248–254, August 15, 1993.

[7] B. M. Levine, K. S. Shaik, and T.-Y. Yan, "Data Analysis for GOPEX Image Frames," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 213–229, August 15, 1993.

[8] R. Q. Fugate, "GOPEX at the Starfire Optical Range," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 255–279, August 15, 1993.

[9] K. P. Klaasen, M. C. Clary, and J. R. Janesick, "Charge-coupled Device Television Camera for NASA's Galileo Mission to Jupiter," *Optical Engineering*, vol. 23, no. 3, pp. 334–342, May/June 1984.

Table 1. GOPEX laser transmitter characteristics.

Characteristic	Table Mountain Facility	Starfire Optical Range
Wavelength, nm	532	532
Pulse energy, mJ	250	350
Repetition rate, Hz	15–30	10
Pulse width, nsec	12	15
Beam divergence, μrad	110	80
Days 1–4	60	40
Days 6–8		
Telescope mirror diameter	0.6	1.5
Primary, m	0.2	0.1
Secondary, m		
Optical train transmission, percent	60	43

Table 2. Summary of detected laser signals.

Day	Shutter speed, ms	Frames received	Frames with detections
1	133	9 of 10	0
	200	24 of 25	0
	400	19 of 20	6
	800	5 of 5	4
2	200	5 of 5	0
	267	15 of 15	0
	533	15 of 15	11
	800	5 of 5	5
3	200	5 of 5	0
	267	10 of 10	0
	533	5 of 5	5
4	200	3 of 3	0[b]
	267	4 of 4	0
	533	3 of 3	2[b]
5		No activity planned	
6[a]	133	3 of 3	0
	267	6 of 6	0
	533	3 of 3	3
7[a]	200	3 of 3	0
	400	3 of 4	3
	800	3 of 3	3
8[a]	267	2 of 2	0
	533	4 of 4	4
	800	2 of 2	2

[a] Adverse weather at Starfire Optical Range precluded laser transmission on this day.

[b] Adverse weather at Table Mountain Facility precluded laser transmission on this day, and it was cloudy at Starfire Optical Range.

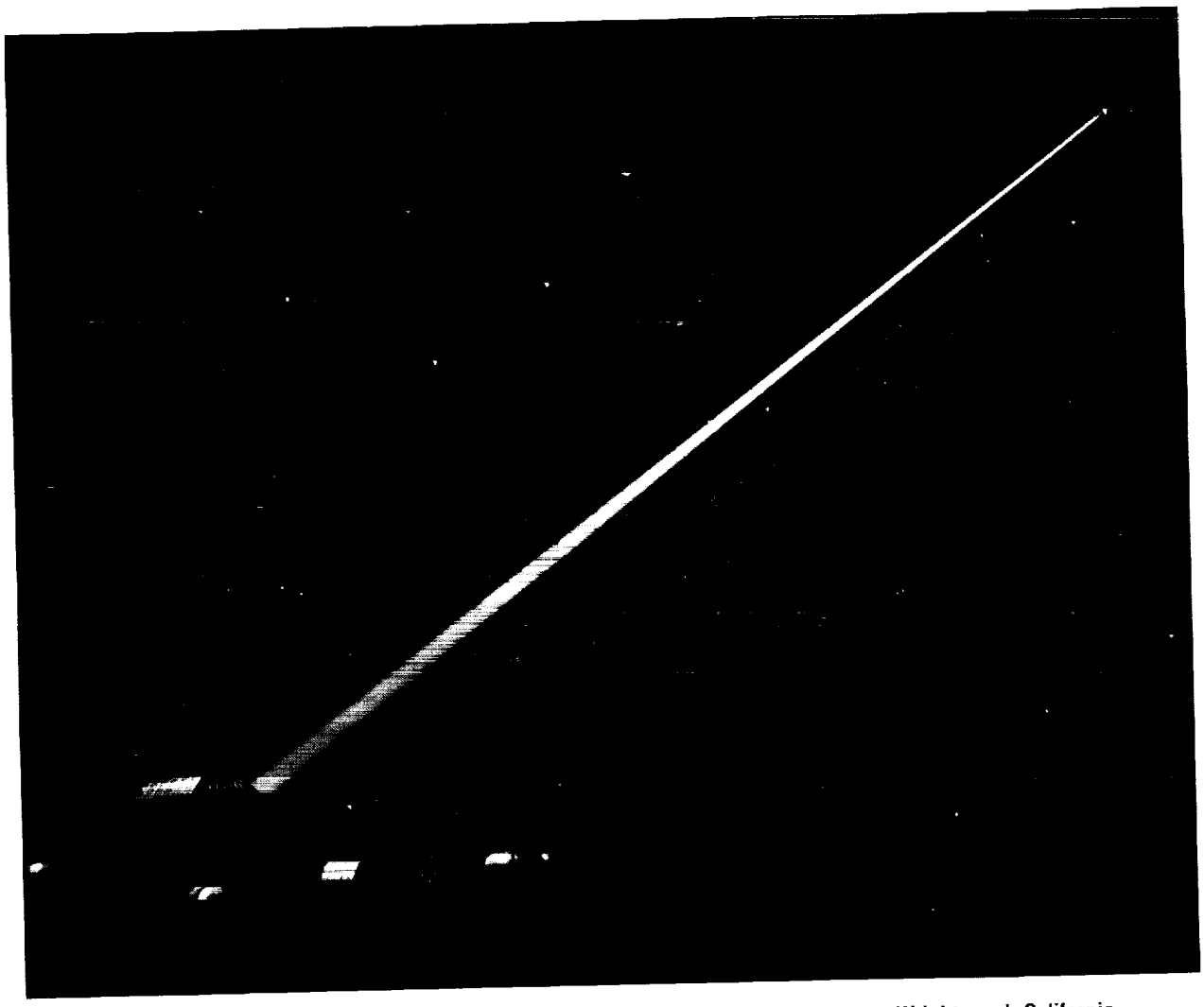


Fig. 1. Laser transmission from the 0.6-m telescope at the Table Mountain Facility, Wrightwood, California.

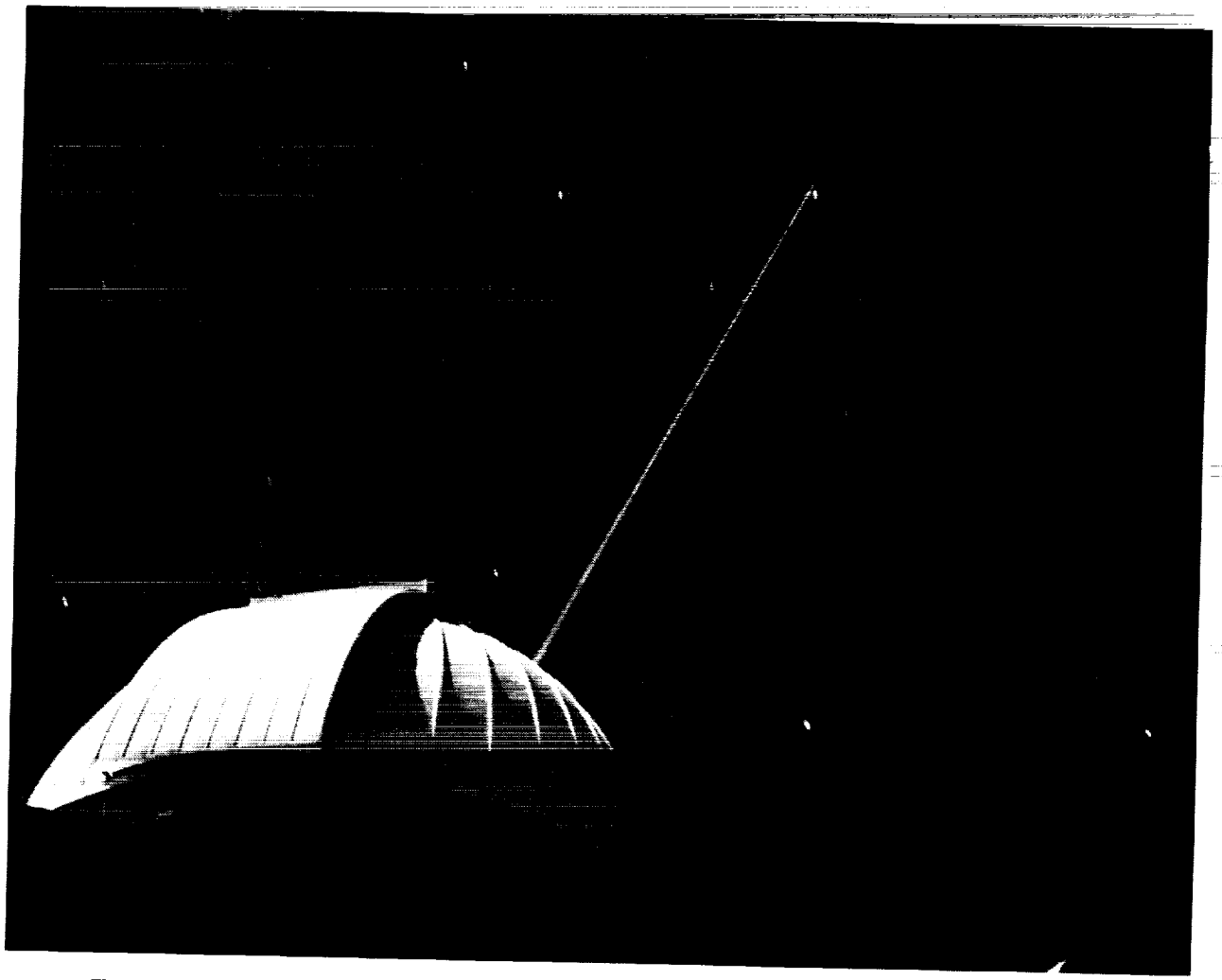


Fig. 2. Laser Transmission from the 1.5-m telescope at Starfire Optical Range, Albuquerque, New Mexico.

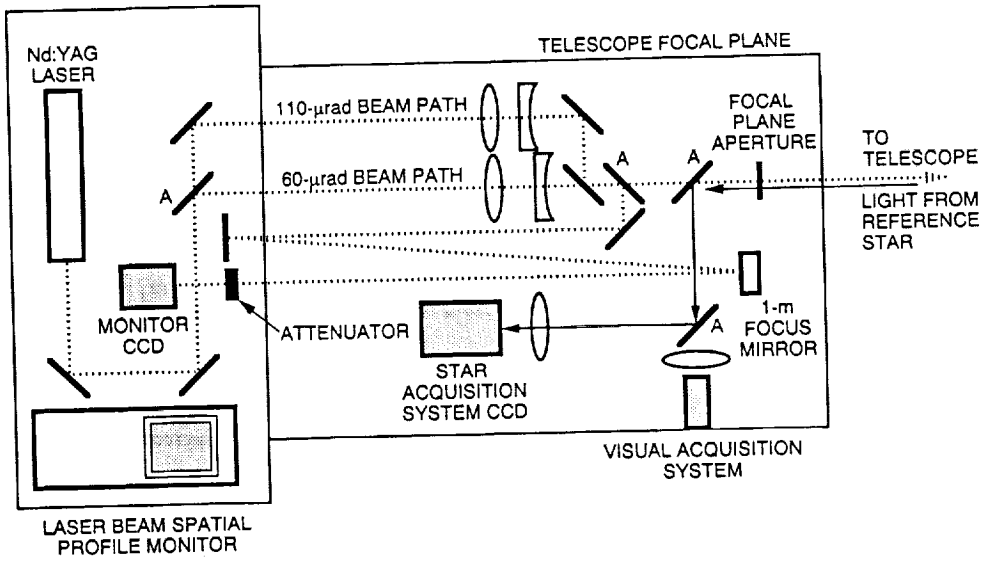


Fig. 3. GOPEX optical train, at the Table Mountain Facility. Relay mirrors (labeled "A" in the figure) are appropriately inserted into the optical train to obtain the required beam divergence.

Fig. 4. Camera image showing detected pulses from TMF and SOR. To the right of the figure is the blurred image of the portion of the Earth illuminated by sunlight. Detections of the TMF laser pulses are shown on the left and of SOR pulses closer to the right. The missing pulses are due to the spatial scanning of the SOR beam.

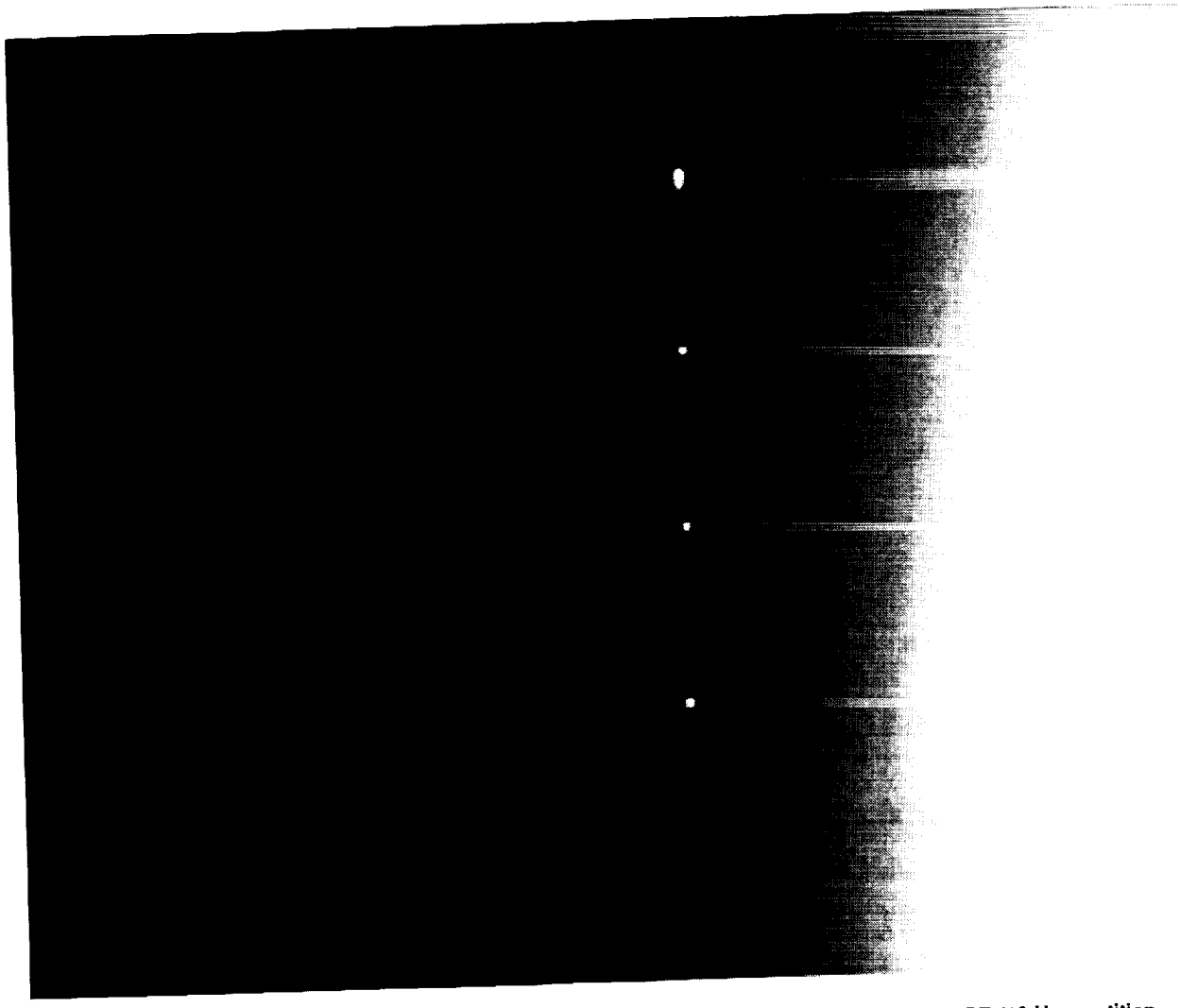


Fig. 5. Camera image showing laser uplink detections from TMF (15-Hz repetition rate) and SOR (10-Hz repetition rate) on Day 2 of GOPEX. SOR scanning was off.

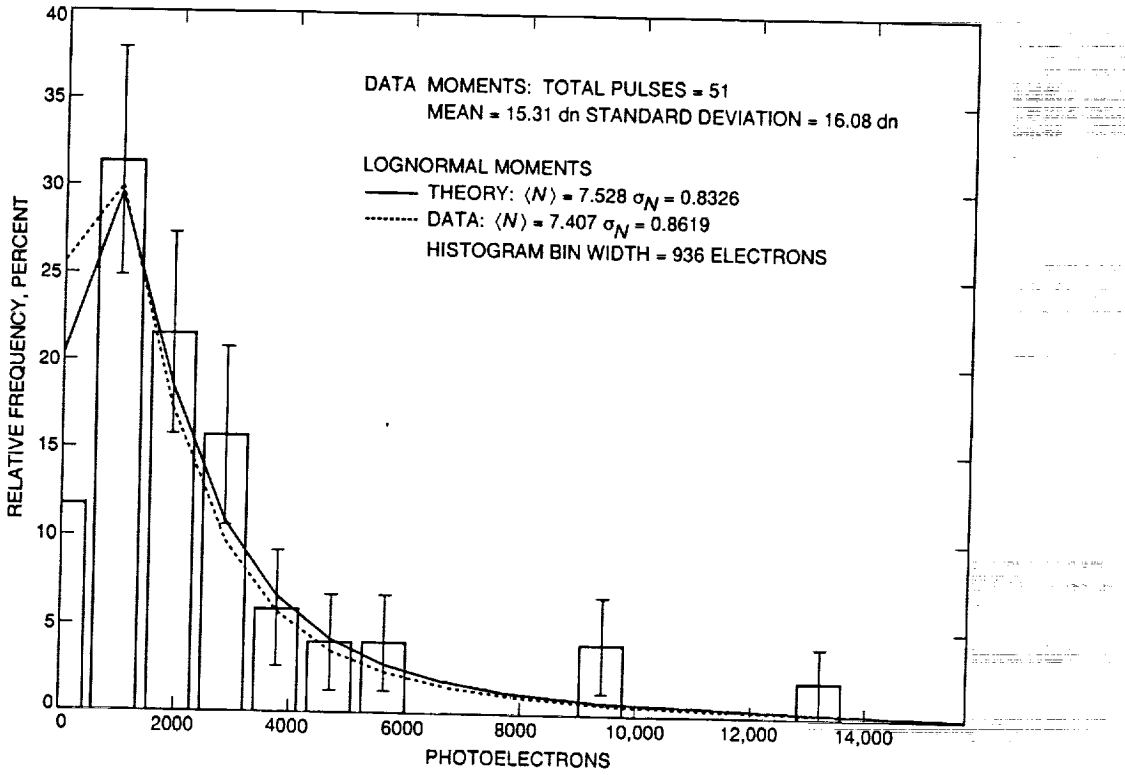


Fig. 6. Histogram of detected signal strengths from TMF on Day 8 of GOPEX. The data show good agreement between the experimental and the theoretical lognormal distribution using parameters for strong turbulence theory.

TDA Progress Report 42-114

August 15, 1993

Calibration of the Receiver Channel for the GOPEX Precursor Experiments

A. Biswas
Space Materials Science and Engineering Section

Calibration measurements for the cooled (253 K) photomultiplier tube (PMT) detector, used on the receiver channel of the transmit/receive (T/R) switch for GOPEX, were conducted in the laboratory. A pulsed frequency-doubled neodymium:yttrium-aluminum-garnet (Nd:YAG) laser was used to direct 532-nm light on the PMT. By monitoring the energy per pulse of the light incident on the PMT, the minimum number of photons detected could be determined. These results agreed with the photon flux back-calculated from the PMT output waveform. Approximately 700 incident photons arriving during a temporal pulse width of ~65 nsec were detected with a signal-to-noise ratio (SNR) of ~1. Other receiver channel characteristics, such as PMT dark currents, optical transmission, and interference filter sensitivity to angle of light incidence, were also measured.

I. Introduction

GOPEX was successfully concluded on December 16, 1992. Key to the success of this experiment were the precursor experiments that preceded GOPEX. These experiments were used to calibrate the laser transmitters from both the Starfire Optical Range (SOR) and Table Mountain Facility (TMO) sites, by retro-reflecting the laser beam from specifically targeted Earth-orbiting satellites. The retro-reflected signal was detected using a gallium arsenide (GaAs) PMT. A polarizing beam splitter was used at SOR and a holey mirror was used at TMF to separate the T/R signal paths.

This article describes the calibration of the T/R switch that was used at TMF to detect the satellite return signals. The switch, shown in Fig. 1, consisted of a 10-cm diameter, dielectrically coated reflecting surface (R > 99 percent at 532 nm for 45-deg incidence) along with two similarly coated relay mirrors that were used to relay the signal to the PMT. An interference filter was positioned in front of the PMT to reduce the amount of background light detected. The receiving train was characterized both on a component level and as an integrated system. The results of these tests are presented here. The T/R switch was characterized over the range of signal levels expected for the satellite returns. These ranged from tens of photons for the distant Etalon satellites (20,000 km) to thousands of photons for the close Earth-orbiting Ajesai satellites (1500 km).

II. Objectives

The four objectives of this work were:

(1) Characterize PMT at room temperature (295 K) and after cooling to 253 K.

(2) Measure optical transmission at 532 nm for the T/R switch with and without the 532-nm interference fil-

ter. This measurement allowed a determination of the transmission losses of the interference filter.

(3) Measure the sensitivity and dynamic range of the PMT using single laser pulses and assess recovery time for the PMT. Scintillation effects due to atmospheric turbulence can result in large variations in the return signal strength. This measurement would enable evaluation of the detector's linear range.

(4) Measure transmission of the 532-nm interference filter as a function of the angle of incidence. This measurement would provide a better estimate of detected signal strength for less than optimal alignment of the relay mirrors, i.e., non-normal incidence of return signal on the filter.

III. Experimental Setup

The optical arrangement depicted in Fig. 2 was used to measure the sensitivity and dynamic range of the receiver channel detector. A Q-switched, frequency-doubled Nd:YAG laser operating at a pulse repetition rate of 3.5 kHz, and a temporal pulse width full width at half maximum (FWHM) of ~65 nsec, was directed through the optical receiving path of the T/R switch and brought to incidence on the GaAs photomultiplier tube (PMT) detector.[1] The PMT output was displayed on an oscilloscope.[2] The 1.06-μm high-reflectivity (HR) flats shown were used to filter the residual fundamental frequency (1.06-μm wavelength) laser light from the signal. The power/energy meter[3] was used to measure the beam energy to the T/R switch. The beamsplitter (BS) reflected a fraction of 532-nm light onto a fast photodiode,[4] the output of which was displayed on the oscilloscope. This photodiode signal served as a reference. A 10× attenuation filter was used in front of the photodiode to extend the range of laser energies over which measurements could be made without saturating the fast photodiode detector. A 532-nm pass interference filter was positioned at the T/R switch exit port. A green-light-blocking filter[5] served as a fixed attenuator at the entrance port of the T/R switch, while neutral density (ND) filters were used to attenuate the beam to the desired power level.

[1] Manufactured by Burle Electron Tubes (PMT Serial number 96684, base serial number 31304/0076/0588; 1874-92).

[2] Tektronix DSA 602A.

[3] Photodyne Model 300 head connected to Model 66XLA readout.

[4] Newport Model 877.

[5] LASER-GARD series LGA, supplied by Newport Research Corporation.

IV. Results

(1) Characterization of the PMT at room temperature and at 253 K.

The PMT dark current was measured using an ammeter,[6] first at room temperature and then at ~253 K. At room temperature the dark current increased monotonically with increasing anode voltage and was ~8 nA at 1400 V. At 253 K and 1400 V, typical dark currents were 0.007 nA with occasional spikes registering as much as 0.02 nA. These measured values were consistent with the manufacturer's specifications. The PMT dark pulses were typically 10 nsec wide (FWHM).

(2) Measurement of the optical transmission.

The first step in measuring the optical transmission of the T/R switch at 532 nm with and without the 532-nm interference filter was to maximize the PMT output signal.

This was done to ensure optimum alignment of the optical train. With the holey mirror rotating at 30 Hz, the PMT signal was observed on an oscilloscope.[7] The expected waveform resulting from periodic light blockage confirmed the optical alignment. The holey mirror for the T/R switch was then disconnected and fixed so as to reflect all the laser light to the PMT. Light transmission measurements, with and without the interference filter, were made using the power/energy meter; these results are shown in Table 1. From the data in Table 1, the return path transmission of the T/R switch was calculated at 75 percent, and at 48 percent with the interference filter.

(3) Measurement of the sensitivity and dynamic range of the PMT using single laser pulses.

The photodiode reference signal was calibrated against the Photodyne energy meter by operating the laser in a single pulse mode and comparing the peak voltage of the detector to the output of the energy meter. This calibration was accurate to within 14 percent; the shot-to-shot variation of the laser pulse width. Figure 3 shows the calibration data. Linear regression of the data yielded a slope of $5.93 \times 10^{-4} \pm 7.84 \times 10^{-5} \mu J/mV$. The calibration constant was used to convert peak millivolts detected by the photodiode to laser energy incident at the T/R switch.

[6] Keithley Model 485.

[7] Tektronix Model 2215A.

Figure 4 shows typical simultaneous acquisitions made with the reference photodiode (positive-going pulse) and the PMT (negative-going pulse) positioned as shown in Fig. 2. The FWHM measured on both traces is approximately 65 nsec. The laser after pulsing detected by the photodiode was not detected by the PMT because of the large attenuation in the optical train of this detector. The alignment of the reference detector (fast photodiode) with respect to the split fraction of light incident on it was maintained and periodically checked to ensure accurate determination of actual laser pulse energies incident on the T/R switch.

Figure 5 is a plot of the PMT response as a function of the microjoules of incident energy. Each data point represents a single 65-nsec FWHM laser pulse. The data indicate that for pulse energies greater than $\sim 1.5\mu$J, the PMT response becomes nonlinear. A linear regression of the data in Fig. 5 over the expected operating range (1 nJ – 1μJ) yields a slope of 0.85 V/μJ.

A LASER-GARD filter was used at the T/R switch entrance to attenuate the incident laser beam and enable characterization of the T/R switch at the low photon levels expected in the retro-reflected returns from the higher satellites. The filter's optical density was 6.53 at 532 nm, and it enabled calibration of the receiver down to a few hundred photons incident on the PMT. The results are shown in Fig. 5 where the PMT peak voltage is plotted against the energy incident on the PMT and against the number of photons on the upper horizontal scale.

The number of photons incident on the PMT was confirmed from a calculation of the area under the waveform. This was done as follows: The FWHM of the laser pulse was multiplied by the anode current (determined from the peak voltage divided by 50 ohms, the output impedance of the PMT). This product represents the charge on the anode. The anode charge was divided by the gain (1.2×10^5) to determine the number of photoelectrons generated at the cathode. This was then divided by the PMT quantum efficiency (QE$\sim$0.145) and the energy per photon to determine the number of photons incident on the cathode. Table 2 shows the results.

The noise equivalent power (NEP) estimated from the low level signals was 10^{-13} W. The recovery times of the PMT over the range of pulse energies used during the investigations were found to be excellent in that no "ringing" was observed.

(4) Measurement of transmission of the 532-nm interference filter as a function of angle of incidence of the laser beam.

To measure the angular dependence of the interference filter transmission, the filter was mounted on a precision rotary mount, ensuring normal (0-deg) incidence by checking for auto-collimation (Fig. 6). The rotary stage was used to increment the angle of incidence while monitoring the transmitted laser power through the filter. Data were recorded by the power/energy meter for clockwise and counterclockwise angular rotation of the filter. The data show that the filter transmission remains >90 percent for up to a 4-deg angle of incidence. These results show that there is no significant loss in signal strength for small (<3 deg) deviations in incident angle introduced as the relay mirrors steer the return beam to the PMT.

V. Conclusions

The T/R switch used for detecting the retro-reflected signals from target satellites was characterized. The results show that the detector assembly would readily measure return signals from low Earth orbiting satellites, such as Ajesai. The PMT's 10^{-13} NEP indicates that this system can measure the tens of photons in a 10-nsec pulse expected from the high Earth orbiting Etalon satellites.

This article also gives the results of the measurements of (1) the performance of the PMT at room temperature versus that at 253 K and (2) the angular dependence of the interference filter transmission. The data show that cooled PMT operation is mandatory if one is to detect the anticipated low signal levels and that the transmission losses through the interference filter remain within 90 percent of peak for angle of incidence less than 4 deg.

Acknowledgments

The author is grateful for the guidance provided by K. Wilson. The assistance provided by K. Masters and M. S. Shumate in obtaining the results reported is much appreciated.

Table 1. Measurement of optical transmission at 532 nm for the T/R switch with and without the 532-nm interference filter.

Power meter location	Mean pulse energy, μJ
Entrance of the T/R switch	34.54 ± 1.76
At the exit port, with the 532-nm interference filter	16.67 ± 2.03
At the exit port, without the 532-nm filter	25.9 ± 2.5

Table 2. A comparison of the number of photons calculated using the deduced single shot laser energy from the photodiode measurement and estimated from the waveform simultaneously recorded by the PMT.

Peak, mV photons	FWHM, nsec	Number of incident photons measured using calibration curve from Fig. 5	Number of incident photons measured using the PMT waveform
2	18	700	638
36	67	1.56×10^4	1.73×10^4
175	63	5.70×10^4	7.92×10^4
412	55	1.55×10^5	1.63×10^5
461	68	1.71×10^5	2.27×10^5
629	60	2.39×10^5	2.72×10^5
2106	78	1.18×10^6	1.18×10^6

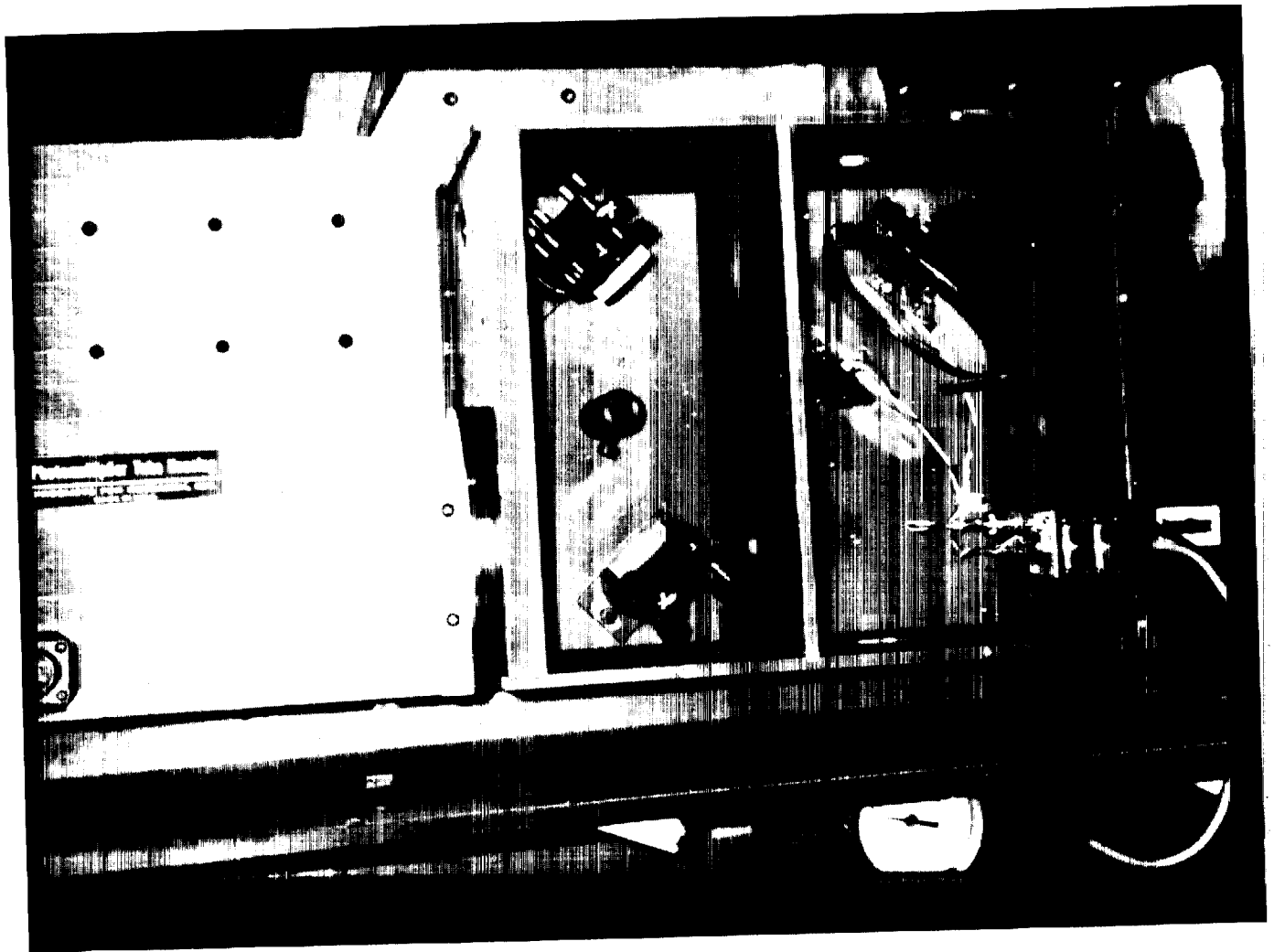


Fig. 1. Photograph of T/R switch with top cover removed. The 10-cm holey mirror consisting of a dielectrically coated surface is shown in the right chamber. The two relay mirrors that direct the return signal to the photomultiplier tube are shown in the left chamber tube. A polarizing beam splitter at SOR and a holey mirror at TMO were used to relay the return beam to the detector.

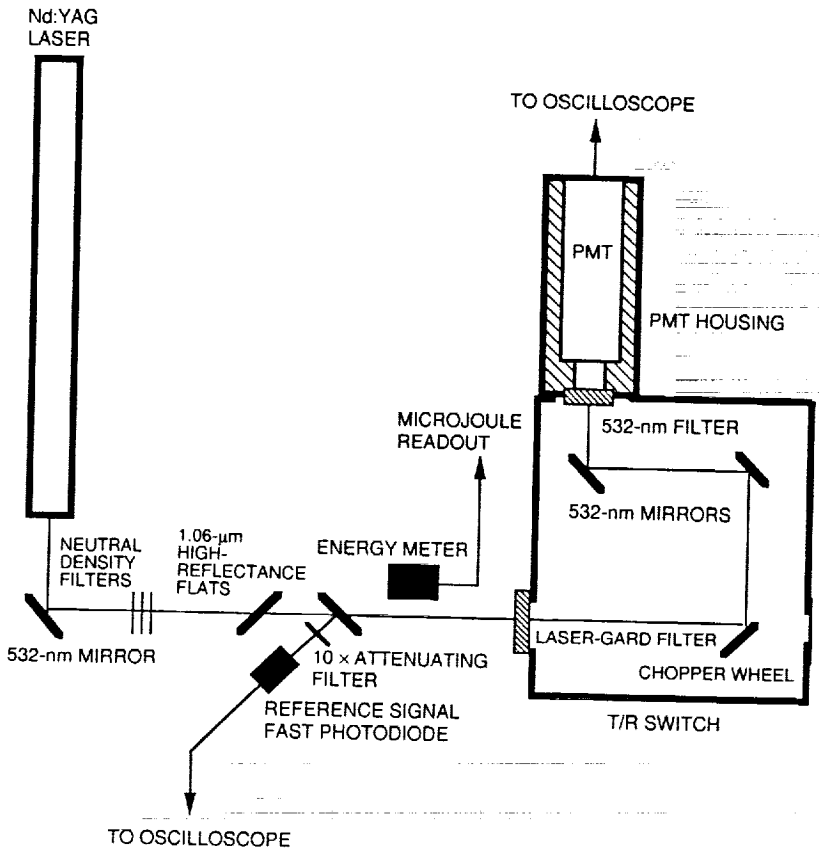


Fig. 2. Schematic experimental arrangement used for calibration of the receiver channel of GOPEX.

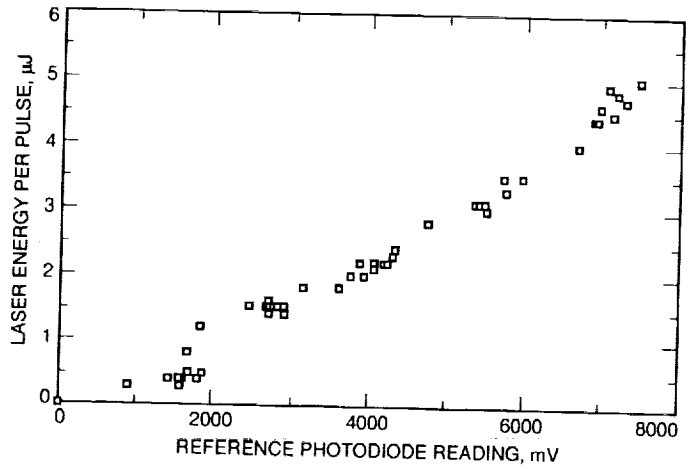


Fig. 3. The calibration curve used to convert photodiode-pulse peak heights to laser energy.

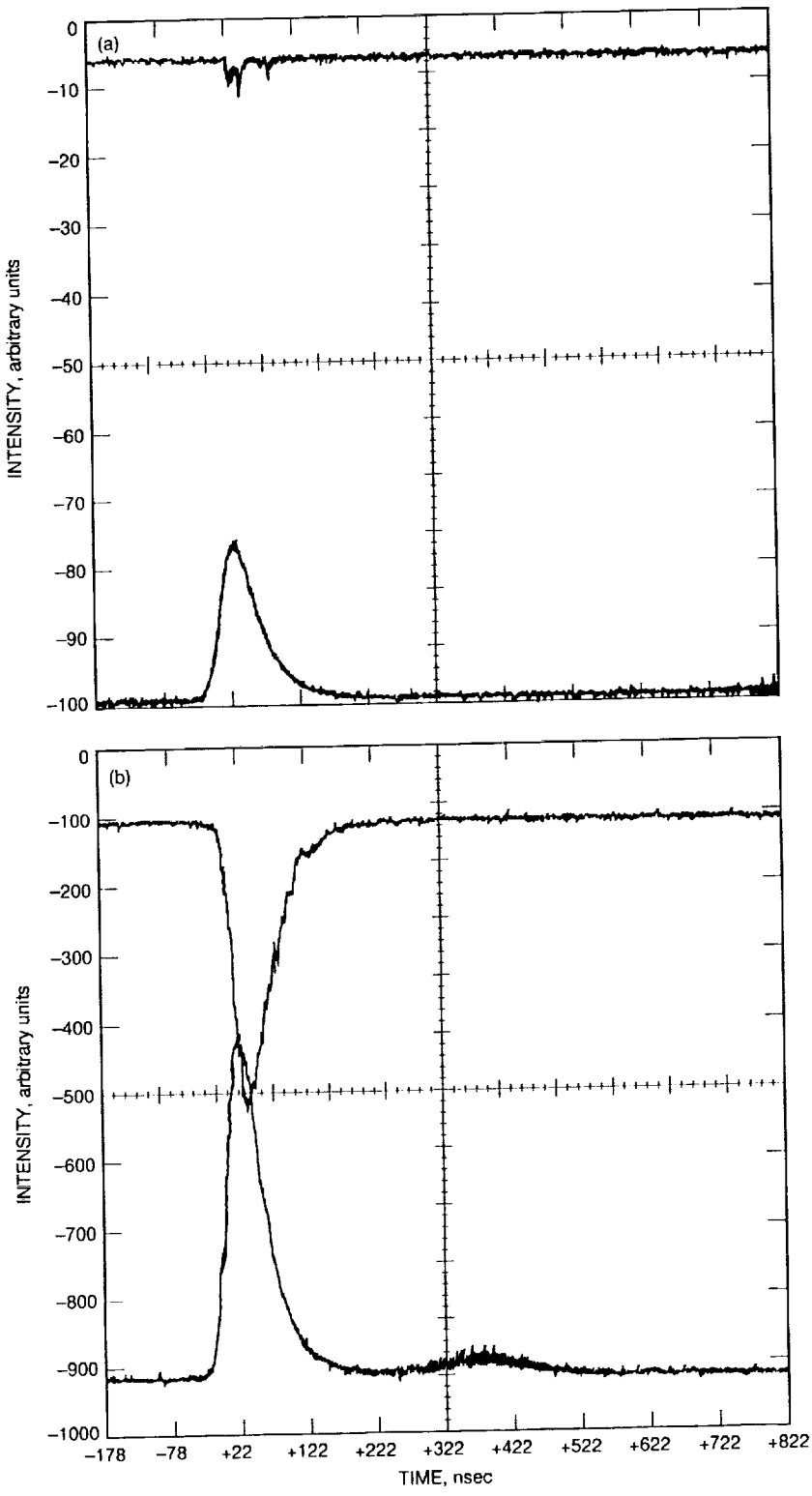


Fig. 4. Representative simultaneous data acquisitions made on the PMT (negative-going pulse) and reference signal (positive-going pulse) using the Tektronix DSA 602A oscilloscope: (a) $\sim$0.001-μJ incident energy, and (b) $\sim$0.3-μJ incident laser energy.

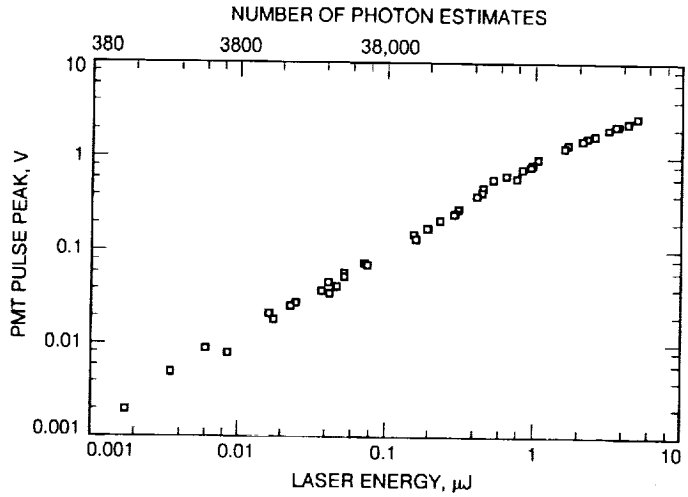


Fig. 5. The PMT response to 532-nm, 65-nsec single laser pulses. The abscissa shows the laser energy incident on the LASER GARD filter at the entrance of the T/R switch. The corresponding photodiode pulse peak heights are shown on the ordinate. The upper right horizontal axis indicates the actual number of photons incident on the PMT after correcting for the attenuation of the LASER GARD filter and T/R switch transmission.

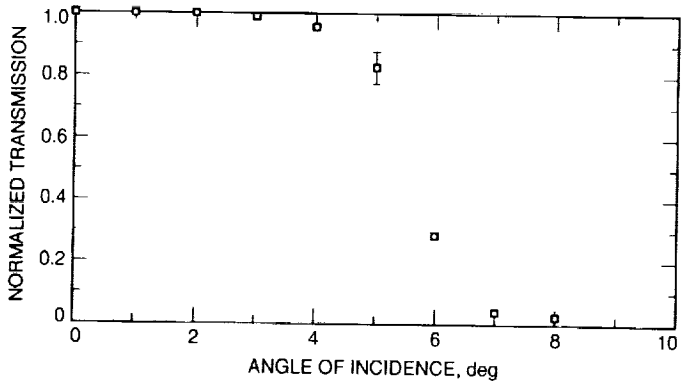


Fig. 6. Transmission of the interference filter as a function on the angle of incidence. The filter transmission remained greater than 90 percent for angles of incidence of 4 deg and less.

Data Analysis for GOPEX Image Frames

B. M. Levine
Optical Sciences and Applications Section
K. S. Shaik and T.-Y. Yan
Communications Systems Research Section

This article describes the data analysis based on the image frames received at the Solid State Imaging (SSI) camera of the Galileo Optical Experiment (GOPEX) demonstration conducted between December 9 and 16, 1992. Laser uplink was successfully established between the ground and the Galileo spacecraft during its second Earth-gravity-assist phase in December 1992. SSI camera frames were acquired which contained images of detected laser pulses transmitted from the Table Mountain Facility (TMF), Wrightwood, California, and the Starfire Optical Range (SOR), Albuquerque, New Mexico. Laser pulse data were processed using standard image-processing techniques at the Multimission Image Processing Laboratory (MIPL) for preliminary pulse identification and to produce public release images.

Subsequent image analysis corrected for background noise to measure received pulse intensities. Data were plotted to obtain histograms on a daily basis and were then compared with theoretical results derived from applicable weak-turbulence and strong-turbulence considerations. This article describes processing steps and compares the theories with the experimental results. Quantitative agreement was found in both turbulence regimes, and better agreement would have been found, given more received laser pulses. Future experiments should consider methods to reliably measure low-intensity pulses, and through experimental planning to geometrically locate pulse positions with greater certainty.

I. Introduction

This article describes the data analysis based on the image frames received at the Solid State Imaging (SSI) camera of the Galileo Optical Experiment (GOPEX) demonstration conducted between December 9 and 16, 1992. Simultaneous pulsed laser transmissions from the Table Mountain Facility (TMF), Wrightwood, California, and the Starfire Optical Range (SOR), Albuquerque, New Mexico, were recorded on the Solid-State Imaging (SSI) camera as a series of illuminated pixels in one image frame. Each pixel was quantized into 256 levels for transmission back to the JPL Multimission Image Processing Laboratory (MIPL) for processing. The image frames were transferred to the authors for analyses of the laser pulse strength.

A top-level summary of the GOPEX experiment, including statistics on the number of pulses received for each

day and for each frame, has been published separately,[1] and this article provides further results of statistical analyses. Section III describes image-processing techniques used at the MIPL and the custom procedures developed using the commercially available MATLAB software package to estimate laser pulse intensities, daily pulse-intensity histograms, and other relevant results. Software algorithms were written to identify each laser pulse and measure its strength above background. A statistical summary was compiled for each day's activities. These summaries were put in a form allowing comparison with an analytic model that predicts the probability distribution, and its moments for each night. Section IV provides a brief description for computing the lognormal variance for using either a weak turbulence theory or a strong turbulence theory. Section V provides a discussion comparing the experimental results with the statistical results using the parameters based on the statistical models discussed in Section IV.

II. GOPEX Summary

GOPEX was conducted on all but one evening between December 9 and 16 December, 1992. No operations were conducted on December 13, hence the data are reported according to days 1–4 and 6–8. On each night, pulsed lasers from TMF and SOR were directed toward the SSI camera of the Galileo spacecraft. During laser transmission, the scan platform of the SSI was activated (at approximately 6 mrad/sec) to spread the laser pulses in proportion to the telescope scan rate and the laser repetition rate. The TMF laser operated at 15 Hz for days 1 and 2, and at 30 Hz for the remainder of the experiment. SOR transmitted at 10 Hz during the days that laser pulses were detected. A simulated result for day 1 is shown in Fig. 1, as generated by MIPL. The two lines of pulses correspond to each transmission site. The distance between the lines is proportional to the distance between the two transmitting sites and the range to the spacecraft. The contrast in the figure has been digitally enhanced to exaggerate the relationship between the transmitting sites and the terminator, the location on the Earth in the shadow of the Sun's illumination. The image in Fig. 1 is displayed to correspond geographically with the more eastern SOR site, which is closer to the terminator. The distance between the transmitting sites and the terminator also changes in relation to the spacecraft range. The smearing of the terminator will influence the measurement of the laser pulse from its neighboring background illumination. This became a more significant factor in the later days of the experiment with a weaker laser signal and when the relatively large range placed both transmitters near the terminator and correspondingly closer to high background levels. Deviations in the laser pulse positions from a straight line were simulated from expected jitter of the scan platform. For the most part, the presence of laser pulses was identified with certainty by their separation from neighboring pulses. The simulated image served as the basis for all the image-analysis programs developed prior to the GOPEX demonstration.

[1] B. M. Levine, "GOPEX Data Products Summary," JPL Interoffice Memorandum (internal document), Optical Sciences and Applications Section, Jet Propulsion Laboratory, Pasadena, California.

III. Image Acquisition and Analysis

A. MIPL Image Acquisition and Processing

All GOPEX images were received at the MIPL according to a predetermined implementation plan.[2] During the first two days, the images were relayed to a real-time display from which identification of laser pulses could be made. Data display for the rest of the experiment was performed as soon as the images were played back from spacecraft storage. Images were modified by MIPL for pulse identification purposes during the first two days, and also for public release images. All data used to measure pulse energies were transmitted unmodified as 8-bit binary numbers for each pixel value. The disadvantage is that corrections for dark counts, responsivity uniformity, and individual pixel defects were not performed. The major reason for not correcting the GOPEX data was that the files used to make such corrections were out of date and would not be revised until significantly after GOPEX.[3]

Examples of raw and processed images are given in Fig. 2. The raw image contained numerous Reed–Solomon (RS) coding errors in the first two days of GOPEX. These errors appeared as horizontal bars across the frame and complicated a predetermined automated algorithm based on thresholding pixel values above the mean of each vertical column. The incidence of these errors practically disappeared after day 2, however, the intermittent incidence of errors forced the visual inspection of each frame to identify laser pulses throughout the remainder of the experiment.

B. Image Analysis

A set of programs to measure the intensity of all laser pulses was written in MATLAB [2], a commercial software

[2] L. Wanio, "SDT-MIPS GOPEX Implementation Plan–Version 2," JPL Interoffice Memorandum PA 6-384-92-LAW66.law (internal document), Jet Propulsion Laboratory, Pasadena, California, November 23, 1992.

[3] L. Wanio, private communication, Image Processing Applications and Development Section, Jet Propulsion Laboratory, Pasadena, California, November 1992.

package for numeric computation and data visualization. Program input consisted of a list of coordinate locations which specified the approximate peak of each pulse. One program would automatically find the centroid of the pulse and then would sum the data number (dn) values about a 5×5 window from the centroid. This size window was chosen because it encompassed the entire signal pulse found in the data simulations. Background would be subtracted by an average background 5×5 window which was computed from windows immediately to the left and right of the window containing the laser pulse data. It was originally desired to form the background window from data above and below the laser pulse window; however, the presence of coding error lines too frequently corrupted the result. In practice, signals above the background enclosed a 3×3 window because of the added noise in the actual data frames. A smaller window for adding all laser signals would have increased detection sensitivity. Additive measurement errors occur when a background measurement is subtracted from data containing both signal and background; thus, minimizing the window size would also minimize the accumulated measurement errors. After all calculations were made, the data were automatically separated into lists representing the two transmitter sites sorted by pixel location, and saved onto a file containing image analysis from all the day's image frames.

The result of such a calculation for the data illustrated in Fig. 2 is given in Table 1. The pulses attributed to SOR are uniformly spaced. The footnoted pulse numbers were added through calculations made independently of the computer program. In the case of TMF pulse 2, a modified value was substituted due to a line of dn's of 255 which corrupted the data. Originally the distance between pulse 5 and pulse 6 was 112.5 pixels, roughly equal to the distance between 3 laser pulses, as no pulses were visually observed on the thresholded image. Pulses 5.1 and 5.2 were added to the list from a hand calculation of the data about the expected location of the pulses.

For the exposure time of 800 msec, one would expect more pulses to be detected, up to 8 from SOR and up to 15 from TMF. It was suspected that the scan platform movement was not synchronized with the laser transmission times, which resulted in the loss of a number of pulses on each frame. Because of lack of correspondence between the exposure time and the expected number of laser pulses, it was not possible to correlate locations on the frame where no detection was recorded. This is important for two reasons. First, single detection events near a detection threshold cannot be reliably determined. This results in the under-reporting of pulses during the experiment. Second, there is a finite probability that the entire signal could be lost due to turbulence. Hence, the total number of zeroes is an important statistic in fitting histogram data to hypothesized probability distributions. The only certain way in which zero intensities could be determined is if they were to occur between two unambiguously identified laser pulses, as shown in the data in Table 1.

The other major program took a vector of intensity values for each day of GOPEX and plotted a histogram of dn values together with the expected lognormal probability distribution using parameter values determined from the analysis given in the next section, and also from the actual data. The results of all the histograms shown in this article were made by this program.

IV. Theoretical Development

The analytical model assumes that the optical beam possesses a Gaussian profile and the communication channel has lognormal scattering characteristics. The atmosphere-induced jitter is modelled as two independent zero-mean Gaussian random variables. By modelling the system parameters as a set of independent and identically distributed (iid) random variables, the combined impact of uncertainties due to system parameters and the turbulent atmosphere is approximated by a lognormal distributed signal intensity at the spacecraft [3].

The mean number of photoelectrons distributed over the pixels of the SSI camera has been computed for the laser uplink by Kiasaleh and Yan [3]. The corresponding lognormal variance, however, is not given, and must be computed separately from their analyses. The two theories given below provide the required lognormal variance. The first theory follows the work of Tatarski and is applicable only in the limit of weak turbulence, that is for a lognormal variance much less than unity. The seeing conditions at TMF during GOPEX were measured within this limit only on one of the observing nights. The second theory developed is applicable under conditions of strong atmospheric turbulence. It relies on statistically describing the effect of laser intensity fluctuations as a large number of strong independent disturbances on the received photoelectron count measured by the SSI. This limiting behavior was observed at TMF on all nights but one.

A. Weak Turbulence Theory

The lognormal variance taken from Tatarski [4] is given by

$$\sigma_N{}^2 = 2.24k^{-7/6}\sec(\theta)^{11/6}\int_0^z C_n^2(h)h^{5/6}dh \qquad (1)$$

where $k = 2\pi/\lambda$ is the wave number, θ is the zenith angle, $C_n^2(h)$ is the atmospheric structure constant, and z is the height of the atmosphere. Since no measurements of the atmospheric structure constant were made during the experiment, the above equation cannot be used to obtain good estimates of lognormal variance. However, rough estimates of seeing on the days when the GOPEX demonstration was conducted are available, and it is possible to rewrite Eq. (1) to make use of the seeing estimates. Assuming a zenith height, z_0, for the atmosphere, with $C_n^2(h)$ a constant, Eq. (1) simplifies to

$$\sigma_N{}^2 = 10.3\lambda^{-7/6} z^{11/6} C_n^2 \tag{2}$$

where $z = z_0 \sec(\theta)$. An expression for r_0, the atmospheric coherence parameter, is given below [5]:

$$r_0 = 0.185\ \lambda^2/(zC_n^2)^{\ 3/5} \tag{3}$$

Note that the seeing, α, can be simply related to the zenith atmospheric coherence length by the relation $\alpha = 4\lambda/(\pi r_o)$.[4] Eliminating C_n^2 from Eqs. (2) and (3), one gets

$$\sigma_N{}^2 = 0.4(z_0/\lambda)^{5/6} [\sec(\theta)\alpha^2]^{5/6} \tag{4}$$

Using this expression and assuming $z_0 = 7$ km with $\lambda = 0.532\ \mu$m, Eq. (4) can be further reduced to

$$\sigma_N^2 = 1.0 \times 10^8\ \sec(\theta)\alpha^{2\ \ 5/6} \tag{5}$$

Equation (5) calculates lognormal variance in terms of the seeing (in radians) and the zenith angle. This equation can now be used to compute log-variance from the estimated seeing for the GOPEX demonstration days. The lognormal mean, m_n, of the distribution is given by the expression

$$m_n = \ln < N_{tot} > \ - 1/2\sigma_N^2 \tag{6}$$

where $< N_{tot} >$ is the expected number of photoelectrons at the detector which can be calculated using the equations developed by Kiasaleh and Yan [3].

[4] K. S. Shaik, JPL Interoffice Memorandum 331.6-91-191 (internal document), Communications Systems Research Section, Jet Propulsion Laboratory, Pasadena, California, September 3, 1991.

B. Strong Turbulence Theory—Photoelectric Count Probability Distribution

Consider the charge-coupled device (CCD) on the SSI as a detector that converts incident photons into counts of photoelectrons. The amplifiers and digitizing circuitry that follow convert the photoelectron count into a dn value, depending on its system parameters. For the SSI, the photoelectron count can be associated to the dn by a gain state, g, which can be set over 4 different values. Specifically, the dn is related to the photoelectron count, $< N_{tot} >$, by the multiplicative relation

$$dn = gN_{tot} \tag{7}$$

Thus, photoelectron counts are discussed interchangeably with dn values in the following discussion. Statistical errors from the quantization noise introduced by Eq. (7) have been ignored.

A photoelectric detector, such as a CCD, generates counting statistics which obey the Poisson probability law, and furthermore the parameter of the Poisson distribution is proportional to the light intensity. Thus, the probability of detecting n photoelectrons obeys Mandel's equation (also known as the Poisson transform) to give the relation [6]:

$$p(N_{tot}) = \int_e^{-1} I_{tot}^N / N_{tot}! p(I) dI \tag{8}$$

For constant laser illumination, $p(I)$ degenerates into a delta function, leaving the familiar Poisson distribution for photoelectric counts. In GOPEX there are fluctuations in laser intensity from frame to frame[5] as well as from the atmosphere. Hence, the random nature of the photoelectron counts must include this second random factor. For any $p(I)$, intensity moments are related to those of photoelectron counts by the identities

$$< I > = < N_{tot} > \tag{9}$$

$$< I^2 > = < N_{tot}(N_{tot} - 1) > \tag{10}$$

$$\sigma_I^2 = < I^2 > - < I >^2 = \sigma^2{}_{N_{tot}} - < N_{tot} > \tag{11}$$

[5] G. Okamoto, "GOPEX Monitor Data Analysis," JPL Interoffice Memorandum 331.6-93-032 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 3, 1993.

The corresponding lognormal variance based on the photoelectron counts is determined by the equations

$$\sigma_N{}^2 = \ln(1 + \sigma_I^2 / < I >^2) \qquad (12)$$

with the lognormal mean following Eq. (16).

The most general probability distribution for the intensity fluctuation is the exponential distribution, and it represents the limiting distribution of modulation from the combination of all sources. Kiasaleh and Yan also recommend using this distribution in the presence of strong turbulence. In this limit, the photoelectrons follow the Bose–Einstein distribution whose mean and variance are

$$< N_{tot} > = \mu \qquad (13)$$

$$\sigma^2_{N_{tot}} = \mu^2 + \mu \qquad (14)$$

Using Eq. (12), the lognormal variance, σ_I^2 is always ln(2) and holds for all values of atmospheric turbulence.

There are further multiplicative factors, such as random atmospheric transmission, and other system parameters (see [3, Section II]) which could further broaden the distribution of photoelectric counts. An analysis by Saleh [7] produces an approximate solution for the detected light intensity after the signal has passsed through a channel with lognormally distributed disturbance characteristics. Its evaluation is beyond the scope of the analysis in this article, however, the exponential approximation should be adequate by assuming constant systematic conditions.

V. GOPEX Results and Discussion

Table 2 provides a summary of site operational conditions and data-derived results at TMF for GOPEX. The former includes the laser repetition rate, its beam divergence, range to Galileo, atmospheric seeing, and the gain state of the SSI. The latter includes the number of pulses detected (including pulses of zero intensity), $< N_{tot} >$, and the lognormal parameters, σ_I and m_n (using both theories). There appears to be good agreement between the data and the strong turbulence theory on day 1. This histogram of data pulses is plotted in Fig. 4. The error bars on the histogram heights show a relative range of probability that one would expect if another GOPEX data set, taken under the same conditions, were to be obtained and similarly displayed. The ± 1 standard-deviation error bars on the histogram are relatively wide due to the limited number of pulses available. As shown in Fig. 4, there is also good agreement between the data and the weak turbulence theory on day 2, the only day its assumptions were applicable in GOPEX. The weak turbulence approximation curve is close to most of the observed frequencies of pulses passing through most of the error bars of its histogram. The other histograms for TMF, Figs. 5–8, show varying agreement. On days 3 and 8, there is good agreement between the computed strong theory lognormal parameters and the data histograms. On day 6, the theory shows a good quantitative fit, but the lognormal fit due to the data does not agree as well. On day 7, there appears to be a bimodal histogram, which would indicate changing conditions during the experiment. For day 8, both histograms show a lack of fit near the smallest data bin. As mentioned in Section III.B, small values and zero values of the laser pulses were difficult to estimate, which probably downwardly biased this particular bin and artificially inflated the percentages of pulses in the other bins.

Table 3 similarly provides a summary of data and results for SOR. Note that the SOR seeing measurements are all in the strong turbulence regime and, consequently, weak theory estimates for the lognormal standard deviation and mean do not apply. The results are displayed in Figs. 9–12. One reason for the large differences between the data and the calculated lognormal moments differences is the lack of data. The best agreement between data and calculation was on day 4, and its histogram is shown below in Fig. 12. Again, there is an underestimation of the first histogram bin.

VI. Summary

Laser pulse data from the GOPEX demonstration have been analyzed for pulse strength and pulse distribution. Procedures were developed in both the weak turbulence regime and in the strong turbulence regime to fit the observed histogram of received pulses to a lognormal distribution. Both procedures were shown to quantitatively fit the data for a subset of data. Systematic differences between the data and lognormal fits obtained using atmospheric structure constant values computed from observatory seeing conditions and SSI camera conditions were attributed to the difficulty in determining small values and zero values of laser pulse energy measured by the SSI camera. Also, the use of nominal values of some attributes (like seeing, for example) as input parameters to the theoretical models may not have been representative of actual conditions. Pulse strengths could be measured with greater confidence if the relative location of pulses were

better known. The timing mismatch between the scan platform movement and the onset of laser pulse transmission deprived the analysis team of important information needed to determine the presence or absence of each pulse. Experimental sensitivity could also be improved through simple SSI camera calibrations. GOPEX was prevented from using such calibrations due to the lack of current data. At MIPL, extensive on-line software exists which uses the GALSOS program to perform this task once the calibration files are updated. The quality of the pulse intensity measurements could be improved, and possibly a greater number of data points could be obtained with the existing GOPEX image data set.

Acknowledgments

The authors are greatly indebted to the Multimission Image Processing Laboratory facilities and staff in general. In particular, the authors thank Lisa Wainio for providing staff for MIPL programming support and for providing the simulated images, and to Danika Jensen for her support during GOPEX operations and for providing all the images used in the analysis. The authors also acknowledge discussions with Keith Wilson, Jim Lesh, and Jim Anderson on data analysis, and the support of Gregory Wanish for his help with critical parts of the MATLAB programming.

References

[1] K. E. Wilson and J. R. Lesh, "An Overview of the Galileo Optical Experiment (GOPEX)," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 192–204, August 15, 1993.

[2] MATLAB, Natick, Massachusetts: MathWorks, Inc., 1992.

[3] K. Kiasaleh and T.-Y. Yan, "A Statistical Model for Evaluating GOPEX Uplink Performance," *The Telecommunications and Data Acquisition Progress Report 42-111, vol. July–September 1992*, Jet Propulsion Laboratory, Pasadena, California, pp. 325–332, November 15, 1992.

[4] V. I. Tatarski, *Wave Propagation in a Turbulent Medium*, New York: McGraw Hill, 1961.

[5] J. W. Goodman, *Statistical Optics*, Chapter 8, New York: John Wiley and Sons, 1985.

[6] J. W. Goodman, *Statistical Optics*, Chapter 9, Sections 9.1 and 9.2, New York: John Wiley and Sons, 1985.

[7] B. E. A. Saleh, *Photoelectron Statistics*, Section 5.2, p. 242, New York: Springer-Verlag, 1978.

Table 1. Data reduction for GOPEX data frame 8 on day 2.

Results for file = Day 2/sec 0165359600.1

Transmitter location = SOR

Pulse number	Peak position		Centroid position		Distance	Background signal	Corrected dn in 25 pixels	Comments
	Row	Col.	Row	Col.	Pixel	dn per pixel		
1	557	325	556.91	325.01	0.0	23.9	352.0	—
2	614	325	613.96	325.01	57.1	22.9	168.5	—
3	671	325	670.99	324.97	57.0	21.7	247.5	—
4	729	326	727.76	325.78	56.8	75.8	409.0	—
5	785	325	785.45	324.98	57.7	34.8	100.0	—

Transmitter location = TMF

Pulse number	Peak position		Centroid position		Distance	Background signal	Corrected dn in 25 pixels	Comments
	Row	Col.	Row	Col.	Pixel	dn per pixel		
1	510	377	510.66	377.00	0.0	64.2	21.0	—
2	547	377	548.07	377.06	37.4	59.1	196.5	—
2[a]	—	—	—	—	—	21.2	44.0	Replace above
3	585	377	585.02	376.99	36.9	20.1	28.0	—
4	623	377	622.99	376.97	38.0	19.7	45.5	—
5	660	377	660.02	376.98	37.0	19.1	21.5	—
5.1[b]	698	377	—	—	38	18.2	28.0	On 2×3 between RS[c] errors
5.2[b]	736	377	—	—	38	17.0	16.0	On 2×3 between RS errors
6	774	377	772.54	376.99	112.5	60.9	36.0	Signal okay

[a] Replace background signal and corrected signal in 2.

[b] Pulse found by manual inspection.

[c] RS = Reed Solomon.

Table 2. GOPEX data results summary at TMF on days data were received.

Parameter	Day 1	Day 2	Day 3	Day 4	Day 6	Day 7	Day 8
Site and spacecraft conditions							
Repetition rate, pulses/sec	15	15	30	—	30	30	30
Beam divergence, μrad	110	110	110	—	60	60	60
Average range, x 10^6 km	0.7	1.45	2.21	—	4.48	5.25	6.01
Average zenith angle, deg	50	51	53	—	55	56	56
Seeing due to turbulence, μrad	20	7.5	20	—	40	20	30
SSI camera gain state, electrons/dn	391	391	156	—	156	156	156
Data products							
Data frames received	60	40	20	—	12	10	8
Number of frames with pulse detections	10	16	5	—	3	6	6
Total received pulses	33	59	41	—	24	60	51
Expected photoelectrons/pulse, μ	67,116	14,860	7267	—	4352	3295	2629
Lognormal mean, m_l (data)	10.86	9.27	8.45	—	7.70	8.36	7.41
Lognormal mean, m_l (strong theory)	10.83	—	8.50	—	8.03	7.75	7.36
Lognormal mean, m_l (weak theory)	—	9.40	—	—	—	—	—
Lognormal standard deviation, σ_N (data)	0.7263	0.5356	0.940	—	1.047	0.8060	0.8619
Lognormal standard deviation, σ_N (strong theory)	0.8326	—	0.8326	—	0.8326	0.8326	0.8326
Lognormal standard deviation, σ_N (weak theory)	—	0.6300	—	—	—	—	—

Table 3. GOPEX data results summary at SOR on days data were received.

Parameter	Day 1	Day 2	Day 3	Day 4	Day 6	Day 7	Day 8
Site and spacecraft conditions							
Repetition rate, pulses/sec	10	10	10	10	—	—	—
Beam divergence, μrad	80	80	80	10	—	—	—
Average range, x 10^6 km	0.7	1.45	2.21	4.48	—	—	—
Average zenith angle, deg	50	51	53	54	—	—	—
Seeing due to turbulence, μrad	No data	13.5	13.5	17.0	—	—	—
SSI camera gain state, electrons/dn	391	391	156	156	—	—	—
Data products							
Data frames received	60	40	20	10	—	—	—
Total received pulses	16	43	12	8	—	—	—
Expected photoelectrons/pulse, μ	74,818	17,178	7233	3447	—	—	—
Lognormal mean, m_n (data)	10.49	10.64	8.70	8.07	—	—	—
Lognormal mean, m_n (strong theory)	10.88	9.41	9.23	7.80	—	—	—
Lognormal mean, m_n (weak theory)	—	9.40	—	—	—	—	—
Lognormal standard deviation, σ_N (data)	1.0153	0.7360	0.9613	0.8505	—	—	—
Lognormal standard deviation, σ_N (strong theory)	0.8326	0.8326	0.8326	0.8326	—	—	—
Lognormal standard deviation, σ_N (weak theory)	—	—	—	—	—	—	—

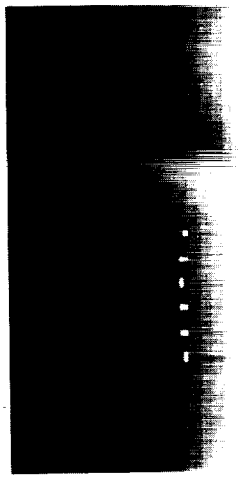


Fig. 1. A portion of a simulated image for day 1 of GOPEX.

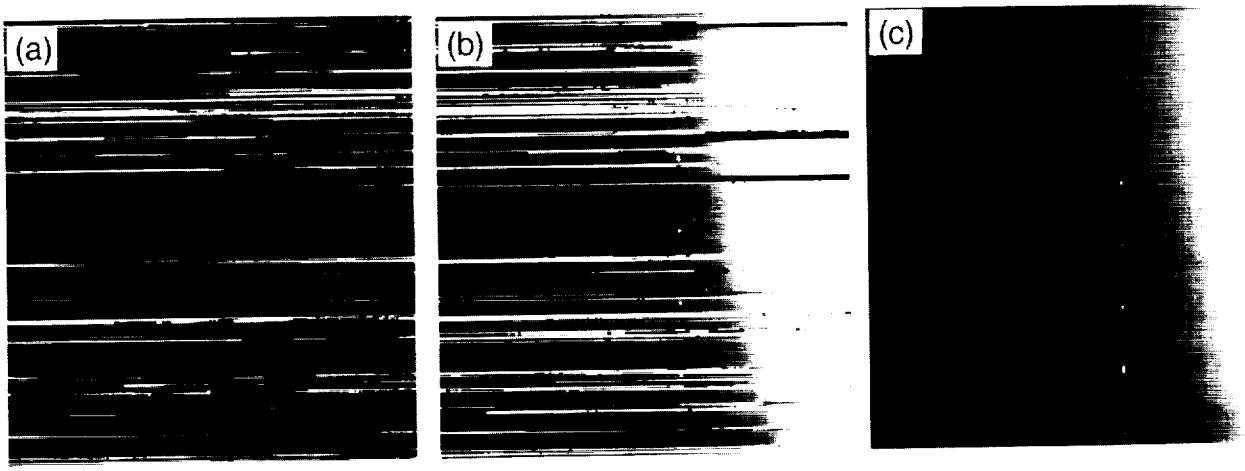


Fig. 2. Raw and processed GOPEX images for day 2, frame 8: (a) raw 8-bit image; (b) contrast stretched image; and (c) MIPL-processed image.

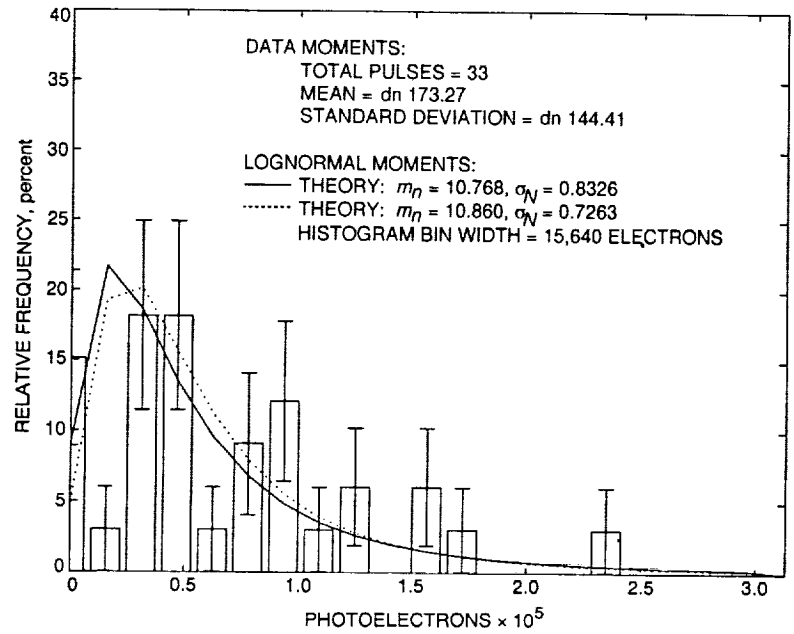


Fig. 3. Histograms for TMF on day 1 of GOPEX. Using the lognormal model, strong turbulence theory parameters agree well with the data parameters.

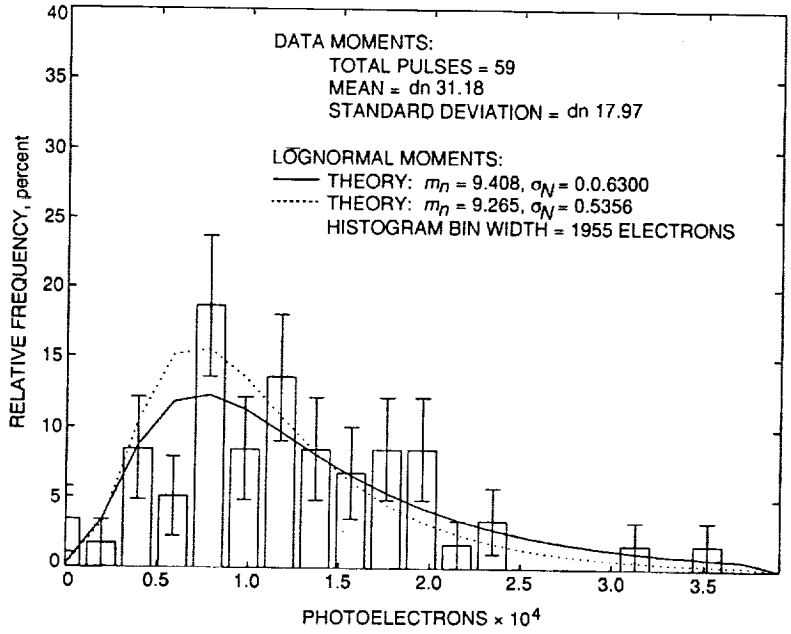


Fig. 4. Comparison of weak turbulence theory for day 2 of GOPEX, at TMF; histogram fit using weak turbulence theory. The error bars represent ±1 standard deviation.

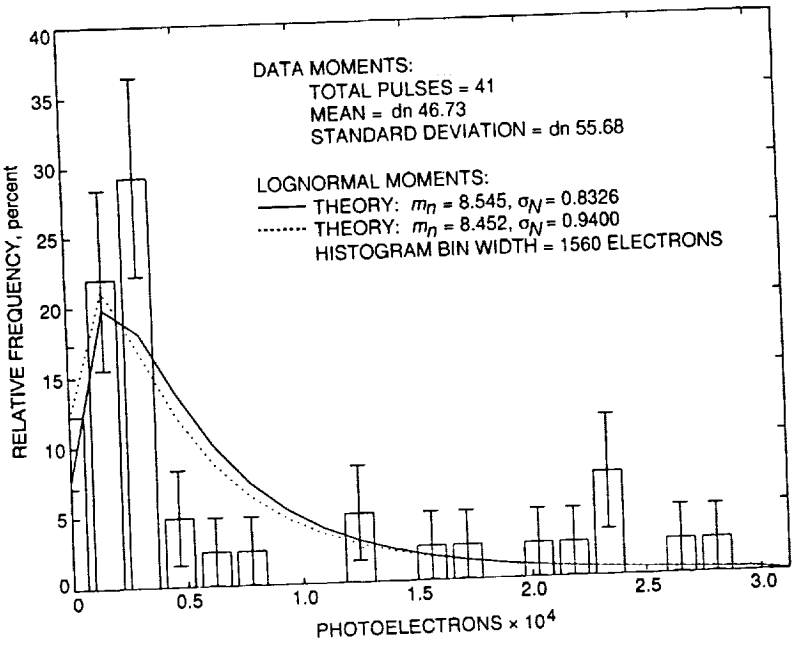


Fig. 5. Histograms for TMF on day 3 of GOPEX.

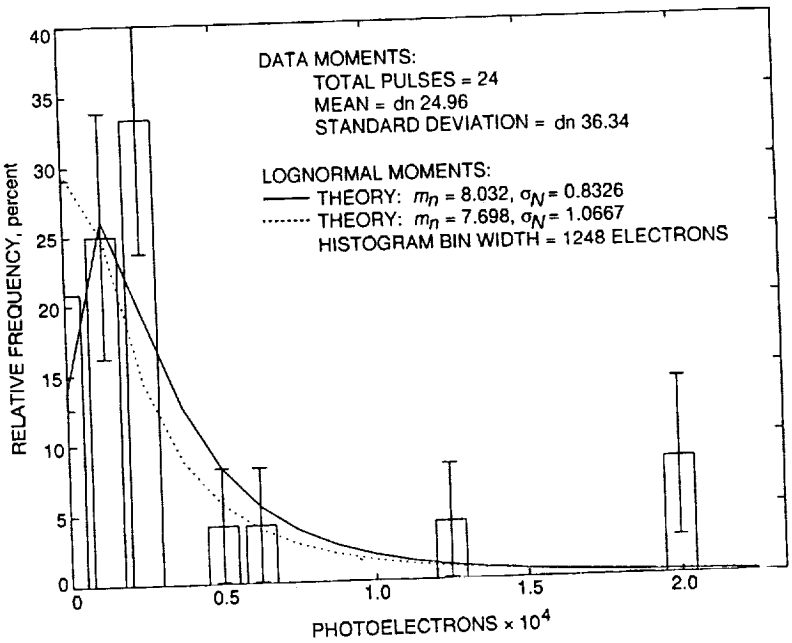


Fig. 6. Histograms for TMF on day 6 of GOPEX.

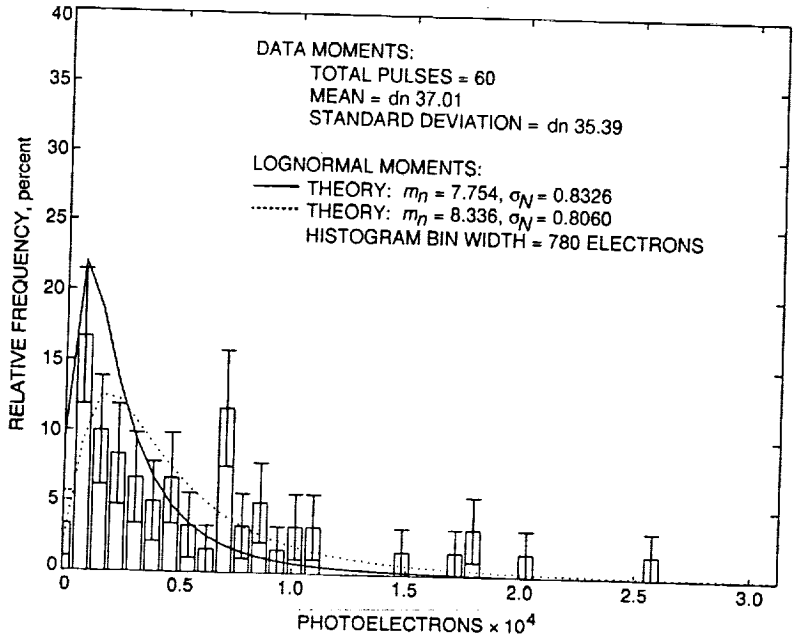


Fig. 7. Histograms for TMF on day 7 of GOPEX.

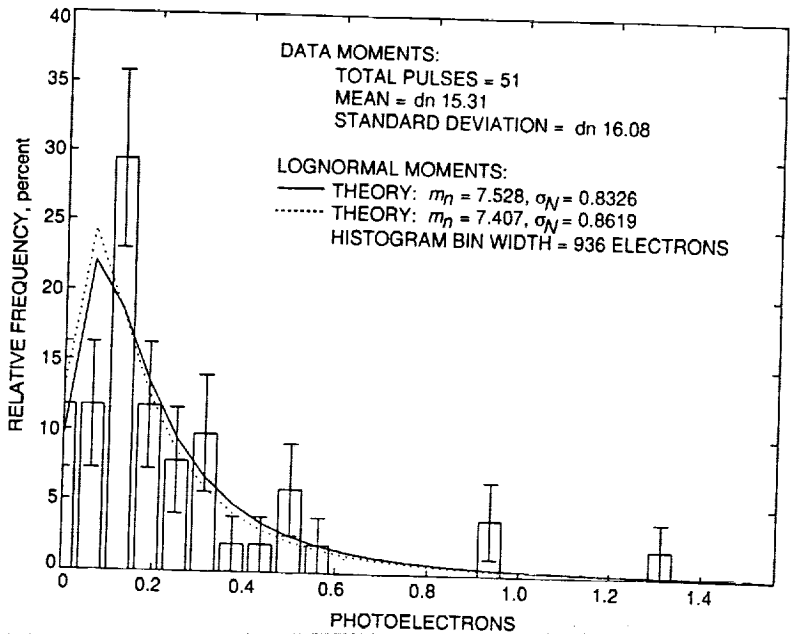


Fig. 8. Histograms for TMF on day 8 of GOPEX. There is good agreement between the data and the lognormal distribution using strong turbulence theory parameters.

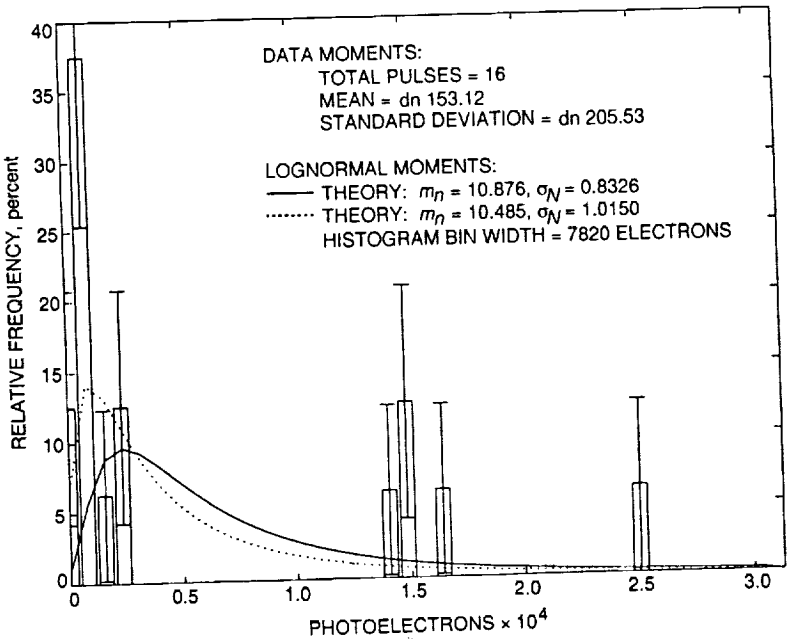


Fig. 9. Histograms for SOR on day 1 of GOPEX.

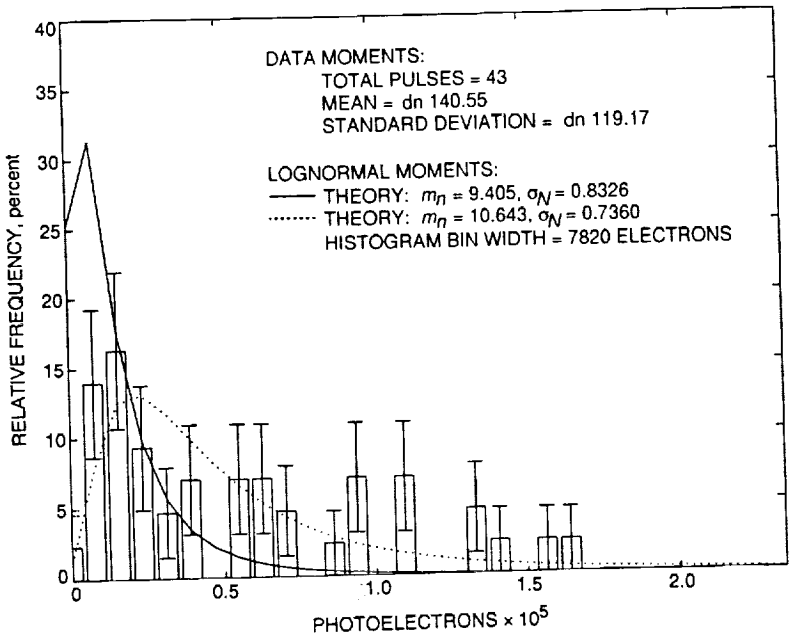


Fig. 10. Histograms for SOR on day 2 of GOPEX.

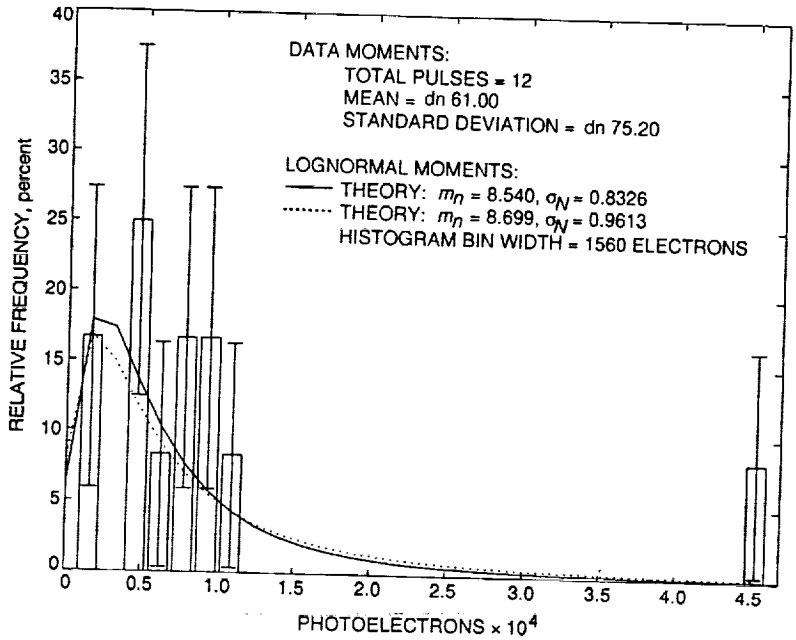


Fig. 11. Histograms for SOR on day 3 of GOPEX.

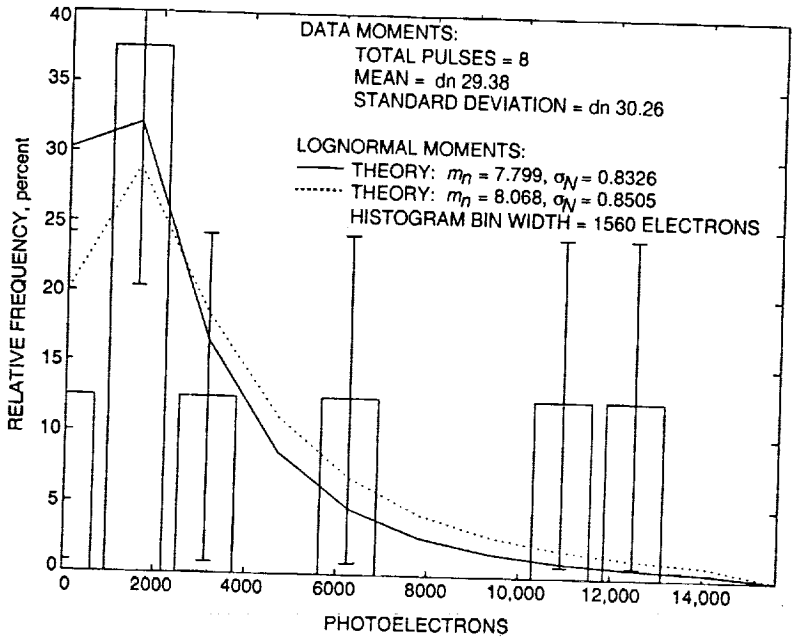


Fig. 12. Histograms for SOR on day 4 of GOPEX.

Appendix

Figures A-1 and A-2 are MATLAB m-files written to plot the observed frequency histogram with its error bars and to compare the data to the lognormal distribution with parameters given by either the weak or strong turbulence theories. The first m-file evaluates the lognormal probability density function, and the second m-file compiles the histogram, error bars, and directs the plotting.

The MATLAB m-file for computing the lognormal probability density function for a vector of photoelectric counts given by x is presented in Fig. A-1.

The MATLAB m-file for plotting the histogram and its related error bars with the lognormal probability density function using parameters derived from either the data or theory is presented in Fig. A-2.

```
function fx=log_norm(mu,sigma,x)
%bml 21 dec 92
%evaluates the lognormal distribution function over the vector x (xmin>0)
%parameters of the log normal are mu and sigma
%
fx=[];
n=size(x,2);
fact=sqrt(2.*pi).*sigma;
fact2=2*sigma*sigma;
mean=mu.*ones(1,n);
arg=fact.*x;
arg2=(log(x)-mean).*(log(x)-mean)./fact2;
fx=exp(-arg2)./arg;
```

Fig. A-1. The MATLAB m-file for computing the lognormal probability density function for a vector of photoelectric counts given by *x*.

```
function des_stat=histo_theory2(freq,xmitr,day,deltab,bmax,gain,mu,sigma)
% output file names fileroot=[day, '.',xmitr,'.'];
% descriptive statistics small=min(freq);
large=max(freq);
avg=mean(freq);
med=median(freq);
stdev=std(freq);
%
%histogram plot
%assumes a vector of intensities
bins=0:deltab:bmax;
[count location]=hist(freq,bins);
%normalize counts total=sum(count);
ndata=total;
count=100.*count./total;
%plot results bar(location,count);
axis([0 bmax 0 60]);
xlabel('DN');
ylabel('relative frequency (%)');
title_string=sprintf('GOPEX encircled energy at %s; %s--bins=%2i DN, Total=%3i
pulses',xmitr,day,deltab,ndata); title(title_string);
%keyboard
print_command =['print ',fileroot,'DN.ps'];


draft filed as: /var8/gopex/memos/tda-image.report
last revision: February 24, 1993 1:55 pm
eval(print_command)
%
%theory lognormal pdf
delta=deltab*gain;
cmax=gain*bmax;
counts=gain.*freq;
bins=0:delta:cmax;
[count location]=hist(counts,bins);
%normalize counts total=sum(count);
count=100.*count./total;
%plot results
x=1:cmax;
pdf1=log_norm(mu,sigma,x);
%integrate over bin width
nbins=floor(cmax/delta);
for i=1:nbins
  pdf(i)=100*sum(pdf1((i-1)*delta+1:i*delta));
end
```

Fig. A-2. The MATLAB m-file for plotting the histogram and its related error bars with the lognormal probability density function using parameters derived from either the data or theory.

```
pdf=[pdf'; 100*sum(pdf1(nbins*delta+1:cmax))]';
%sample lognormal pdf
sigma2=sqrt(log(stdev*stdev/avg/avg+1));
mean2=log(avg*gain)-sigma2*sigma2/2;
pdf1=log_norm(mean2,sigma2,x);
%integrate over bin width
nbins=floor(cmax/delta);
for i=1:nbins
  sample_pdf(i)=100*sum(pdf1((i-1)*delta+1:i*delta));
end
sample_pdf=[sample_pdf'; 100*sum(pdf1(nbins*delta+1:cmax))]';
%plot results
ebar=100.*sqrt(count.*(100.-count)./10000./total);
errorbar(location,count,ebar);
hold on;
plot(location,count,'k');
bar(location,count);
axis([0 cmax 0 40]);
xlabel('photoelectrons');
ylabel('relative frequency (%)');
%keyboard
title_string=sprintf('GOPEX encircled energy for %s on %s',xmitr,day); title(title_string);
plot(bins,pdf);
plot(bins,sample_pdf,':');
string1=sprintf('data moments:  total pulses = %3i', ndata);
string2=sprintf('mean     = %5.2f DN  std dev  =%5.2f DN',avg,stdev);
string3=sprintf('log normal moments');
string4=sprintf('theory:  mN  = %6.3f  sigmaN = %7.4f',mu,sigma);
string5=sprintf('data   :  mN  = %6.3f  sigmaN = %7.4f',mean2,sigma2);
string6=sprintf('histogram bin width = %g electrons',delta); x=.4*cmax;
text(x,39,string1);
text(x,37,string2);


draft filed as: /var8/gopex/memos/tda-image.report
last revision: February 24, 1993 1:55 pm
text(x,34,string3);
plot ([x-.05*cmax; x-.005*cmax],[32; 32]);
text(x,32,string4);
plot ([x-.05*cmax; x-.005*cmax],[30; 30],':');
text(x,30,string5);
text(x,28,string6);
% output histogram
print_command =['print ',fileroot,'counts.ps'];
eval(print_command)
%print counts.ps
hold off;
%save day2.histogram.mat bins freq
des_stat=[small; large; med; avg; stdev];
```

Fig. A-2. (Cont'd)

TDA Progress Report 42-114

August 15, 1993

Telescope Pointing for GOPEX

W. M. Owen, Jr.
Navigation Systems Section

In order for photons emitted by the GOPEX lasers to be detected by Galileo's camera, the telescopes at Table Mountain Observatory and Starfire Optical Range had to be pointed in the right direction within a tolerance less than the beam divergence. At both sites nearby stars were used as pointing references. The technical challenge was to ensure that the transmission direction and the star positions were specified in exactly the same coordinate system; given this assurance, neither the uncertainty in the star catalog positions nor the difficulty in offset pointing was expected to exceed the pointing error budget. The correctness of the pointing scheme was verified by the success of GOPEX.

I. Introduction

The GOPEX experiment [1] depended critically upon accurate pointing of the transmitting telescopes at Table Mountain Observatory (TMO) and Starfire Optical Range (SOR), for if a telescope had been pointed in the wrong direction, its laser emission would not have reached Galileo. Accurate pointing requires accurate spacecraft and transmitter ephemerides, correct algorithms for calculating the transmission direction, and reliable procedures at the telescope to point in the desired direction. The desired accuracy was determined by the beam divergence of the lasers (minimum 60 μrad or 12 arcsec) and atmospheric seeing (several arcsec); the GOPEX project assigned an error budget of 5 arcsec for pointing.

Tests of the TMO 24-in. (0.61-m) telescope mount, both hardware and software, indicated that one could not point at an arbitrary direction in the sky without introducing errors larger than GOPEX's tolerance. The causes of this error remain unknown; possible explanations include unmodeled mount misalignment, flexure, or mismodeled atmospheric refraction. Whatever its source, the absolute pointing error ruled out using absolute pointing.

Relative pointing, however, proved much more reliable. One can point the telescope at one star and then offset to another nearby star with an error of only a few arcseconds, provided that the two stars are separated by less than one or two degrees. This behavior suggested that the best way to point the telescope for GOPEX was to point first at a star close to Galileo, then to move from the star to Galileo. Therefore, GOPEX required a catalog of suitable reference stars surrounding Galileo's position during the experiment.

The remaining sections of this article present in detail the various aspects of the pointing problem: predictions of the uplink direction, the actual pointing of the telescopes, and the development of the reference star catalog.

II. Pointing Prediction

The "raw material" for predicting pointing consists of a planetary ephemeris file, a spacecraft ephemeris file (known as a P-file), the Earth-fixed coordinates of the telescope, and the Earth orientation model. Updated P-files were received just prior to the first day of the experiment

and again after three days; each contained more tracking data and was, therefore, more precise than its predecessor.

The coordinates for the uplink direction were found as follows:

(1) A P-file was received from the Galileo Navigation Team and converted, using the standard navigation utility program PVTOEXP, from the B1950 coordinate system to the J2000 coordinate system. Since a P-file by construction uses an inertial reference frame, this transformation is accomplished by a simple 3×3 rotation matrix applied to both positions and velocities.

(2) The position of Galileo at the midpoint of each Solid-State Imaging (SSI) camera exposure was found relative to the barycenter of the Solar System by interpolating and adding vectors from the P-file and from the DE202 planetary ephemeris.

(3) The position of the telescope, also relative to the Solar System barycenter, was similarly found at the midpoint of the *transmission* window. The one-way light time, and therefore the position, were determined iteratively. The Earth was oriented using the J2000 precession model [2], nutation angles from the DE202 planetary ephemeris, and timing and polar motion angles from the P-file. For GOPEX, the light time amounted to just over two seconds on the first day and about 20 seconds on the last day.

(4) The difference between the above two vectors gives the *true* direction referred to the J2000 system. The *transmission* direction is derived from this by removing the effect of classical stellar aberration: one subtracts $\mathbf{v}/c$ from the true unit vector, where $\mathbf{v}$ is the Solar System barycentric velocity of the telescope and c is the speed of light. This correction accounts for the fact that the outgoing photons are displaced forward by the transmitter's velocity, as seen by an observer at the Solar System barycenter. For GOPEX, since Galileo's motion was almost exactly radially outward relative to the Earth, the effects of light time and aberration almost exactly cancel. This happy situation will not always hold, however, so both effects were included in the software.

(5) The transmission vector just found is in *apparent* coordinates in the J2000 system. The *mean* position of a star as given in a catalog, however, is the direction from which an observer at the Solar System barycenter would observe the starlight. Any other observer would see the starlight emanating from a different direction, the *apparent* direction, due again to stellar aberration. One must, therefore, either compute the apparent coordinates of the reference stars or else transform the transmission direction to match the mean positions of the stars. The latter option, to produce *astrographic* coordinates, was chosen for the sake of simplicity, to avoid having the star coordinates change as the telescope's velocity changes. One therefore subtracts $\mathbf{v}/c$ again from the unit vector; this reverses the change in position from astrographic to apparent due to the motion of the observer.

(6) The telescope mount is nominally aligned with the true north pole of date. In order that computed offset directions may match those commanded to the telescope, one precesses [2] the above vector into *mean-of-date* coordinates with the epoch chosen to be the midpoint of the GOPEX experiment. Choosing a fixed epoch again keeps the star positions constant over time while having a negligible effect on the direction to celestial north.

The above procedure, performed for every transmission, resulted in a file of transmission directions that were in the same coordinate system as the star catalog. A plot of the transmission directions for TMO appears in Fig. 1. The longest line represents the first day of the GOPEX experiment.

Pointing predictions using the above algorithm were also produced for SOR but not used there, as SOR had its own prediction capability. SOR's geocentric algorithm included the following steps:

(1) JPL provided SOR with a list of positions and velocities of Galileo relative to the Earth, referred to J2000 coordinates, at the midpoint of each SSI exposure. The vectors were interpolated directly from the geocentric P-file.

(2) SOR rotated each vector into Earth-fixed true-of-date coordinates by applying precession, nutation, and sidereal time matrices, these angles being computed for the uplink time.

(3) The direction from SOR to Galileo was found in true-of-date coordinates by vector subtraction. This was taken to be the transmission direction.

This algorithm, while not rigorously correct, gave answers that agreed with JPL's predictions to within a fraction of an arcsecond. Using a different algorithm at each site helped to guard against error at one site or the other.

III. Telescope Pointing

A. Procedure at TMO

The TMO telescope control program includes a "retrieve" command to point the telescope at coordinates which appear in a file of positions. Both the uplink directions and the coordinates of the reference stars were produced at JPL in the format used at TMO, and the files were transmitted electronically to TMO for use at the telescope. However, the retrieve command did not work properly during checkout, so all pointing commands were issued manually by typing in the coordinates. Having constant star coordinates proved to be a wise decision.

During the experiment, the pointing engineer would first command the telescope to the chosen reference star, one that was northwest of the transmission direction, visible through the eyepiece, and as close to Galileo as possible. The reference stars actually used are listed in Table 1. Once the star was centered, the telescope right ascension and declination were reset to match the star's catalogued coordinates, thus calibrating the pointing, and the telescope was then offset to the predicted transmission direction. A second command to point anew at the transmission direction was issued just before each transmission, in order to eliminate any accumulated tracking errors. Selecting a star northwest of Galileo helped to combat the effect of atmospheric refraction: The star during calibration would appear at nearly the same altitude as Galileo would have during uplink, and so both would be affected nearly equally by refraction.

B. Procedure at SOR

SOR also pointed its telescope by offsetting from a nearby reference star. The SOR mount is much more stable than TMO's, and modeled better, so that it did not need to return to the reference star and recalibrate before every transmission. Rather, after the initial acquisition of the reference star, the telescope would move from one transmission direction to the next. Occasionally, the reference star would be observed as a check, but no recalibration was ever required.

IV. Star Catalog Development

The characteristics of the TMO telescope mount dictated that the reference stars be as close as possible in the sky to Galileo. This in turn meant that the reference catalog needed to be fairly dense, including many dim stars, in order to assure the availability of a suitable star every night. However, the stars also had to be bright enough to be seen visually through TMO's telescope at the coudé focus. The GOPEX project thus required a catalog of all stars brighter than magnitude 12.0 within a 5- by 5-deg region surrounding the uplink direction.

Existing star catalogs generally fall into two categories: very precise but sparse, and less precise but more complete. The *Fifth Fundamental Catalogue* (FK5) [3], with 1535 stars over the whole sky and precisions of several hundredths of an arcsecond, is an example of the former; the *Astrographic Catalogue Reference Stars* (ACRS) catalog [4], with 300,000 stars and precisions of half an arcsecond, is the most recent example of the latter. Furthermore, many star catalogs, particularly photographic ones, such as the ACRS, have a *maximum* brightness as well as a minimum brightness, as too-bright stars are overexposed and do not yield reliable image centroids.

Three catalogs were used to construct the GOPEX star catalog, in decreasing order of preference:

(1) The ACRS supplied 112 stars, most of which were between seventh and tenth magnitude.

(2) *The Smithsonian Astrophysical Observatory (SAO) Star Catalog* [5] supplied three stars: two that were too bright for the ACRS, and one that was the secondary component of a double star system.

(3) *The Hubble Guide Star Catalog (GSC)* [6–8] supplied 260 stars that were too faint for the ACRS.

There were no FK5 stars in the region of interest.

A comparison of ACRS and GSC coordinates for stars that appeared in both catalogs revealed no statistically significant differences between the coordinate systems of the two catalogs over the GOPEX region. This check was important because the GSC is known to have occasional large systematic biases. The final merged catalog, referred to J2000 coordinates, is believed to be accurate to half an arc second, well below the error budget.

The GOPEX star catalog was precessed to mean coordinates at epoch 1992 December 12 for use at TMO. In this way the stars' coordinates were in the same system as the uplink directions, so that one could offset the telescope from a star to Galileo without having to perform any calculations at the telescope.

The original J2000 version was transmitted to SOR, where the catalog was transformed into true-of-date positions in order to match their own uplink predictions which were likewise expressed in true-of-date coordinates.

V. Conclusion

All the components of telescope pointing for GOPEX, from the construction of a star catalog and prediction algorithms to the procedures at the telescopes at TMO and SOR, were spectacularly validated when the first long exposures taken by Galileo showed that laser light from both sources had reached the spacecraft. Even on the final days of the experiment, when the beam divergence at TMO was reduced to 60 μrad, the consistent uplink successes confirmed the accuracy of the telescope pointing strategy.

References

[1] K. E. Wilson, J. R. Lesh, and T.-Y. Yan, "An Overview of the Galileo Optical Experiment (GOPEX)," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 192–204, August 15, 1993

[2] J. H. Lieske, T. Lederle, W. Fricke, and B. Morando, "Expressions for the Precession Quantities Based upon the IAU (1976) System of Astronomical Constants," *Astron. Astrophys.*, vol. 58, pp. 1–16, 1977.

[3] W. Fricke, H. Schwan, and T. Lederle, *Fifth Fundamental Catalogue (FK5)*, Veröff. Astron. Rechen-Inst. Heidelberg, Karlsruhe: Verlag G. Braun, vol. 32, 1988.

[4] T. E. Corbin and S. E. Urban, *Astrographic Catalogue Reference Stars*, Washington, D. C.: United States Naval Observatory, 1991.

[5] Smithsonian Astrophysical Observatory, *Smithsonian Astrophysical Observatory Star Catalog*, Washington, D. C.: Smithsonian Institution, 1966.

[6] B. M. Lasker, C. R. Sturch, B. J. McLean, J. L. Russell, H. Jenkner, and M. M. Shara, "The Guide Star Catalog. I. Astronomical Foundations and Image Processing," *Astron. J.*, vol. 99, no. 6, pp. 2019–2058, June 1990.

[7] J. L. Russell, B. M. Lasker, B. J. McLean, C. R. Sturch, and H. Jenkner, "The Guide Star Catalog. II. Photometric and Astrometric Models and Solutions," *Astron. J.*, vol. 99, no. 6, pp. 2059–2081, June 1990.

[8] H. Jenkner, B. M. Lasker, C. R. Sturch, B. J. McLean, M. M. Shara, and J. L. Russell, "The Guide Star Catalog. III. Production, Database Organization, and Population Statistics," *Astron. J.*, vol. 99, no. 6, pp. 2082–2154, June 1990.

Table 1. GOPEX Reference Stars Used at TMF

Date Dec., 1992	Transmissions	Right Ascension h min sec	Declination deg arcmin arcsec	Magnitude
09	1 to 50	11 40 11.4	−5 21 02	10.35
09	51 to 52	11 40 29.9	−5 09 34	8.02
10	All	11 38 55.1	−5 11 03	9.45
10	1 to 6	11 39 18.2	−4 56 43	8.77
11	7 to 8	11 34 37.3	−4 19 20	6.39
11	9 to 20	11 39 18.2	−4 56 43	8.77
12	All	11 36 51.1	−4 33 12	8.39
14	All	11 36 51.1	−4 33 12	8.39
15	All	11 36 51.1	−4 33 12	8.39
16	All	11 36 51.1	−4 33 12	8.39

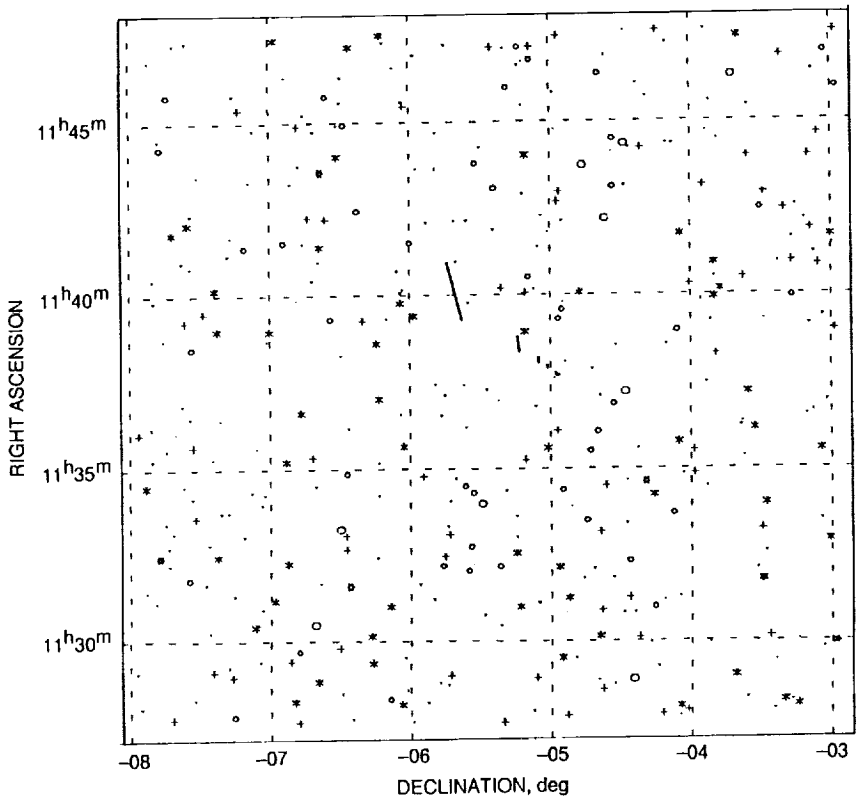


Fig. 1. The uplink direction from TMF to Galileo plotted against the GOPEX star catalog. Galileo's apparent motion was east to west (left to right) daily with a slow northern drift; the longest line corresponds to the first day of the GOPEX experiment. Stars are plotted by magnitude: five- and six-pointed stars for magnitudes 5 and 6; large and small open circles for magnitudes 7 and 8; asterisks for magnitude 9; plus signs for magnitude 10; dots for magnitude 11. Coordinates are referred to the mean equinox and equator of December 12, 1992.

TDA Progress Report 42-114

August 15, 1993

Galileo Optical Experiment (GOPEX) Optical Train: Design and Validation at the Table Mountain Facility

J. Yu and M. Shao
Optical Sciences and Applications Section

The Galileo Optical Experiment (GOPEX) has demonstrated the first laser communications uplink to a deep space vehicle. This article describes the optical design and validation tests performed at the Table Mountain Facility (TMF) transmitter site. The system used a 0.6-m telescope and an optical system at coudé focus to produce the uplink beam. The optical system used a pulsed neodymium:yttrium-aluminum-garnet (Nd:Yag) laser and beam diverger optics to produce the required optical output. In order to validate the optical design, a number of uplinks were performed on Earth-orbiting satellites (e.g., Lageos 1 and 2).

I. Introduction

The Galileo Optical Experiment (GOPEX) performed a deep-space optical uplink with the Galileo satellite during its second flyby of Earth (December 9–16, 1992) [1]. The uplink was performed using transmitter sites at the Table Mountain Facility (TMF), Wrightwood, California, and the Starfire Optical Range (SOR), Albuquerque, New Mexico.

The requirements for the laser transmitter at the TMF site were that a pulsed-doubled neodymium:yttrium-aluminum-garnet (Nd:Yag) laser (l = 532 nm) be used as the optical source. The pulses were to be 14 nsec wide, with a maximum energy of 250 mJ. The 0.6-m telescope was used to point the laser at Galileo and transmit the laser beam. The output beam was to have a divergence of 110 μrad during days 1–4 of the uplink schedule and a divergence of 60 μrad during days 5–8. In addition, the boresighting between the laser and the telescope had to be better than 10 μrad. This article describes the optical design of the GOPEX optical train, the alignment procedure, and the results of the tests performed to validate the system.

The optical design utilized a two-lens system to produce the correct beam divergence. An alignment and boresight procedure was developed and verified through range tests and uplinks to Earth-orbiting satellites. A number of point and shoot experiments were performed on Agessai and Lageos 1 to demonstrate boresighting of the laser with the telescope. In addition, satellite-tracking uplinks were performed on Lageos 1 and 2 and Etalon. All tests were successful and produced returns from the satellites.

During GOPEX, the optical train was aligned and verified every night prior to the uplink time. The beam divergence and laser output power were measured. GOPEX successfully produced uplink data on six of the seven days

of the experiment. Inclement weather prevented laser uplinks on day 4.

II. Optical Train Design

Figure 1 shows a schematic of the GOPEX optical train used at the TMF site. The design used the focal plane iris as the reference point for the entire optical system. During nominal operations, the telescope would be calibrated by focusing a star at the center of the iris. The laser beam would, in turn, be centered on the iris, which would guarantee boresighting of the laser beam with the pointing of the telescope. The output beam divergence would be varied by changing the size of the laser beam at the plane of the iris.

The 0.6-m telescope was used in an off-axis configuration. The laser beam uses only a 15-cm-diameter portion of the primary mirror and exits the telescope unvignetted by the secondary. To accomplish this, the laser beam was incident into the telescope at an angle with respect to its optical axis.

The optical train was subdivided into five components: the laser, the beam diverger, the laser diagnostic system, the charge-coupled device (CCD) acquisition system, and the satellite return system.

A. Laser System

The laser was a pulsed-doubled Nd:Yag laser emitting at a wavelength of 532 nm and capable of producing 14-nsec pulses at 15 and 30 Hz. The average laser output at 30 Hz was 6 W. A more complete description of the laser can be found in [2].

B. Beam Diverger System

Beam diverger optics were used to convert the raw laser beam output to a beam which exits the telescope at the appropriate divergence. GOPEX required output beam divergences of 110 and 60 μrad $\pm$10 percent. These requirements translated to requirements of beam diameter size at the focal plane of 3.36 mm and 1.84 mm, respectively.

The beam diverger also had a requirement that the beam divergence at the focal plane reference iris be approximately 4 μrad. The focal plane divergence was chosen as a compromise between beam vignetting and the power density on the telescope optics. The beam divergence was made as large as possible in order to maximize the beam size and hence minimize the power density at the coudé and telescope optics. However, divergences larger than about 4 μrad would result in the laser beam being vignetted by either the edge of the telescope or the secondary mirror.

Based on measurements of the output laser parameters made by Big Sky Corporation, the optical system was designed to meet the GOPEX requirements. The beam diverger design for both the 110- and the 60-μrad configurations consists of two lenses: a plano-convex and a plano-concave lens. Key elements of the output from the Code V optical design are given in Table 1.[1] The Code V output for the beam diverger optical design is given in Appendix A.

All lenses were purchased from CVI Laser Corporation. The part numbers of the lenses and specifications for the two configurations are given in Table 1. The plano-convex lens was antireflection coated at 532 nm. Because the plano-concave lenses continually suffered damage during initial testing, it was recommended that these lenses be used uncoated.[2] A series of experiments verified that the uncoated surfaces were capable of withstanding the laser energies at both 110- and 60-μrad configurations. In addition, it was determined that laser light reflected from the surface of the plano-concave lens would not focus at or near the plano-convex lens.

C. Laser Diagnostic System

A laser diagnostic system was used to determine the shape of the laser beams at various parts of the optical train. The Big Sky Software Corporation Beam Code Laser Diagnostic System was used. The optical train was set up so that the laser beam at the focal plane iris was monitored. A CVI Laser Corporation concave mirror with a focal length of 2 m was used in the diagnostic system. The divergence and spot size were measured by moving the focusing mirror into one of two locations (labeled A and B in Fig. 1). Location A imaged the focal plane onto the beam-monitoring CCD. The magnification factor was determined by measuring the distances between the principal planes and the focusing mirror with a tape measure. Location B was chosen one focal length away from the monitoring CCD and was used to measure the laser beam divergence.

[1] M. Chrisp, "GOPEX Beam Diverger Optical Design," JPL Interoffice Memorandum OSG 92-46 (internal document), Jet Propulsion Laboratory, Pasadena, California, March 26, 1992, and "GOPEX Redesigned Beam Divergers," JPL Interoffice Memorandum 92-138 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 29, 1992.

[2] G. Al Jumaily, "High Power Lenses for GOPEX," JPL Interoffice Memorandum OSG 92-142 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 26, 1992.

In addition to monitoring the laser beam, the diagnostic system was used to boresight the laser beam with the telescope. The alignment procedure is described in Section III.

D. The Vision Acquisition Systems

The visual acquisition system was used to center a calibration star at the center of the focal plane iris thereby calibrating the telescope to the laser-uplink system. This system consisted of a pick-off mirror to divert the starlight beam into the visual acquisition arm, lens, and eyepiece. The lens and eyepiece were aligned with the line of sight of the telescope and adjusted so that the observer would focus on the plane of the iris. This was important so that the position of the focal plane of the telescope would automatically be collocated with the plane of the iris. During laser uplink, the pick-off mirror would be flipped up to allow the laser beam to enter the telescope.

The CCD acquisition system was used to acquire star images and measure the seeing condition during GOPEX. The system used a Photometrics Star I Camera to perform the image acquisition. A flat mirror was manually placed to pick off the starlight which is normally used for the visual acquisition system and redirect it to the entrance of the camera. A 70- to 210-mm camera lens with a 4-cm extender tube was used to reimage the focal plane iris onto the CCD. The stellar image fell on an approximately 5×5 pixel array. Stellar targets of visual magnitude 4–6 were picked for the seeing measurement so that the integration time was in the 1- to 3-sec range. The long integration times served to average out the speckle caused by the atmosphere.

E. Satellite Return System

The satellite return system was only used during the precursor experiments. Laser uplinks to Earth-orbiting satellites were detected by a two-part transmit and receive (T/R) switch (Fig. 2).

The transmission side consisted of a rotating chopper blade and a light-emitting diode (LED). The chopper blade (labeled A) had a dielectric coating with a maximum reflection at 532 nm, which reflected the return signal from the satellite into the receiver side. In addition, the chopper blade had a 1-cm hole to allow transmission of the outgoing laser pulse through the chopper wheel. The wheel was rotated at 30 Hz. An LED (labeled F) which illuminated a smaller 5-mm hole in the chopper blade and a photodetector in the chopper motor (labeled A) was used to produce a synchronization signal. This signal was fed into the chopper motor driver (labeled D) and then into a delay generator (labeled E) which was then used to control the synchronization between the laser firing time and the position of the 1-cm hole in the chopper blade.

The returned laser beam was reflected by the chopper into the receiver side of the T/R switch where the signal was routed by fold mirrors (labeled B) to a photomultiplier tube (PMT; labeled C). The output of the PMT was then connected to a DSA 602A Digitizing Signal Analyzer for storage of the return pulse.

The signal analyzer was configured to detect only the laser return pulse. The trigger of the analyzer was enabled by the outgoing laser pulse. The trigger delay offset was set at the minimum transit time of the laser, which guaranteed that the analyzer would not interpret scattered light from the outgoing pulse as satellite returns. The trigger level was set so that the frequency of dark events was approximately one per second, which minimized the chance that a dark event would trigger the analyzer prior to the actual satellite return. The satellite returns were stored in the internal memory of the analyzer and later transferred to floppy disks and used to produce hard copies (see Figs. 3–6).

III. Optical Alignment Procedure

The following is the alignment procedure for the 60-μrad configuration. The alignment for the 110-μrad configuration is similar, with the addition of the fold mirrors (M2 and M3) :

(1) The laser output was optimized for pulse energy and beam spatial profile (Gaussian shape) using the Big Sky Corporation's diagnostic system.

(2) The focal plane iris was stopped down and the visual acquisition system was aligned. The lens (L3), eyepiece, and the flip mirror (M4) were moved so that the observer could focus on the focal plane iris.

(3) An aperture target board was placed at the output of the coudé tube. This was used to align the laser beam with the optical axis of the telescope

(4) A low-power raw laser beam was aligned to both the center of the focal plane iris and the coudé target. The tilt and position of mirror M1 were adjusted so that the light reflected from the telescope secondary was centered on the coudé target.

(5) A sheet of white paper was placed on the dome, and the edges of the telescope tube and secondary mirror were marked on the paper.

(6) The coudé target was removed and mirror M1 was adjusted to allow the laser beam to pass through the center of the focal plane iris and produce a spot on the dome paper, which is roughly centered between the telescope edge markings.

(7) The plano-convex lens (L1) was placed in its designed position and adjusted in centering and tilt by using the light reflected backwards from the surface of the lens.

(8) The plano-concave lens (L2) was placed in its designed position and also adjusted using back reflections.

(9) The laser-beam-monitoring system was then adjusted by focusing the focal plane iris onto the diagnostic CCD.

(10) The plano-concave lens was adjusted slightly so that the beam was centered on the iris. Then the beam diameter was measured and adjusted by either moving lens M5 or both lenses as a unit.

(11) The laser beam divergence at the focal plane was controlled by changing the spacing between the convex and concave lens. The beam diameter at the focal plane was adjusted by varying the distance from the concave lens and the focal plane. With these two degrees of freedom, it was possible to meet the GOPEX requirement.

(12) The laser beam output was then checked at the dome target to make sure there was no vignetting by the telescope edges.

IV. Optical System Boresight Tests

A. Range Test

Numerous range tests were used to verify the optical design and the alignment procedure. In these tests, a target was positioned on an adjoining slope (distance = 2.5 km). The target had a flashlight in its center. The telescope was focussed and centered on the telescope using the visual acquisition system. The laser was turned on and the laser beam pattern on the target was recorded on videotape. Both the laser beam diameter and centering on the target were studied.

B. Cross Hair Test

The cross hair test was used to verify the boresight of the optical system without the use of an external target. Instead, the cross hairs of the finder scope were used to determine the system's pointing accuracy.

The cross hair test was conducted as follows:

(1) A relatively bright star was centered on the center of the focal plane iris.

(2) The cross hair of the 0.6-m telescope's finder scope was then centered on the same star.

(3) Laser pulses were fired and the pointing of the laser with respect to the cross hair was checked to determine boresight.

(4) Adjustments were made to the optical system (with Beamcode) if the boresight was found to be in error.

V. Precursor Experiments

Prior to the GOPEX encounter, numerous precursor experiments were performed to test laser uplinks to Earth-orbiting satellites. This section summarizes the results of these tests.

A. Agessai Point and Shoot Experiments

The first laser return experiments were performed on the Agessai satellite. This was a low altitude satellite which travelled horizon to horizon in about 15 minutes. Because of this rapid velocity, satellite tracking using the TMO 0.6-m telescope could not be done. Consequently, an uplink experiment was performed that consisted of pointing the laser at a point on the satellite track and waiting for a return as the satellite crossed the beam. In order to guarantee at least one returned pulse, a 30-Hz pulse repetition rate and a 300-μrad beam diameter were chosen to be used. This output beam divergence was implemented by moving the 110-μrad lens pair away from the focal plane iris (see Fig. 2).

The experiment consisted of finding the closest *Fifth Fundamental Catalogue* star to a particular point in the satellite track (typically the middle of the track) and calibrating the telescope on that star. At a time close to when the satellite would cross that point in the track, the telescope was off-pointed to the predicted satellite location and the laser was fired. At the time the satellite crossed the laser beam, a return pulse would be observed.

Two experiments were performed on October 1 and 6. Both resulted in successful returns from Agessai. Figure 3(a) shows the satellite return on October 1. The return shows a typical output from a Q-switched laser, which is distinctly different from a photomultiplier dark event, as shown in Fig. 3(b). Figure 4 shows two returns obtained on October 6. The satellite return on this date was markedly larger than in the previous experiment. This

was attributed to the shorter time (~5 sec) between the pointing of the telescope and the time of laser firing. In the first experiment, the delay between pointing and firing was approximately 2 minutes, which allowed the pointing of the telescope to drift significantly.

B. Lageos 1 Point-and-Shoot Experiment

An experiment similar to the Agessai experiments was performed on the Lageos 1 satellite on November 1. Lageos is a higher flying satellite and traverses horizon to horizon in about 30 minutes. For this experiment, the output beam divergence was decreased to 110 μrad, with a repetition rate of 30 Hz. The laser was shot at the satellite at three different positions in its track. For all three shots, a laser return was observed. Figure 5 shows the laser returns for the different satellite positions.

C. Lageos 1 and 2 Tracking Experiment

A set of satellite tracking experiments was performed on November 2. In this set of experiments, laser pulses were uplinked as the telescope was set to track the satellites along its path. A 110-μrad beam divergence and a 15-Hz repetition rate were used. A description of the setup for the tracking experiment follows:

(1) Satellite predict files were obtained for both point-ahead and non-point-ahead files. The point-ahead file was to be used for laser uplink and the non-point-ahead file was to be used for visual acquisition.

(2) The satellite track was divided into four sections. The first and third sections used the non-point-ahead file, and the second and fourth sections used the point-ahead file. This merged satellite file was used to control the telescope pointing.

(3) The telescope mount was calibrated along the satellite track using the LEARN command in the telescope control software. During this procedure, numerous stars close to the track path were found and the offset from the mount pointing was recorded in the program.

(4) The satellite was visually acquired during the first section of the track and centered on the focal plane iris. Because the satellites only showed up as sporadic glints of light, two people (J. Young and B. Hornaday) were needed for visual acquisition: one to visually search for the satellite in the finder scope and the second to acquire it in the coudé room.

(5) Before the second section of the track began, the laser beam would begin firing and the uplink returns would be observed and recorded on the storage scope.

(6) The visual acquisition procedure was repeated for the third section of the track and the laser was uplinked during the fourth section of the track.

On November 2, tracking laser uplinks were performed on both Lageos 1 and 2. Since both satellites were passing over at the same time, it was decided that tracking would be performed on only half of the Lageos 1 track and about three-quarters of the Lageos 2 track. In all cases, successful returns from both satellites were observed. Figure 6 shows a small sample of the many pulse returns from the two satellites.

D. Etalon Tracking Experiment

On November 30, a 60-μrad divergence beam with a 15-Hz repetition rate was used to track the satellite Etalon. This satellite took approximately 2.5 hours to traverse from horizon to horizon. Again the track was divided into four parts with two acquisition and uplink pairs. This experiment was performed after the internal dry run that had been performed earlier that night. Successful returns were obtained from Etalon. It was noted that the signal was sporadic; for some of the time the returns were strong and continuous, followed by a period where there were no returns, followed again by a strong series of returns. This was attributed to wobble in the pointing of the telescope, which was corroborated during the visual acquisition phase where the satellite was observed to wobble in the field as the telescope tracked.

VI. Galileo Encounter Support

During Galileo encounter, the optical train was checked every night and the beam parameters were measured. The checks included measuring the output laser power and beam divergence. In addition, boresighting of the laser beam with the telescope was checked using both a range test and a cross hair test.

As described previously, the range test involved firing the laser beam at a target located on a slope facing TMF. This test was performed on day 1 and day 5 of the experiment. The beam pointing and shape were checked to see if they were qualitatively within the specifications for that night's experiment. Range tests for the other nights were not done since the optical system had not been reconfigured.

On most nights, no problems were encountered and no adjustments to the optical system were made during the time of the experiment. However, during day 3, when

the repetition rate for the 110-μrad configuration was increased from 15 to 30 Hz, laser damage was observed on the coudé turning mirror. In this situation, the mirror was rotated to an undamaged section of the mirror just prior to the time for Galileo uplink. By the end of the experiment, evidence of laser damage was found on the turning mirror. Day 4 experiments were cancelled at the TMO site due to unfavorable weather conditions.

VII. Conclusion

The optical design and alignment procedures for the GOPEX optical train at TMF have been presented. The results for the first optical uplink to a deep space satellite were very successful. This was due, in part, to an extensive verification process that used uplinks to Earth-orbiting satellites and range tests to validate the performance of the optical train.

Acknowledgments

The authors would like to recognize Michael Chrisp and Ghanim Al-Jumaily for their contributions to the optical design. Thanks also to Kevin Masters, William Hornaday, Jim Young, Garrett Okamoto, and Ron Dotson for their contributions during GOPEX and the precursor experiments. The contributions of Jim Lesh, Keith Wilson, and Tsun-Yee Yan, as well as the Table Mountain Facility staff, including Dan Sidwell and Pam Glatfelter, are also gratefully acknowledged.

References

[1] K. E. Wilson and J. R. Lesh, "An Overview of the Galileo Optical Experiment (GOPEX)," *The Telecommunication and Data Acquisition Progress Report 42-114, vol. April–July 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 192–204, August 15, 1993.

[2] G. Okamoto and K. Masters, "GOPEX Laser Transmission and Monitoring Systems," *The Telecommunication and Data Acquisition Progress Report 42-114, vol. April–July 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 248–254, August 15, 1993.

Table 1. Optical design parameters.

Beam divergence, μrad	Convex	Concave	Lens spacing, mm	Distance to focal plane, mm
110	PLCX–25.4/128.8-UV	PLCC–25.4/77.3-UV	84.70	75.00
80	PLCX–25.4/128.8-UV	PLCC–25.4/38.6-UV	186.23	71.54

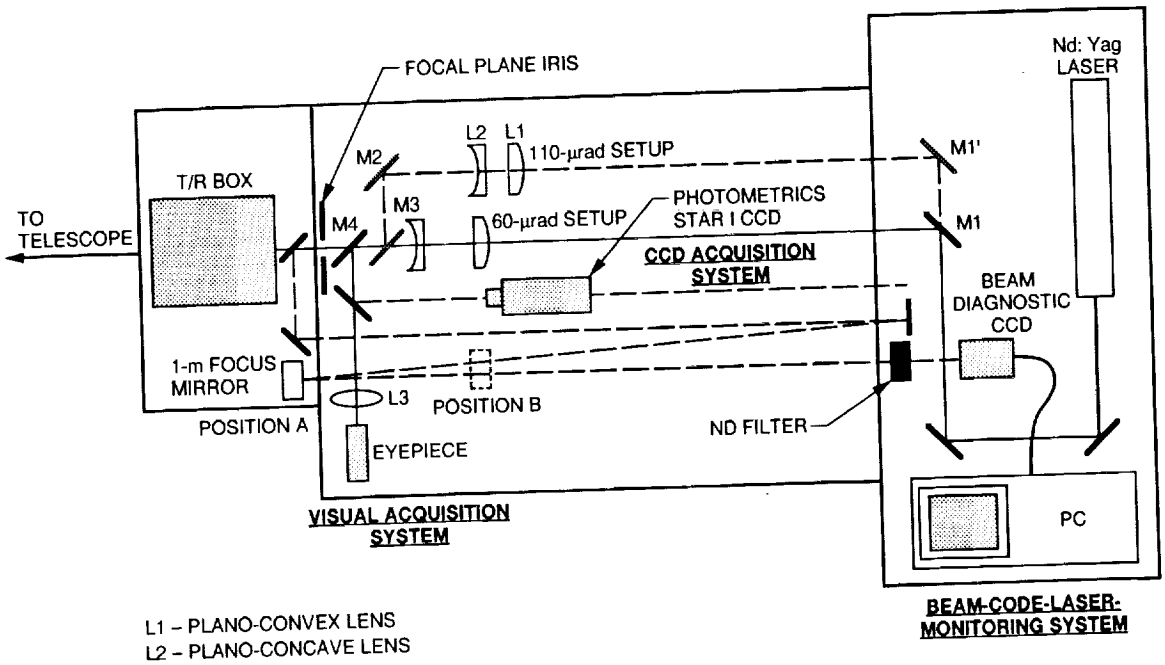


Fig. 1. GOPEX Optical System layout.

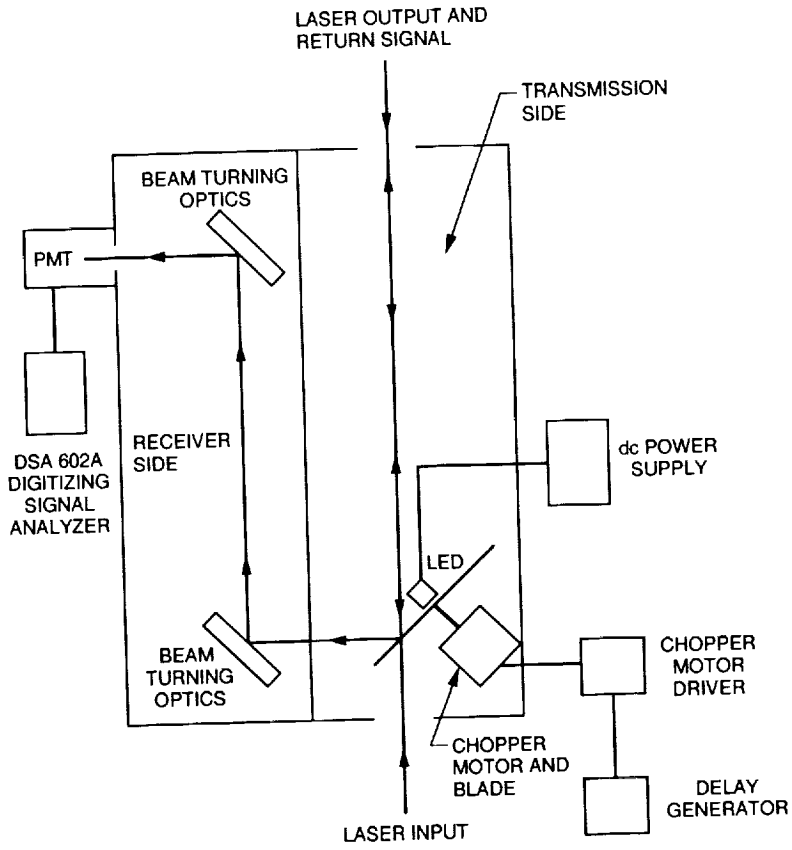


Fig. 2. Transmit/Receive switch.

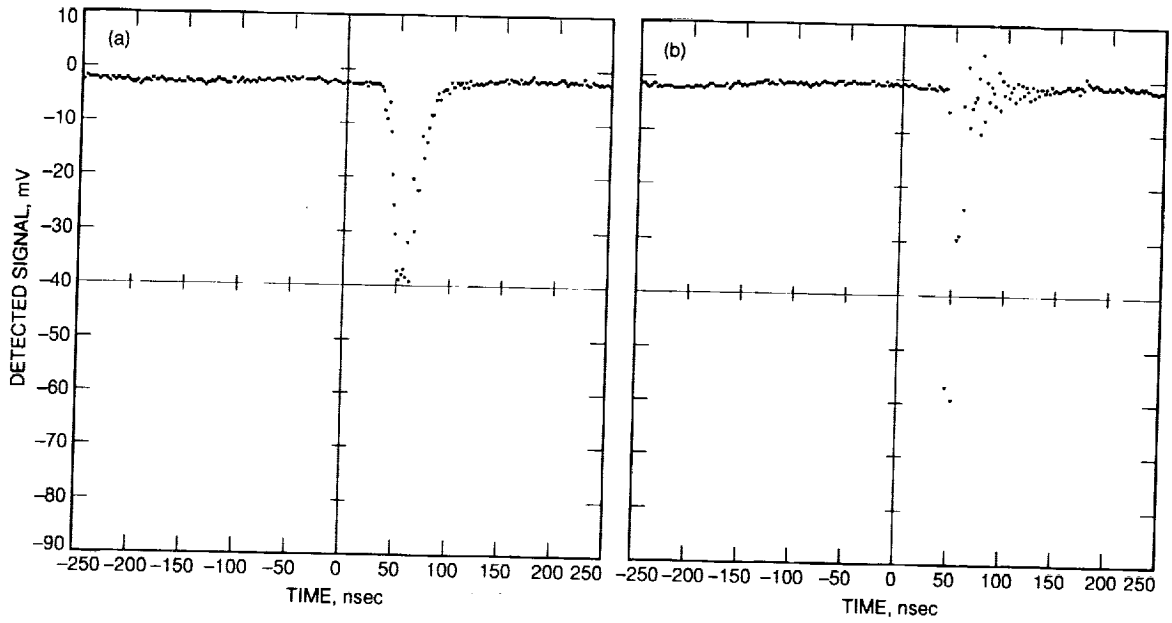


Fig. 3. Agessai Point and Shoot Experiment I (October 1, 1992): (a) Satellite return and (b) Photon detector dark event.

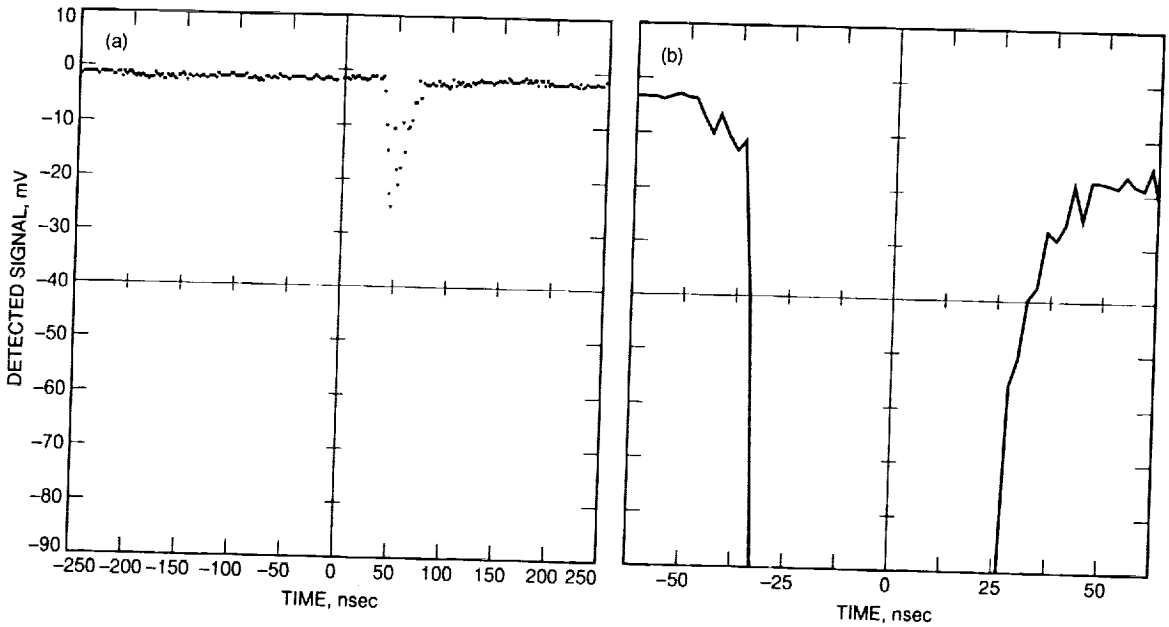


Fig. 4. Agessai Point and Shoot Experiment II (October 7, 1992): (a) Agessai return, highest point and (b) Agessai return, descending point.

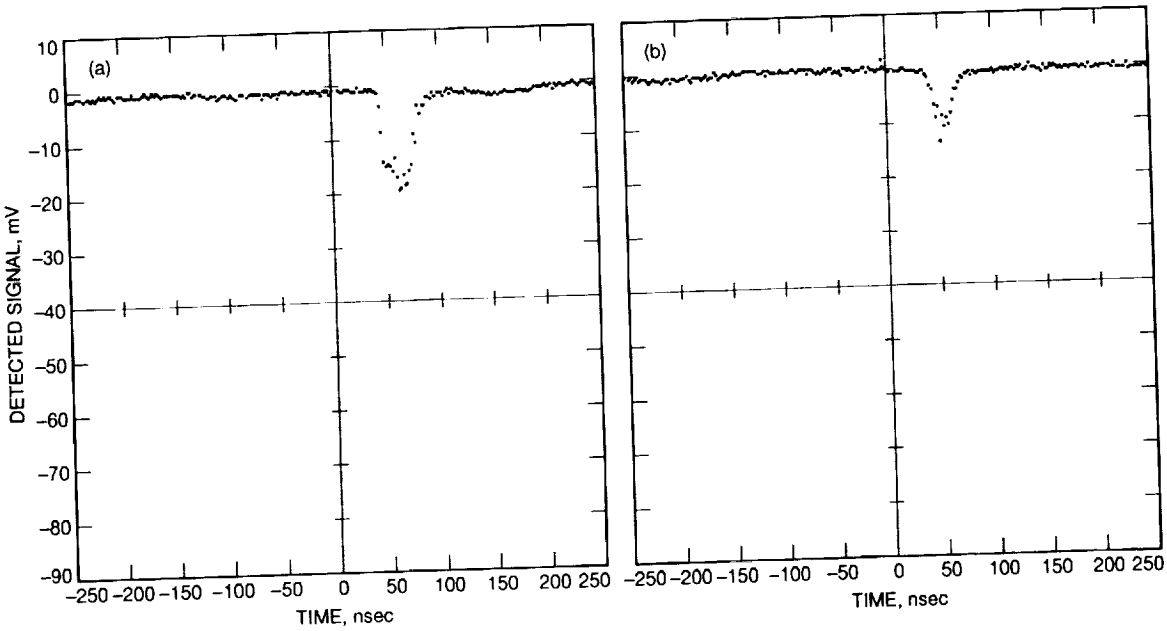


Fig. 5. Lageos Point and Shoot Experiment (November 1, 1992): (a) Lageos return, highest point and (b) Lageos return, descending point.

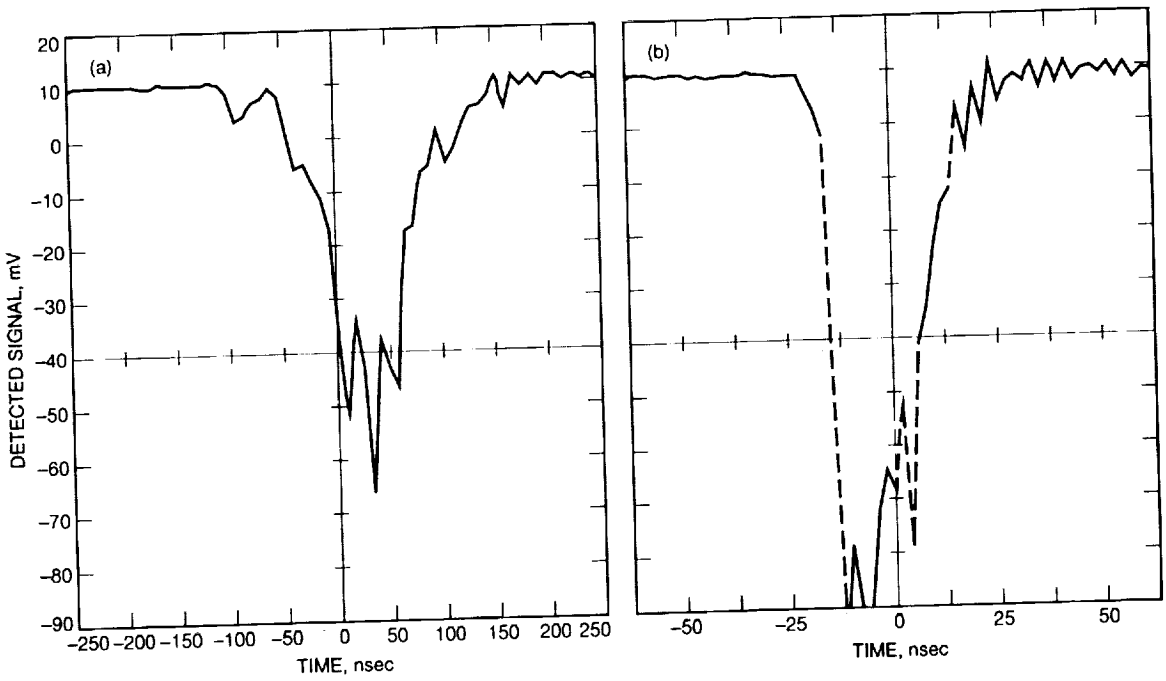


Fig. 6. Satellite returns from Lageos Tracking Experiment (November 2, 1992): (a) Lageos 1 uplink and (b) Lageos 2 uplink.

Appendix A

Code V Design of Beam Divergers

```
     File [MPC.GOPEXJGOOD110.LEN(2) has been restored
     Lens title: "REDESIGNED 110 MICRORAD SYSTEM 9/28/92"
 CODE V> LIS
          LIS
      REDESIGNED 110 MICRORAD SYSTEM 9/28/92
                 RDY              THI        RMD        GLA
 > OBJ:          INFINITY      4766.815342
   STO:          INFINITY         0.000000
     2:         128.80000         4.000000        SILICA_SPECIAL
     3:          INFINITY        84.704658
     4:          INFINITY         2.300000        SILICA_SPECIAL
     5:         77.30000          0.000000
     6:          INFINITY        75.000000
   IMG:          INFINITY         0.000000

 SPECIFICATION DATA
    NAO          0.00046
    DIM               MM
    WL            532.00
    REF                1
    WTW                1
    XAN          0.00000
    YAN          0.00000
    VUY          0.00000
    VLY          0.00000

 APERTURE DATA/EDGE DEFINITIONS
    CA
    CIR S1                        6.550000

 REFRACTIVE INDICES
     GLASS CODE                        532.00
    SILICA_SPECIAL                    1.460711

 No solves defined in system

 INFINITE CONJUGATES
    EFL       -2060.3872
    BFL       -1404.3365
    FFL        3153.5222
    FNO        -466.4748
 AT USED CONJUGATES
    RED          -0.2601
    FNO        -280.7455
    OBJ DIS    4766.8153
    TT         4932.8200
    IMG DIS      75.0000
    OAL          91.0047
    PARAXIAL IMAGE
     HT           0.0000
     THI       -868.3498
    ANG           0.0000
    ENTRANCE PUPIL
     DIA          4.4169
     THI          0.0000
    EXIT PUPIL
     DIA          2.8858
     THI        -58.1606
 CODE V> RSI 1 0 0 0
          RSI 1 0 0 0
     REDESIGNED 110 MICRORAD SYSTEM 9/28/92
     Position  1, Wavelength =    532.0 NM
              X             Y             Z          TANX          TANY          LENGTH
 OBJ       0.00000       0.00000       0.00000       0.00046       0.00000
 STO       2.20847       0.00000       0.00000       0.00046       0.00000       0.00000
   2       2.20847       0.00000       0.01894      -0.00509       0.00000       0.01894
   3       2.18820       0.00000       0.00000      -0.00744       0.00000       3.98112
   4       1.55823       0.00000       0.00000      -0.00509       0.00000      84.70700
   5       1.54644       0.00000       0.01547       0.00178       0.00000       2.31550
   6       1.54641       0.00000       0.00000       0.00178       0.00000      -0.01547
 IMG       1.68000       0.00000       0.00000       0.00178       0.00000      75.00012
                              OPD =     2.813 Waves
```

Fig. A-1. A 110-μrad beam diverger Code V listing.

Appendix B

GOPEX Code V Optical Design

```
        RES [MPC.GOPEX]GOOD60
    File [MPC.GOPEX]GOOD60.LEN(3) has been restored
    Lens title: "REDESIGNED 60 MICRORAD SYSTEM 9/28/92"
CODE V> LIS
        LIS
     REDESIGNED 60 MICRORAD SYSTEM 9/28/92
                 RDY            THI      RMD        GLA
> OBJ:          INFINITY     4668.743625
  STO:          INFINITY        0.000000
    2:         128.80000        4.000000        SILICA_SPECIAL
    3:          INFINITY      186.231488
    4:          INFINITY        2.300000        SILICA_SPECIAL
    5:         38.60000         0.000000
    6:          INFINITY       71.544887
  IMG:          INFINITY        0.000000

SPECIFICATION DATA
   NAO          0.00046
   DIM               MM
   WL            532.00
   REF                1
   WTW                1
   XAN          0.00000
   YAN          0.00000
   VUY          0.00000
   VLY          0.00000

APERTURE DATA/EDGE DEFINITIONS
   CA
   CIR S1                          6.550000

REFRACTIVE INDICES
                                      532.00
    GLASS CODE
   SILICA_SPECIAL                    1.460711

No solves defined in system

INFINITE CONJUGATES
   EFL       -4470.1230
   BFL       -1423.4322
   FFL       14636.2829
   FNO       -1033.3016
AT USED CONJUGATES
   RED           -0.2316
   FNO         -249.8945
   OBJ DIS    4668.7436
   TT         4932.8200
   IMG DIS      71.5449
   OAL         192.5315
   PARAXIAL IMAGE
    HT           0.0000
    THI       -388.3650
   ANG           0.0000
   ENTRANCE PUPIL
    DIA          4.3261
    THI          0.0000
   EXIT PUPIL
    DIA          1.3212
    THI        -58.1949

CODE V> RSI 1 0 0 0
        RSI 1 0 0 0
     REDESIGNED 60 MICRORAD SYSTEM 9/28/92
     Position  1, Wavelength =    532.0 NM
              X            Y            Z          TANX         TANY        LENGTH
OBJ       0.00000      0.00000      0.00000      0.00046      0.00000
STO       2.16303      0.00000      0.00000      0.00046      0.00000       0.00000
  2       2.16304      0.00000      0.01816     -0.00498      0.00000       0.01816
  3       2.14321      0.00000      0.00000     -0.00727      0.00000       3.98189
  4       0.78842      0.00000      0.00000     -0.00498      0.00000     186.23642
  5       0.77693      0.00000      0.00782      0.00200      0.00000       2.30785
  6       0.77691      0.00000      0.00000      0.00200      0.00000      -0.00782
IMG       0.92000      0.00000      0.00000      0.00200      0.00000      71.54503
                                OPD =     1.729 Waves
```

Fig. B-1. A 60-μrad beam diverger Code V listing.

S22-36
185882

TDA Progress Report 42-114

August 15, 1993

P-9

N94-14391

GOPEX Laser Transmission and Monitoring Systems

G. Okamoto and K. Masters
Communications Systems Research Section

This article describes the laser transmission and monitoring system for the Galileo Optical Experiment (GOPEX) at the Table Mountain Facility (TMF) in Wrightwood, California. The transmission system configuration and the data measurement techniques are described. The calibration procedure and the data analysis algorithm are also discussed. The mean and standard deviation of the laser energy transmitted each day of GOPEX show that the laser transmission system performed well and within the limit established in conjunction with the Galileo Project for experiment concurrence.

I. Introduction

The Galileo Optical Experiment (GOPEX) was conducted from December 9–16, 1992, when laser transmissions from the Table Mountain Facility (TMF) in Wrightwood, California, and the Starfire Optical Range (SOR) in Albuquerque, New Mexico, were successfully detected by the Galileo Solid State Imaging (SSI) camera [1]. The laser system at TMF was operated and maintained consistent with the Galileo Project's concurrence requirement of mean energy of 250 mJ, with a 10-percent margin. The laser was electronically controlled and monitored, and pulse output characteristics were recorded at both TMF and SOR. This article describes the laser-transmission-control and laser-emission monitoring systems at TMF. Section II discusses the laser equipment at TMF; Section III discusses the data measurement techniques; Section IV contains the results and analysis; and Section V contains the conclusion. Analysis of the data showed that the laser performed as expected and was maintained within the requirements.

II. Laser Equipment

Figure 1 shows the operational setup of the laser and optics system in the coudé path of the TMF 0.6-m telescope. The laser (labeled A in the figure) was a dual ellipse flashlamp-pumped neodymium: yttrium-aluminum-garnet (Nd:YAG) oscillator/amplifier with a second harmonic generator [2]. It can produce laser pulses at 12 nsec full-width half-maximum (FWHM), with 0.7-mrad beam divergence and repetition rates of 5, 6, 10, 15, and 30 Hz. The beam-steering mirrors (labeled B in the figure) were high-damage-threshold dielectric coated 2-in. fused silica. The coating was dielectric stack with a high reflection at 532 nm and a high transmission at 1064 nm for 45-deg incident radiation. The convex and concave lenses (labeled D and E in the figure) have an antireflecting coating for 532 nm. These lenses were used to control the beam divergence leaving the telescope (110 μrad for the case shown). The mirror (labeled B in the figure) and both lenses (labeled D and E) were moved to position 3, and both mirrors (labeled F) were removed for the 60-μrad

case, as noted by the dotted components and transmission path. The other beam-steering mirrors (labeled F) were high-damage-threshold dielectric coated 1-in. fused silica, coated for maximum reflection for 532 nm at near-normal incidence. A flip mirror (labeled H in the figure) was used to direct the light from the telescope into an eyepiece (labeled J in the figure) for viewing and aligning stars in the field-stop aperture (labeled G in the figure). The alignment was done only when the laser was off. The flip mirror (labeled K in the figure) was used to direct the laser beam into the camera (labeled P in the figure), which was connected to the Beam-code computer (labeled Q) in the figure. Beam-code is a beam-profile-analyst software and hardware setup from the Big Sky Software Company that was used to adjust and fine-tune the laser. It was also used to align the outgoing laser beam to the field-stop aperture. Detailed information about the optical setup is presented in [3] in this volume.

The GOPEX Monitoring Program recorded all the data presented in this article. The monitoring software was developed by the JPL Instrumentation Group to enable the user to control the Tektronix DSA 602A Digitizing Signal Analyzer. The software used LabWindows by National Instruments as the programming environment. This created an interactive graphical interface for the program. Configuration files were created so that the proper configuration could be loaded before each frame of data was to be taken. A total of 133 frames of usable data were recorded, with each frame containing 46 or 92 laser pulses, depending on the transmission duration. Each laser pulse waveform was stored in a separate file and each frame was stored on a different disk. Data from these disks were later archived to an optical disk for permanent storage.

III. Data Measurement

Because the GOPEX team needed to record both pulse width and energy of each laser pulse transmitted to Galileo, a fast detector with a rapid data acquisition and storage system was needed. The Hamamatsu R1193u photo detector combined with the Tektronix DSA 602A digital storage oscilloscope was used to rapidly record each outgoing laser pulse form. From these data both the pulse width (FWHM) and the energy per pulse calculated from the area under the pulse form were obtained.

The photodetector was positioned at position 1, as seen in Fig. 1. At this point the detector monitored the leakage through the turning mirror B_v. A 532-nm bandpass interference filter rejected the 1.06-laser light.

By measuring the reflected energy at position 2 (not shown in Fig. 1 because it was removed after calibration) using an Ophir power meter and comparing that with the integrated area measured by the DSA 602A, the energy per pulse transmitted to the spacecraft was deduced. Ten measurements were made at each of seven different laser power levels with the Hamamatsu at position 1 and the Ophir at position 2. Each set of Ophir power meter readings was averaged and recorded along with the corresponding integrated pulse area displayed by the DSA 602A. The results (shown in Fig. 2) are plotted for various laser power levels. To ensure that the laser beam characteristics (i.e., pulse shape and beam profile) were unchanged at the various power levels, the laser oscillator settings were held constant and the laser output energy was changed by detuning the delay time of the firing of the amplifier flash lamp away from its optimum setting.

All transmission data were recorded on floppy diskettes using the floppy drive in the DSA 602A. Each data file included a variable-length header of approximately 300 bytes in a text format followed by two-byte data points in a binary format. The header contains such information as the time and date the data were recorded, the waveform name and number, and the scope settings. The multiplier factors to convert the horizontal and vertical scales into physical units were computed from the information contained in the header, the information recorded on each night of transmission, and the specifications of the DSA 602A [4].

By using an IBM-compatible computer, data were extracted, scaled, and integrated to find the area of each laser pulse. The results were compared with the calibration data to translate the area computed into laser energy values. Measurements taken daily of the GOPEX transmission system using the power meter and the DSA 602A were used to validate these results.

The conversion factors needed to compute the energy of the laser pulses from the raw data values of the stored laser pulses were computed as follows: The interval of time between data points in seconds was computed, as were the proper scaling factors required to convert the raw data values into voltage readings. The time and voltage values were then used to compute the area of the laser pulses in units of picovolts squared times seconds (pVVsec) to be consistent with the units in which the calibration data were taken by the DSA 602A. The computed area value was translated into an energy value by using the conversion scale shown in Fig. 2 and linearly interpolating between the nearest two data points to find the appropriate scaling factor. The area was calculated as follows:

The interval of time between data points in seconds on the DSA 602A, x_{mult}, was calculated by dividing the time in the interval by the number of sample points in the interval using

$$x_{mult} = D_x(X/X_r) \tag{1}$$

where D_x is the number of horizontal divisions, X is the horizontal sensitivity of the oscilloscope set by the monitoring software in seconds, and X_r is the number of valid sample points recorded in the timing interval.

Data were digitized over a 64-K value range (R) and scaled into voltage values by y_{mult}, calculated by

$$y_{mult} = D_y(Y/R) \tag{2}$$

where D_y is the number of vertical divisions and Y is the vertical sensitivity of the oscilloscope set by the monitoring software in volts per division. The data values were then converted into volts (V) by

$$V = \text{data}(y_{mult}) \tag{3}$$

The area of the laser pulse was computed by multiplying the voltage squared of each sample point by the time interval of the sample point and then summing over all sample points. Using the calibration measurements, the area of the laser pulse was converted into energy, as shown in Fig. 2. The exact energy reading was then interpolated linearly between the nearest two data points in Fig. 2 to get the most appropriate scaling factor.

From these data, the mean laser output energy for each frame and day given by [5] is

$$\mu = \frac{1}{N}\sum_{i=1}^{N} E_i \tag{4}$$

where μ is the mean energy, E_i is the energy of the particular laser pulse, and N is the number of laser pulses in the desired frame or day. The variance (σ^2) of the set of laser pulses given by [5] is

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N} (E_i - \mu)^2 \tag{5}$$

IV. Results and Analysis

Table 1 shows the mean and standard deviation for the output of the laser for each of the GOPEX days. No data were available for days 4 and 5 since TMF was shut down on day 4 due to weather and there were no planned GOPEX activities on day 5. Figure 3 shows a plot of the mean laser output for each day, and Figs. 4–9 plot the mean for each frame in each day, with the overall mean for the day noted by a dotted line. Certain frames on the first three days are not shown because no transmission took place due to U.S. Space Command prohibitions, weather problems, or equipment failures. The intensity fluctuations from pulse to pulse, as seen in Figs. 4–9, were expected since the transmission system used a flashlamp-pumped laser.

The mean laser energy for the first two days, as shown in Table 1, was significantly less than the other days because of the aging flashlamps in the laser system. The laser energy for the second day was approximately 20 percent lower than the first day because of a timing problem that was discovered and corrected after the second day of transmission. This reduction in energy resulted in a noticeable drop in the intensity level of the signals received at the spacecraft, much lower than the signal levels received from SOR for that day. The flashlamps were replaced before the third day of GOPEX. This boosted the laser power to the expected level (250 mJ). The laser power was increased slightly on the seventh day to compensate for the long transmission distance. The laser power was increased again on the eighth day for the same reason. On all days, the mean laser power remained within the 10-percent margin allowed by the experiment concurrence negotiated with the Galileo Project.

V. Conclusion

The GOPEX laser transmissions at TMF were successfully detected by the Galileo SSI camera. The output of the laser was monitored by a photodetector which sent the detected signal to a Tektronix DSA 602A Digitizing Signal Analyzer which, in turn, was controlled by the GOPEX monitoring system. The monitoring system then transferred the data to storage diskettes. Software was created to extract all the data and compute the mean and standard deviation for each frame of each day, as well as for each day overall. These results show that the GOPEX laser transmission system performed as expected and was maintained within specifications.

Acknowledgments

The authors gratefully acknowledge the support provided by J. McGregor, F. Loya, and E. Baroth of the Instrumentation Section's Measurement Technology Center at JPL. The authors would like to thank the staff at TMF for their excellent support. They also thank T.-Y. Yan of the Communications Systems Research Section at JPL for his helpful suggestions about this article.

References

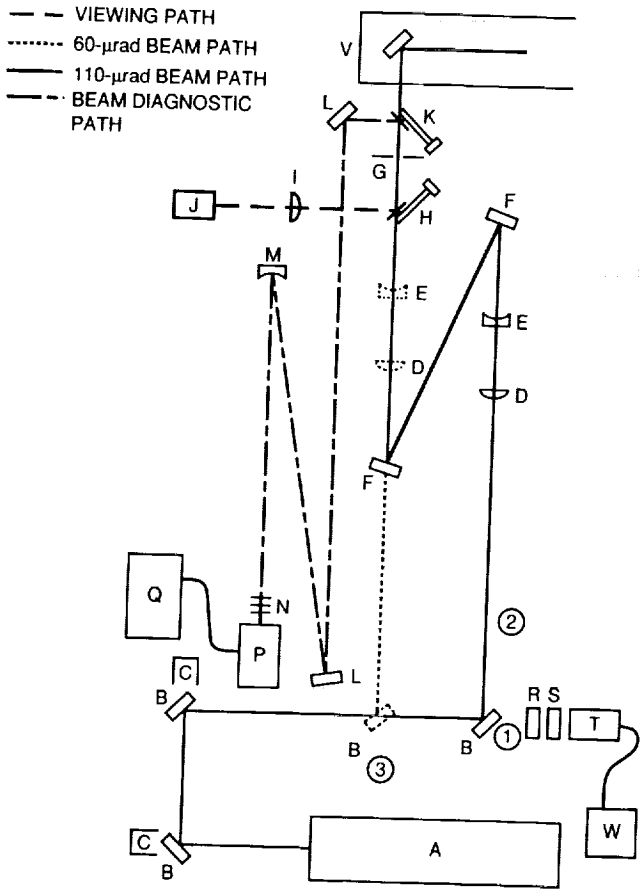
[1] K. E. Wilson and J. R. Lesh, "An Overview of the Galileo Optical Experiment (GOPEX)," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 192–204, August 15, 1993.

[2] *YG580 Operation and Maintenance Manual*, Santa Clara, California: Quantel International, 1986.

[3] J. Yu and M. Shao, "The Galileo Optical Experiment (GOPEX) Optical Train: Design and Validation at the Table Mountain Facility," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 236–247, August 15, 1993.

[4] *The DSA 601A and DSA 602A Digitizing Signal Analyzers User Reference*, Beaverton, Oregon: Tektronix, February 1991.

[5] P. R. Bevington, *Data Reduction and Error Analysis for the Physical Sciences*, New York: McGraw-Hill, 1969.



A. LASER CONTINUUM, yg581-30 w/SHG
B. MIRROR, R_{max} 532 nm
T_{max}, 1060 nm
C. LASER BEAM DUMP
D. CONVEX LENS
E. CONCAVE LENS
F. MIRROR, R_{max} 532 nm
G. FIELD STOP APERTURE
H. FLIP-IN FRONT SURFACE MIRROR
I. IMAGING LENS
J. EYEPIECE
K. FLIP-IN BARE GLASS
L. MIRROR, R_{max} 532 nm
M. MIRROR, 2 mcc R_{max} 532 nm
N. ND FILTERS
P. CAMERA
Q. BEAM-CODE COMPUTER
R. GROUND-GLASS DIFFUSER
S. 532-nm BANDPASS FILTER
T. PHOTODETECTOR
W. DSA 602A TEKTRONIX ANALYZER
V. 0.6-m TELESCOPE

1. ENERGY AND TRANSMITTED DIAGNOSTICS
2. TEMPORARY POSITION OF POWER METER FOR CALIBRATION
3. MIRROR POSITION FOR 60-μrad BEAM PATH

Fig. 1. Optics setup.

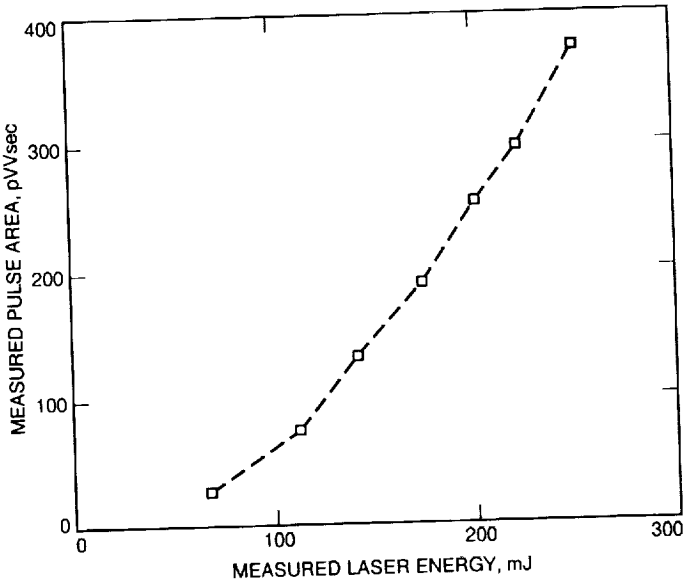


Fig. 2. Conversion curve for calibration data.

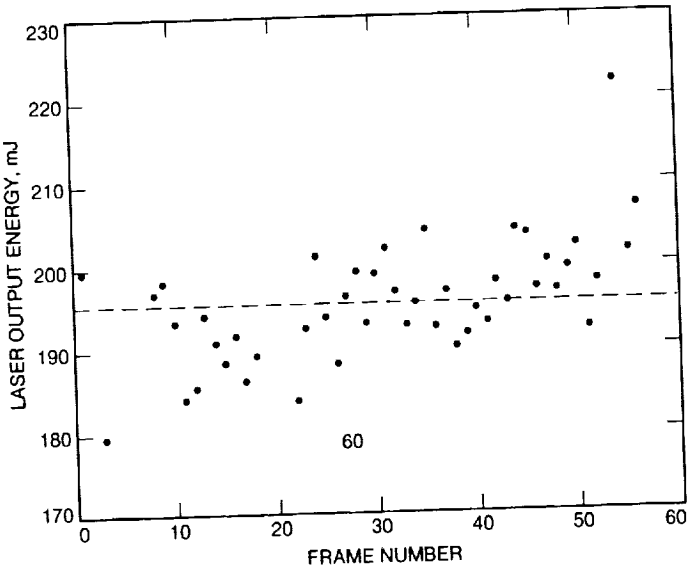


Fig. 4. GOPEX laser output for day 1.

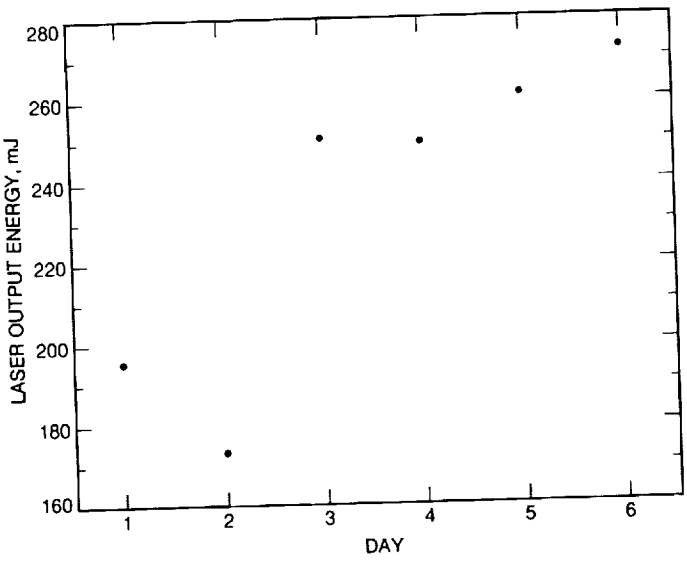


Fig. 3. GOPEX laser energy output per day.

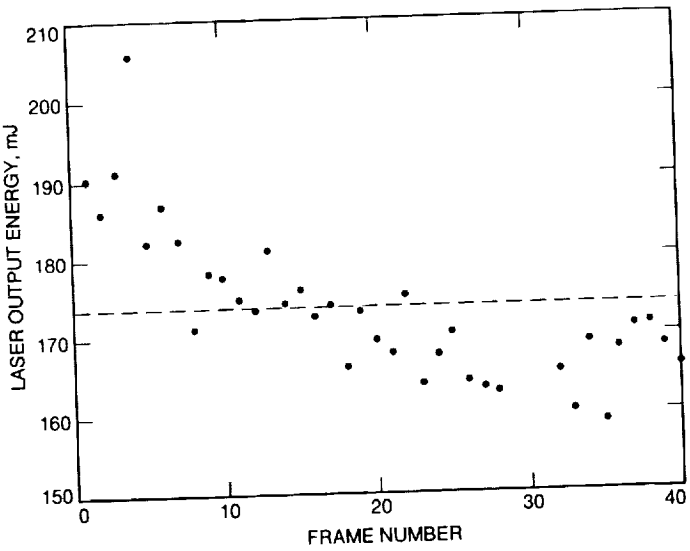


Fig. 5. GOPEX laser output for day 2.

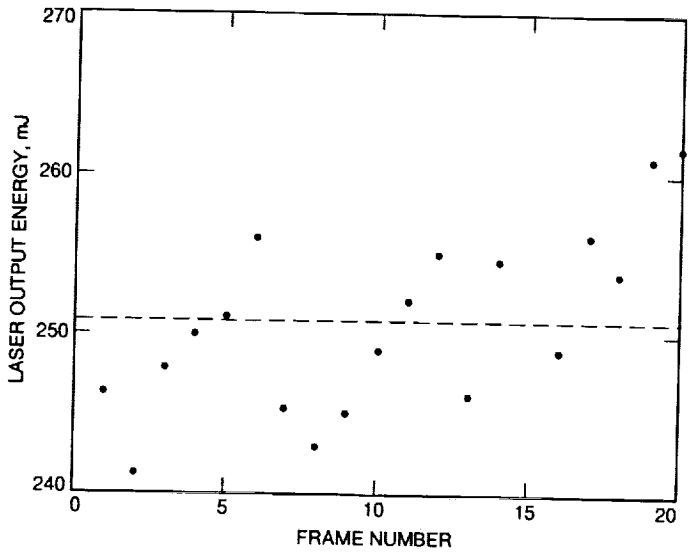


Fig. 6. GOPEX laser output for day 3.

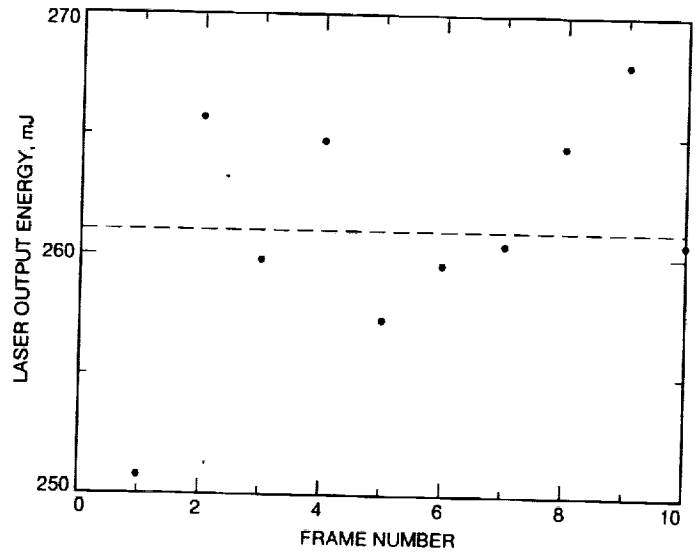


Fig. 8. GOPEX laser output for day 7.

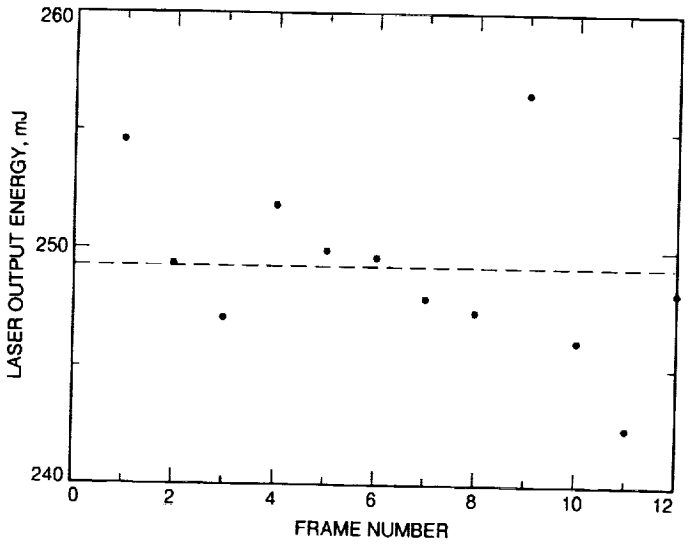


Fig. 7. GOPEX laser output for day 6.

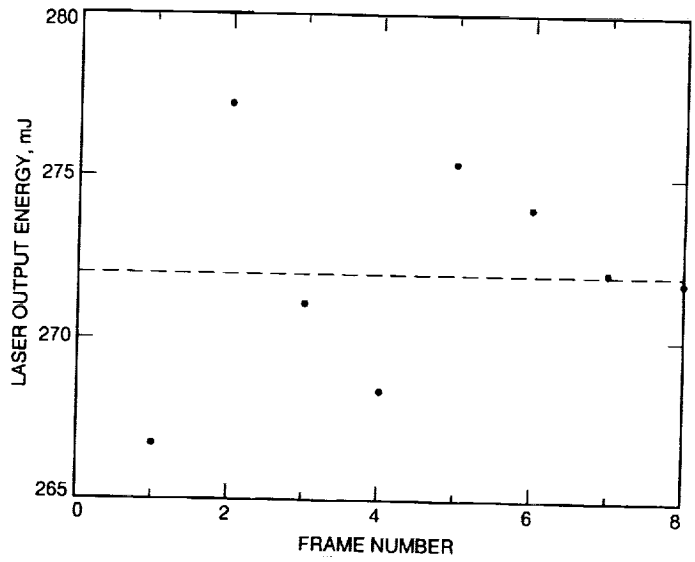


Fig. 9. GOPEX laser output for day 8.

TDA Progress Report 42-114

August 15, 1993

GOPEX at the Starfire Optical Range

R. Q. Fugate
Starfire Optical Range, Phillips Laboratory, Kirtland Air Force Base, New Mexico

The Starfire Optical Range successfully conducted laser uplink experiments to the Galileo spacecraft during the early morning hours of December 9, 10, 11, and 12, 1992, when the spacecraft was at ranges between 700,000 and 3 million km from Earth. Analysts at JPL have reported as many as 79 pulse detections by the spacecraft. The best weather conditions occurred on the second night when 37 pulses were detected with as many as five on one frame. Signal levels at the spacecraft generally agree with predictions.

I. Introduction

This article summarizes the experiment requirements, design, operations, and results obtained in the Galileo Optical Experiment (GOPEX)[1], conducted by the U.S. Air Force Phillips Laboratory at the Starfire Optical Range (SOR) near Albuquerque, New Mexico. SOR was chosen by JPL, the sponsoring agency, as a second site to complement their operations at Table Mountain Facility (TMF), in Wrightwood, near Los Angeles, California, and to provide geographic diversity, increasing the probability of success in case of bad weather.

The primary objective of GOPEX was to demonstrate that a narrow laser beam pointed at the Galileo spacecraft as it receded from Earth could be detected by the on-board Solid-State Imaging (SSI) camera. This objective was indeed achieved at ranges of approximately 700,000 to six million km from Earth. SOR successfully illuminated the spacecraft on the first four nights of the test, but unfortunately bad weather at the site halted the experiment on the last three nights. Site diversity proved to be advantageous in the experiment, since TMF was weathered out on the fourth night. A secondary objective was to measure the level and fluctuation in the laser irradiance at the spacecraft and compare the results with theoretical predictions. In general, this objective was also met with a high degree of success.

II. Experiment Requirements

The TMF and SOR sites were each required to transmit bursts of laser pulses on a preset schedule. Each burst lasted approximately three seconds and was computed to start so that pulses arrived at the spacecraft centered about the camera's shutter opening. Individual laser pulses were synchronized within one millisecond of WWV time. Spacecraft-camera shutter-opening times varied from 133–800 msec on a preprogrammed schedule that operated from the internal clock, which was also synchronized with WWV time. The camera was programmed

to scan along a path parallel to the Earth's terminator to spatially separate individual laser pulses on the focal plane. TMF and SOR never operated at the same laser pulse rate, making it possible to uniquely determine each site by measuring the pixel spacing between laser pulse detections.

Uplink operations occurred just before dawn on December 9, 10, 11, 12, 14, 15, and 16, 1992. The uplink times put the SOR very close to the terminator. Table 1 lists for each experiment day the start and end times, the number of transmissions, and the time between transmissions. At a pulse rate of 10 pulses per second, 4710 pulses in total were scheduled to be transmitted toward the Galileo spacecraft from SOR.

The GOPEX Task Manager required that certain diagnostic information be recorded during the uplink transmissions. This information included the energy and pulse width of every laser pulse; the time, to the nearest millisecond, of every laser pulse transmitted; the telescope coordinates during every pulse transmitted; the position of the steering mirror (explained below); and the coherence diameter (Fried's parameter r_0) of the atmosphere. The laser beam divergence at SOR was required to be 80 μrad full-angle during the first four nights and 40 μrad during the last three nights. SOR was required to develop an experimental technique for setting the full-angle beam divergence to better than ± 10 percent.

Navigational data for the spacecraft were given to SOR by JPL in terms of J2000 geocentric state vectors (position, velocity, and acceleration) and mean-of-date pointing predictions for SOR. The state vector data were converted to mean-of-date local mount coordinates by algorithms developed at SOR, and results were compared with JPL pointing predictions. In general, agreement was better than 2 μrad. Consequently, the SOR algorithms were used to point the telescope since they continuously updated the mount pointing. The mount model was validated and occasionally updated by centering the image of a nearby guide star in the field of a CCD camera between propagations. SOR was required to develop a technique to boresight the laser to the CCD guide-star camera to within 5 μrad. SOR was also required to demonstrate these capabilities during precursor tests using high-altitude Earth-orbiting artificial satellites during a dry run.

III. Description of Experiment Hardware

A. General Layout

Figure 1 shows the overall arrangement of the experimental setup at SOR. The laser-transmitting aperture is a 1.5-m (60-in.) Cassegrain telescope with a coudé path, mounted on elevation-over-azimuth gimbals set on an 8-m-tall hollow pier. The laser and tracking sensors are located in the coudé room on the ground floor of the facility. Three fiber-optic source simulators, located in the pier, are used to set the two values of the laser beam divergence and to represent a star at infinity. The source simulators can be moved into and out of the optical beam path to an angular accuracy of approximately 0.5 μrad, as measured in the output space of the telescope.

B. Telescope and Optics

The 1.5-m telescope is a classical Cassegrain with a parabolic primary mirror and a hyperbolic secondary mirror. The primary mirror has a focal length of 2.2882 m. It is coated with aluminum and a protective silicon monoxide overcoat. The secondary mirror has a focal length of -0.1486 m and a conic constant of -1.028072. The output of the telescope is an f/217 beam, approximately 10 cm in diameter (an angular magnification of ~ 15). The secondary mirror and all coudé mirrors are coated with Denton Vacuum enhanced silver FSS-99 coating.

Light from the telescope (or a laser beam projected by the telescope) is relayed through a coudé path in the center of the pier to the optics room, which is located on the first floor of the facility. Since the telescope is normally used with adaptive optics, the relay optics reimage the primary mirror of the telescope onto a deformable mirror located on the optics table in the coudé room. No adaptive optics were used in this experiment and the deformable mirror was kept in a "system-flat" mode which removed systematic optical aberrations (approximately 1/10 wave) in the system. Figure 2 shows the coudé path optics and M8, the first element in the imaging relay, a spherical mirror having a focal length of 6.21 m used at a 3.2-deg angle of incidence. This figure also shows the image plane for objects at infinity and the locations of the movable source simulators. Two of the simulators were used to set the beam divergence of the laser to either 80 or 40 μrad, as described later. The simulator representing a source at infinity is at a location along the coudé path that produces the minimum wavefront curvature at the output of the wavefront sensor, as compared with a reference wavefront source located on the optics table. By definition, this sets the location of the infinity source simulator. During telescope operations, the secondary mirror position of the telescope is adjusted (while observing a star) to minimize wavefront curvature as reported by the wavefront sensor.

Figure 3 is a schematic diagram showing the layout of components on the optics table in the coudé room. The

diverging beam from the pier is recollimated by an 8.45-m focal length off-axis parabolic mirror, OAP#1. The beam then reflects from a fast-steering mirror onto the deformable mirror (which is preset with a static figure to remove small residual aberrations in the system). An 11.2-cm diameter image of the telescope's primary mirror is formed on the deformable mirror. Another off-axis paraboloid, OAP#2, and a lens reimage the deformable mirror on an array of lenslets in the Shack–Hartmann sensor. This sensor is used to set the 1.5-m telescope focus by observing a bright star just prior to operations.

The pulsed laser beam is injected into the coudé path by means of a thin-film plate polarizer located between OAP#2 and the recollimating lens. The total optical transmission from the output of the laser to the atmosphere is estimated to be 43 ±3 percent. Just prior to laser propagation, the telescope is pointed to a nearby guide star. Light from the guide star passes through the laser-aperture sharing element and is imaged onto a low-noise, high-resolution CCD camera to verify telescope pointing. This camera is the primary sensor for laser boresighting and telescope pointing.

The reference source for the wavefront sensor is placed at the focus of OAP#2, since this point is optically conjugate to infinity. The laser-aperture sharing element is located in the converging beam ahead of the infinity focus. Since the thin-film plate polarizer is used in a converging beam, a glass plate was placed behind it to compensate for the astigmatism in images of the guide star at the CCD camera and during telescope defocus measurements made with the wavefront sensor.

The fast-steering mirror was used to offset the laser pointing direction in a predetermined pattern to increase the probability of detection in the event that the navigation data were in error. The mirror was repositioned between laser pulses to generate either a hexagonal or square pattern, as shown in Fig. 4. These scan patterns were used only on the first night of operations. The scan patterns put the nominal position of the spacecraft in the edge of the beam.

C. Optical Alignment

The basic optical alignment requirements for GOPEX were to (1) establish the optical axis of the system, (2) set the full-angle laser beam divergence to either 80 or 40 μrad, and (3) accurately boresight the laser to the optical axis of the system.

The optical axis of the system was defined in tilt by the CCD guide-star camera and in translation by the center of the entrance pupil of the telescope. The required laser beam divergence was generated by focusing the 1.5-m-diameter beam in the atmosphere at ranges of 18.75 and 37.5 km, respectively. These ranges can be simulated at the appropriate conjugate points in the path of the relay-imaging optics in the pier. Based on the optical design of the relay optics, these points are 64.14 cm and 32.703 cm below the location of the infinity focus where a fiber-optic star simulator is located on a stepper motor-driven stage. The laser beam will come to focus at these points in the coudé path when the divergence is properly adjusted. Furthermore, a source accurately positioned at these points is a fiducial for boresighting the laser to objects at infinity imaged on the optical axis of the telescope. Two 50-μm-diameter optical fibers were placed on precision slide stages at these points. The arrangement of the source simulators is shown in Fig. 5. Light transmitted by the fiber was imaged by the CCD guide-star camera and allowed positioning of the stages to approximately 0.5 μrad in the output space of the telescope. The vertical position of the fiber was measured mechanically with an uncertainty of ±5 mm.

Beam divergence was set by using a knife-edge test on the focused beam and observing the pattern in the plane of the fiber. This technique produces no more than ±0.5 wave of focus error. The telescope focus error is less than ±0.25 wave, including higher order aberrations in the optical system between the star simulator and the telescope exit. Assuming worst-case additive errors, the divergence error is ±2.8 μrad or 6 percent at 40 and 3 percent at 80 μrad full-angle beam divergence. Final beam boresighting was set by maximizing the light injected into the fiber from the focused laser beam. Beam motion of ±0.5 μrad completely extinguishes laser light coming out of the fiber. It was estimated that all error sources would make the worst-case boresight error ±1.75 μrad for the 40-μrad beam-divergence case and ±2.25 μrad for the 80-μrad beam-divergence case. The actual beam divergence was verified by scanning the beam across high-altitude Earth-orbiting satellites equipped with retroreflectors.

D. Laser Characteristics

The laser used for these experiments was a frequency-doubled neodymium:yttrium-aluminum-garnet (Nd:Yag), Spectra-Physics Quanta-Ray DCR-2A, field-modified to the equivalent of a DCR-3G. The laser was equipped with Spectra-Physics' unstable Gaussian Coupled Resonator using Radially Variable Reflectivity coatings. This resonator produces a beam profile shaped more like a "top hat" than gaussian. This feature makes it easier to relay

C-4

through the optics and produces a more uniform intensity pattern at long ranges. The measured intensity profile in a plane equivalent to approximately 700,000 km is shown in Fig. 6. The laser pulse width was 14.5 nsec (full-width half-maximum), and the energy per pulse was 318 ±10 mJ per pulse.

E. Telescope Pointing

The two-axis mount of the 1.5-m telescope is controlled by a microcomputer that is designed to accept data on an object's position from imaging cameras or a track processor. The microcomputer is equipped with relatively simple, but very effective, algorithms that compute angular positions and rates of low and high Earth-orbiting artificial satellites, as well as astronomical objects. The computer code is able to modify in real time the orbital parameters of satellites based on measurements of the satellite's position by imaging cameras or trackers.

The routines for pointing the telescope at a selected guide star and the computed position of Galileo were automated in a script that was executed by the telescope control computer. Thirty seconds prior to propagation, the script automatically pointed the telescope to the computed position of Galileo, and ten seconds after the end of the propagation, it repointed the telescope to the guide star. The position of the azimuth and elevation axes were recorded at the transmission time of each pulse.

F. Laser Diagnostics

The laser pulse width and energy were monitored by a Hamamatsu vacuum photodiode,[2] calibrated against a thermopile radiometer. Light to the vacuum photodiode consisted of the leakage through a turning mirror in the laser-beam injection optics. The thermopile radiometer was placed in the unattenuated beam. The output of the vacuum photodiode was digitized by a 1-GHz sample-rate digital oscilloscope at 1-nsec intervals and saved to a computer file. The pulse width was then computed from the digital data and the pulse energy was computed from the integral under the power-versus-time plot generated by the oscilloscope. The time of the trace was tagged to an accuracy of one millisecond by reading a WWV clock.

G. Communications and Data Transfer

Real-time communications between GOPEX control and SOR were via a dedicated phone line. A JPL representative was on-site to handle communications and monitor JPL control for permission to propagate, for unexpected abort commands, and to report the status of each propagation to the GOPEX Task Manager. Backup communications consisted of telephones and fax machines over commercial phone lines.

Prior to operations and between experiment days, Galileo navigational data and position predictions were exchanged over the Internet between JPL and SOR computers. This computer network was also used to pass downlinked Galileo images from JPL to the SOR in near real time during operations.

H. Atmospheric Data

Separate instruments were used to monitor the atmospheric conditions during operations. The measurements made included Fried's coherence length, r_0; the isoplanatic angle, θ_0; and the atmospheric extinction using a lidar receiver to measure the strength of the atmospheric backscatter from each laser pulse. The coherence length and isoplanatic angle are measured by making modulation-transfer-function and scintillation measurements of light from a nearby bright star.

I. Data Recorded

The data recorded during the operations included

(1) The time of the laser pulse to the nearest millisecond.

(2) Instantaneous laser power versus time digitized in 1-nsec bins.

(3) The telescope's azimuth position.

(4) The telescope's elevation position.

(5) The scan mirror's position off boresite.

(6) The value of r_0.

(7) The value of θ_0.

(8) The lidar backscatter signal.

IV. Precursor Tests

Several propagation tests were conducted prior to operations with Galileo. SOR used Lageos and the Etalon artificial satellites and observed the retro-reflected signal return with a photomultiplier. The objectives of these tests were to (1) verify laser beam divergence and boresighting, (2) verify proper operation of the fast-steering mirror to scan the beam, and (3) get a rough idea of the beam profile.

[2] Model number R1193U.

On the mornings of October 1 and 2, 1992, successful laser uplink tests to Etalon 2 (Cosmos 2024) were conducted using 80- and 40-μrad full-angle beam divergences. Returned signals were detected by a photomultiplier and outputs were saved on a digital oscilloscope. The photomultiplier was calibrated to allow an estimate of the number of photons detected. The beam was scanned across the satellite to measure beam divergence and boresighting and get a rough idea of the beam profile. The returned signal of the 40-μrad beam was, on average, 3.6 (versus an expected value of 4) times stronger than the 80-μrad beam. This represents a combined beam divergence discrepancy of 5 percent, well within the $\pm$10-percent requirement set by JPL.

The scintillation of the return signal was quite severe, varying more than an order of magnitude. Average returns were approximately 400 detected photons for a 300-mJ laser pulse. The data-recording equipment did not permit collecting the hundreds or thousands of detections required to amass adequate statistics on beam-profile mapping. However, when the beam was moved in 10-μrad steps from boresight, one could easily see a sudden drop in the return signal to an undetectable level at the predicted position at the edge of the beam. Signal return was nearly constant over a 30- to 40-μrad radius for the 80-μrad beam and dropped precipitously below 40 μrad until it was completely undetectable at a 50-μrad radius. A bias of approximately 20 μrad was observed along the track of the satellite, which was consistent with the expected point-ahead angle.

On the morning of October 2, 1992, a 40-μrad beam was propagated to Etalon 2. The telescope had to be pointed 23 μrad ahead of the apparent position on the CCD camera. Without point-ahead correction, no detected signal was seen (consistent with a 40-μrad full-angle beam divergence). By moving the telescope 20 μrad off-center and observing a complete loss of signal, it was further verified that the beam was not more than 40 μrad in diameter. Also, the fast-steering mirror was implemented in a 20-μrad square pattern, which demonstrated the expected effect of scanning the beam. When the beam was centered on the satellite, no periodic time variation was seen in the return signal (mentally averaging the scintillation). When the beam was not centered on the satellite, one could see a definite cyclic temporal pattern in the return signal, which indicated that the satellite was being hit on only one position of the scan.

Additional precursor tests were performed on the morning and evening of October 26 using Lageos and Etalon at the 80-μrad beam divergence. Return signal levels were approximately a factor of 25 times stronger from Lageos than Etalon, as expected from the difference in range to the satellites. The beam was step-scanned again with the fast-steering mirror to demonstrate the desired effect.

A full dress rehearsal was conducted on the morning of November 18. All communications circuits and procedures were effected as planned for actual GOPEX operations. The SOR Test Director conducted operations according to a timeline-based checklist. No major problems were encountered, and the checklist was executed well ahead of schedule. The telescope script worked flawlessly, and with the exception of one 4-μrad correction, telescope-pointing corrections were unnecessary. The timing and the scan mirror scripts worked flawlessly. Laser alignment held throughout the test to better than 0.5 μrad. Atmospheric data were collected, and the weather was perfect. The dress rehearsal resulted in a few minor changes to the checklist and improvements in communications with JPL operations.

V. Galileo Operations

A. Overview

The biggest problem at the SOR during Galileo operations was the weather. Of the seven test nights, it was reasonably clear on only one night (the second night). The site was fogged in during the mornings of the last three experiment days, preventing any propagations. Fog is *not* the norm for Albuquerque, a city that experienced more precipitation in December 1992 than in any December in the previous 100 years! At times researchers were propagating through cloud cover so heavy that the guide star was not visible on the CCD camera. Furthermore, on the first experiment day, the relative humidity was so high that to prevent condensation, between propagations a hand-held heat gun had to be used to blow warm air on the secondary mirror of the telescope. In the worst conditions, snow was falling or fog was condensing into snow and falling into the open dome.

Despite the bad weather, SOR successfully conducted operations on the first four experiment days. Table 2 summarizes SOR pulses detected by Galileo. These data, from an article by B. M. Levine, K. S. Shaik, and T.-Y. Yan of JPL summarize the analysis of the GOPEX images [2]. No pulses were detected by Galileo from TMF or SOR for camera-shutter opening times less than 400 msec. Furthermore, there were always fewer pulse detections than possible for shutter times of 400, 533, and 800 msec. One explanation is that the scan motion of the camera on Galileo was not perfectly synchronized with the shutter opening.

B. Operations Procedures

Activities to prepare for, conduct, and assess the nightly operations were based on a test director's checklist and timeline designed to allow ample time to correct minor problems. Appendix A is a facsimile of the test director's checklist for day 344, the first test day.

In general, a test day involves facility preparation; equipment turn-on and warm-up; functional equipment checkout; computer disk-space and directory setup; optics and laser alignment; integrated system checkout; final preparations and double checks; conducting the experiment; postexperiment debriefing; data quick-look; and identification of problems to be fixed. Many of the details of these tasks can be gleaned from the timeline in Appendix A.

C. The First Test Day, December 9, 1992

Sixty propagation sequences were planned for the first test day. The first propagation was at 11:13:35 UTC and every three minutes thereafter until 14:12:32 UTC. Thirty pulses were transmitted during each sequence. On many of these sequences, the fast-steering mirror was stepped between pulses to generate one of the two patterns shown in Fig. 4.

Appendix B contains a sample of the summary of the propagation sequences, two graphs showing plots of each pulse in each propagation sequence of the measured pulse energies and pulse widths, and a sample output from a spreadsheet summarizing the laser diagnostic and telescope-pointing data for each pulse transmitted.

The propagation sequence summary that appears in Appendix B also lists the sequence number; the day number; the time of the first pulse, to the nearest millisecond; a propagation-time correction offset, if needed; the Galileo shutter time; the number of shots in a repeating sequence with no scan-mirror offset; the number of shots in a sequence at some offset radius; the radius size; and comments made during operations after each propagation sequence.

Appendix C contains plots of environmental conditions recorded at the site during Galileo operations. The weather was generally not good the first night. It had been cold (a few degrees above freezing) and rainy all day. After sunset, massive fog set in and in the early part of the evening the relative humidity was nearly 100 percent. It was not possible to open the facility for temperature conditioning, as scheduled, due to the high humidity. At around 08:30 UTC, the sky began to clear and the wind picked up, blowing low-lying clouds to the southeast. However, the sky was too cloudy to permit using a star to set the focus of the telescope with the wavefront sensor. It was necessary to focus the telescope just before the first propagation, based on previous experience and the best image at the guide-star CCD camera. During the propagation sequences, the relative humidity averaged 82 percent. Between propagations, a person (standing atop a stepladder in the dark) directed warm air over the secondary with a hand-held heat gun in order to prevent condensation on the secondary mirror's surface. The temperature plot of the secondary mirror in Appendix C (the plot for temperature sensor TS037, December 9, 1992) shows this process. The data in Appendix C also show that the temperature in the pier (sensor TS030 at the source simulators) averaged a little over 13 deg C, while the outside temperature (sensor TS006) was approximately −1.5 deg C, a very large gradient indeed. These large temperature variations had an unknown, but certainly degrading, effect on the optical quality of the transmitted beam. It was not possible to make any r_0 or θ_0 measurements on the first night due to equipment malfunction.

B. M. Levine, of JPL's Optical Sciences and Applications Section, has analyzed the images from Galileo to determine which frames show detections and to measure their strength with respect to the background. He reports that Galileo detected pulses from SOR on propagation sequences 1, 13, 16, 17, 20, 28, and 32.[3] Note from the comments in the propagation sequence table in Appendix B that the cloud cover was so thick that it was not possible to see the guide star between sequences 4 and 12. The scan mirror was on during sequences 16, 20, 28, and 32, and off during the other sequences. A summary appears in Table 3. The signal levels reported by JPL are included in this table. The average signal from TMF was data number (dn) 199.8, a value comparable to dn 173.8 from SOR. The high standard deviation (dn 212.2) of the signal variability could be due to the fact that most of the pulses were transmitted while the beam was being scanned.

D. The Second Test Day, December 10, 1992

This was the best test day at the SOR. The sky was nearly clear except for a very thin subvisible cirrus cloud layer at the 17.5-km range, which was present during the first 19 or 20 propagations. The relative humidity was still much higher than normal, averaging nearly 70 percent during the propagations. Generally, everything worked per-

[3] B. M. Levine, private communication, Optical Sciences and Applications Section, Jet Propulsion Laboratory, Pasadena, California, December 22, 1992, updated by further private communication.

fectly on this night. Every pulse was transmitted, and atmospheric data were collected for every laser transmission.

Table 4 summarizes the pulse detections by Galileo. There were 37 detections with the average signal dn 143, a factor of more than three times higher than the average signal from TMF. The standard deviation was dn 187 and the maximum signal was dn 354. The laser energy was a bit higher, on average, for this day, and the sky was generally clear although not a "photometric night." The atmospheric seeing was not exceptional, in fact it was less than average for this site.

E. The Third Test Day, December 11, 1992

The weather was again a problem on the third night. The first 11 propagations were into very heavy clouds, and in most cases it was not possible to see the guide star. At propagation sequence number 12, the clouds thinned enough for a detection by Galileo. Detections were also made on sequences 16 and 20, which were the only other shutter openings of 533 msec. The very last propagation was into a fairly clear sky.

Table 5 summarizes the pulse detections for test day 3. Only 11 pulses from SOR were detected. The average signal level was dn 66.0 (compared with dn 54.5 from TMF). Five pulses were detected on the last sequence when the weather was clearest.

F. The Fourth Test Day, December 12, 1992

The cloud cover was variable on the fourth night. Only 10 propagation sequences were conducted. Only three of the sequences were 533 msec. The sky was clear on the first few propagations but became very cloudy after the sixth propagation.

Table 6 summarizes the detections by Galileo on frames 3 and 6. Only 5 pulses were detected. The average signal level was dn 33.6. No TMF data are available for comparison since the facility was weathered out completely on that night.

G. The Last Three Nights, December 14–16, 1992

There is nothing to report for these nights since SOR was completely fogged in on all three nights. The last recorded fog in December in Albuquerque occurred in 1937.

VI. Conclusions

GOPEX was a major success, with 268 pulse detections from TMF on six nights at 15 and 30 Hz, and 76 pulse detections from SOR on four nights at 10 Hz. The signal levels were close to those expected.

Acknowledgments

This experiment would not have been possible without the dedication and expertise of the SOR GOPEX Team. The team members and their areas of responsibility under the author's team leadership and test direction were C. Batcheller, aircraft spotter; B. Boeke, emission control and monitor engineer; R. Cleis, telescope pointing and monitor engineer; J. Drummond, lidar operation and data reduction; B. Ellerbroek, theoretical calculations; F. Gallegos, facility operator and weather monitor; J. Glover, scan mirror operator and countdown announcer; H. Hemmati, JPL representative; G. Jones, CCD guide-star camera operator; Joe Lange, propagation safety officer; P. Leatherman, lidar instrumentation; C. Morgenstern, atmospheric instrumentation; M. Oliker, data compilation and reduction; R. Ruane, laser equipment engineer; J. Spinhirne, optics design, setup, and alignment; P. Stech, Spectra Physics laser maintenance; D. Swindle, wavefront sensor operator and secondary mirror warmer; and S. Torney, aircraft spotter.

The team is also grateful to the programmatic support of Phillips Laboratory management, including John Anderson, SOR Division Chief; Bill Thompson, Ground-Based Laser Technology Program Manager; and Barry Hogge, Chief Scientist of the Lasers and Imaging Directorate. This project could not have been accomplished without their financial and managerial support. GOPEX Task Manager, Keith Wilson, and his supervisor, James R. Lesh, of the Optical Communications Group at JPL were instrumnetal in the success of the SOR Team.

References

[1] K. E. Wilson and J. R. Lesh, "An Overview of the Galileo Optical Experiment (GOPEX)," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April-June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 192–204, August 15, 1993.

[2] B. M. Levine, K. S. Shaik, and T.-Y. Yan, "Data Analysis for GOPEX Image Frames," *The Telecommunications and Data Acquisition Progress Report 42-114, vol. April-June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 213–229, August 15, 1993.

Table 1. GOPEX operations schedule.

Test day, December 1992	Start time, UTC	End time, UTC	Number of transmissions	Time between transmissions, min
9	11:13:35	14:12:32	60	3
10	11:06:21	13:04:38	40	3
11	11:10:06	12:07:44	20	3
12	10:25:24	11:19:59	10	6
14	10:42:08	11:37:45	12	5
15	10:39:54	11:25:24	10	5
16	10:39:41	11:15:04	8	5

Table 2. SOR pulses detected by Galileo.

Test day, day of year	Number of pulses detected
1, 344	16
2, 345	43
3, 346	12
4, 347	5
5, 349	No propagations due to fog
6, 350	No propagations due to fog
7, 351	No propagations due to fog

Table 3. Results for the first test day, day 344, December 9, 1992.

Propagation sequence	Sky condition	r_0, cm	θ_0, μrad	Atmospheric transmission from lidar data	Average energy per pulse, mJ	Beam scan radius, μrad	Galileo shutter time, msec	Number of pulses detected
1	Partly cloudy	No data	No data	0.77	310	0	400	1
13	Cloudy	No data	No data	0.92	312	0	400	2
16	Partly cloudy	No data	No data	0.81	312	60	800	4
17	Partly cloudy	No data	No data	0.80	315	0	400	2
20	Mostly cloudy	No data	No data	0.94	311	60	800	3
28	Good	No data	No data	0.80	315	30	400	1
32	Clear	No data	No data		317	30	400	1

Total number of detections	14
Minimum dn	10
Maximum dn	631
Average dn	173.8
Standard deviation dn	212.2

Table 4. Results for the second test day, day 345, December 10, 1992.

Propagation sequence	Sky condition	r_0, cm	θ_0, μrad	Atmospheric transmission from lidar data	Average energy per pulse, mJ	Beam scan radius, μrad	Galileo shutter time, msec	Number of pulses detected
4	Subvisible cirrus	6.45	7.52	0.67	342	0	800	3
5	Subvisible cirrus	6.77	9.16	0.69	340	0	533	2
6	Subvisible cirrus	6.45	9.59	0.71	342	0	533	2
8	Subvisible cirrus	7.21	8.25	0.74	337	0	800	5
9	Subvisible cirrus	7.55	8.88	0.76	336	0	533	2
10	Subvisible cirrus	6.45	7.45	0.79	337	0	533	2
12	Subvisible cirrus	4.43	8.49	0.78	337	0	800	4
13	Subvisible cirrus	4.07	7.58	0.76	338	0	533	2
14	Subvisible cirrus	4.74	8.93	0.75	337	0	533	2
16	Subvisible cirrus	6.48	6.76	0.76	338	0	800	5
17	Subvisible cirrus	5.06	5.81	0.77	336	0	533	2
18	Subvisible cirrus	4.96	6.64	0.76	338	0	533	1
20	Clear	6.75	8.06	0.78	337	0	800	2
28	Clear	5.73	5.62	0.80	338	0	533	1
32	Clear	6.32	8.92	0.79	336	0	533	2

Total number of detections	37
Minimum dn	14
Maximum dn	354
Average dn	143
Standard deviation dn	187

Table 5. Results for the third test day, day 346, December 11, 1992.

Propagation sequence	Sky condition	r_0, cm	θ_0, μrad	Atmospheric transmission from lidar data	Average energy per pulse, mJ	Beam scan radius, μrad	Galileo shutter time, msec	Number of pulses detected
12	Clouds	No data	No data	0.75	327	0	400	4
16	Very thick clouds	No data	No data	0.76	328	0	400	2
20	Fairly clear	9.18	4.67	0.75	327	60	800	5
Total number of detections		11						
Minimum dn		14						
Maximum dn		292						
Average dn		66						
Standard deviation dn		76						

Table 6. Results for the fourth test day, day 347, December 12, 1992.

Propagation sequence	Sky condition	r_0, cm	θ_0, μrad	Atmospheric transmission from lidar data	Average energy per pulse, mJ	Beam scan radius, μrad	Galileo shutter time, msec	Number of pulses detected
3	Clear	Est. 7.5	Est. 5.0	0.69	299	0	533	2
6	Very thick clouds	Est. 7.5	Est. 5.0	0.59	295	0	533	3
Total number of detections		5						
Minimum dn		6						
Maximum dn		81						
Average dn		33.6						
Standard deviation dn		30.1						

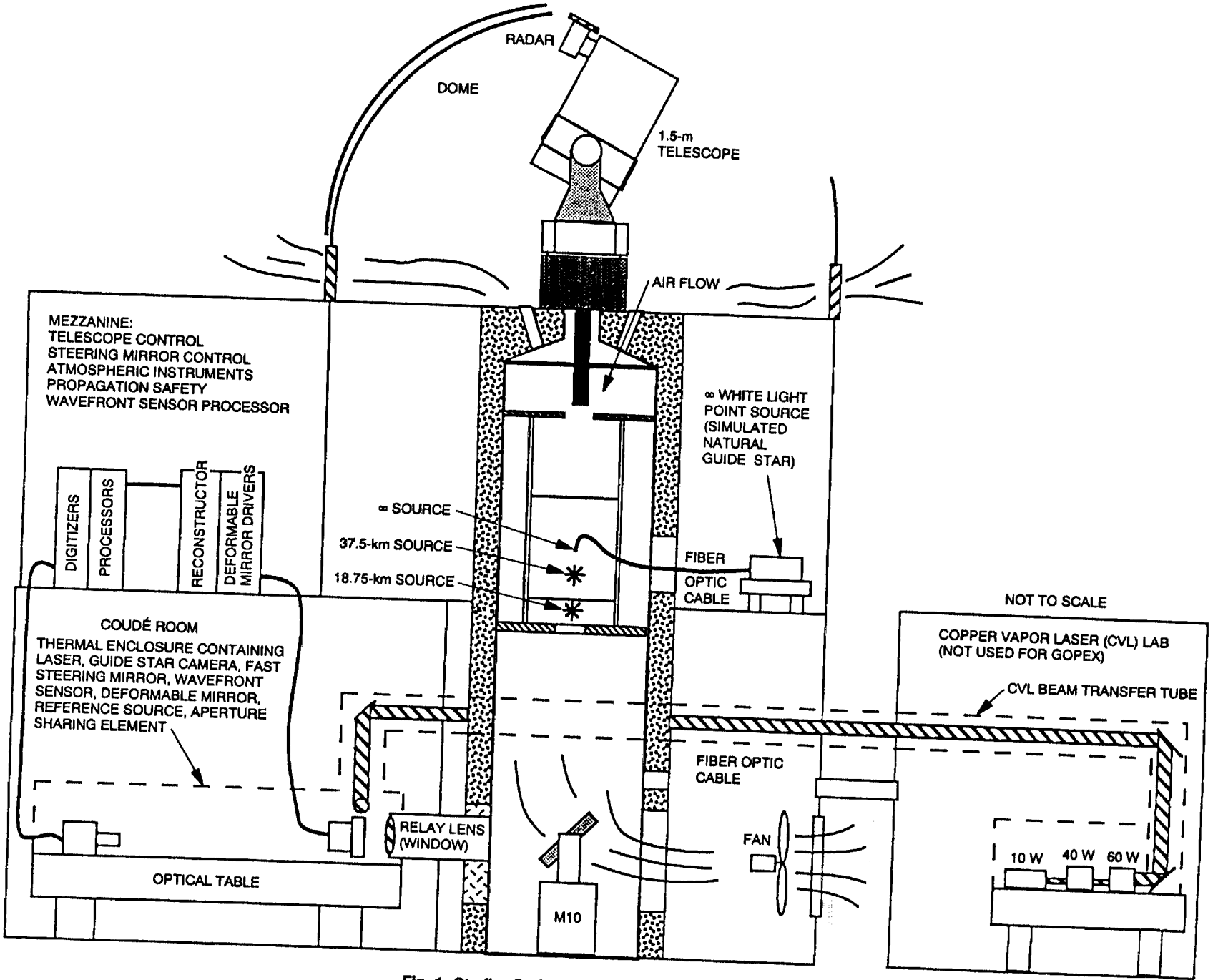


Fig. 1. Starfire Optical Range 1.5-m telescope facility.

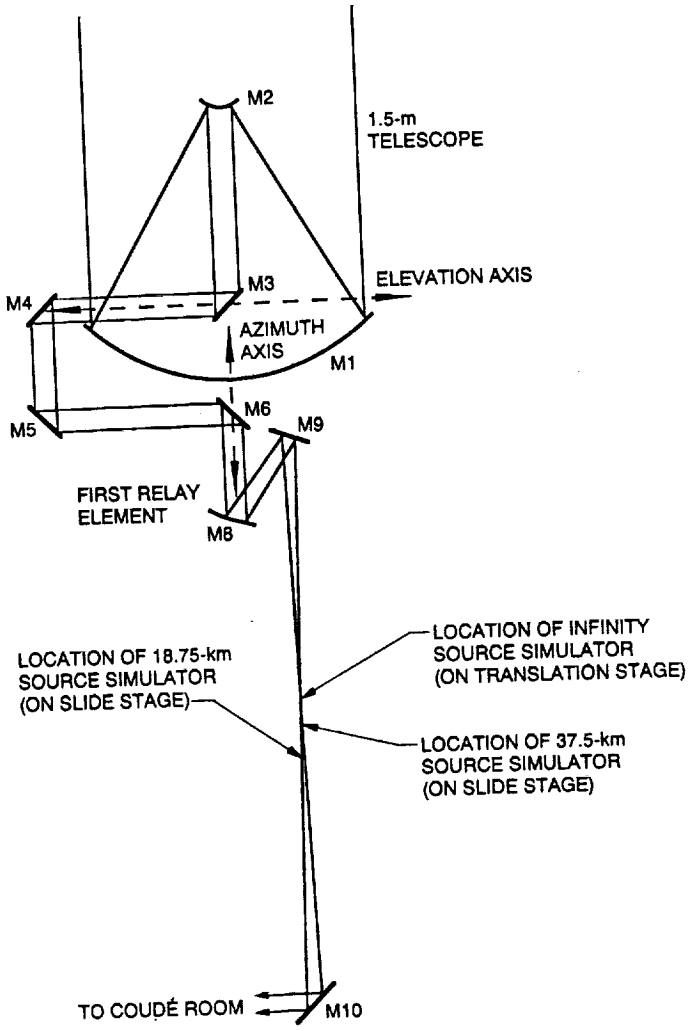


Fig. 2. Coudé path relay optics and source simulators for GOPEX.

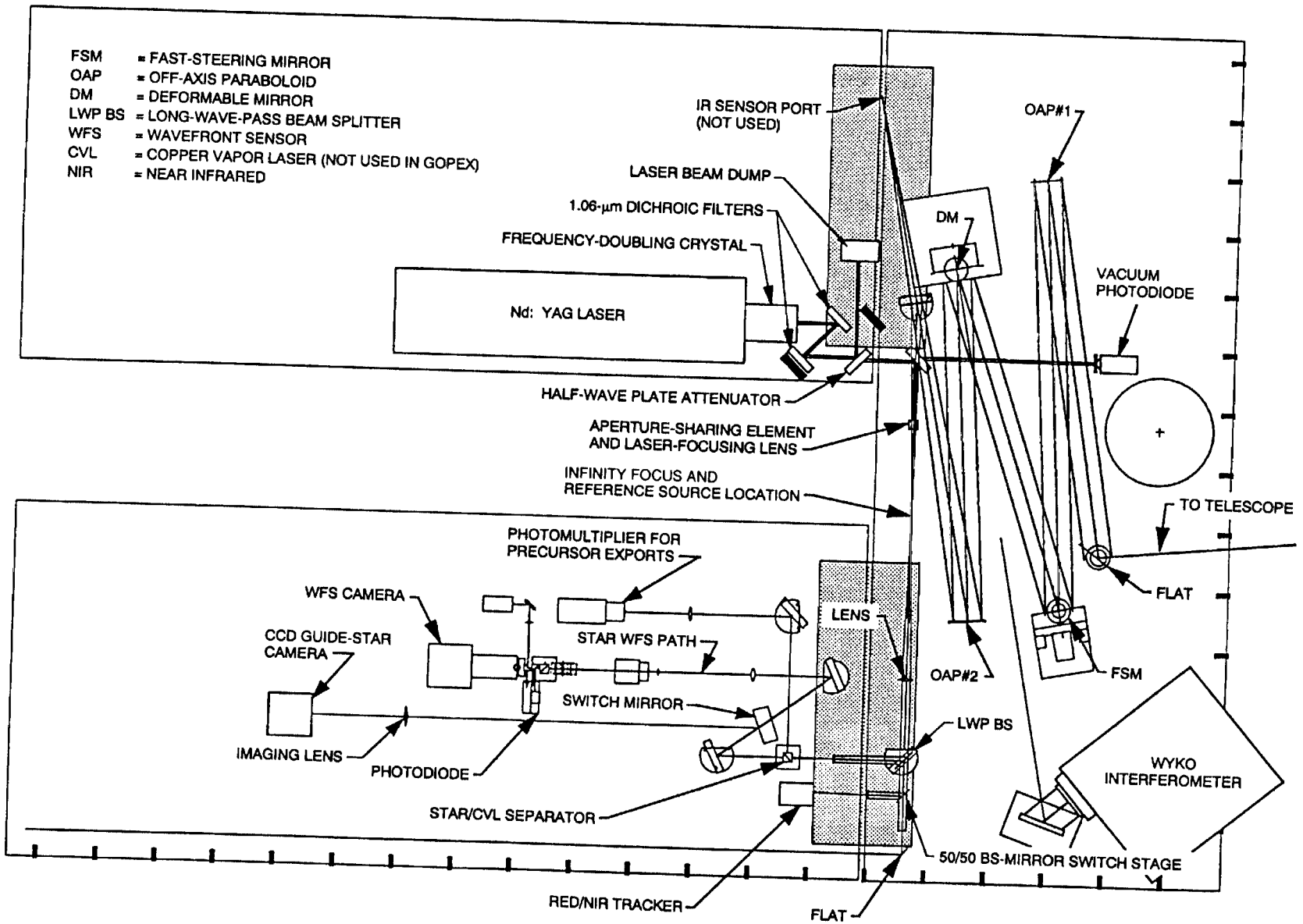


Fig. 3. Coudé room optics layout at SOR, as configured for GOPEX.

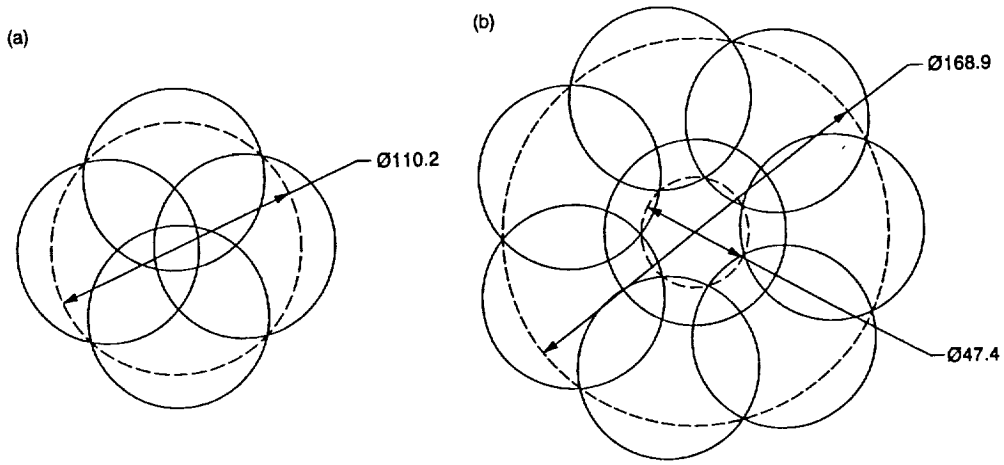


Fig. 4. GOPEX scan patterns used during the first test night, with an 80-μrad beam divergence: (a) 4-pulse mode, no pulse on center, 30-μrad offset and (b) 8-pulse mode, 1 pulse on center, 7 pulses at 60-μrad offset.

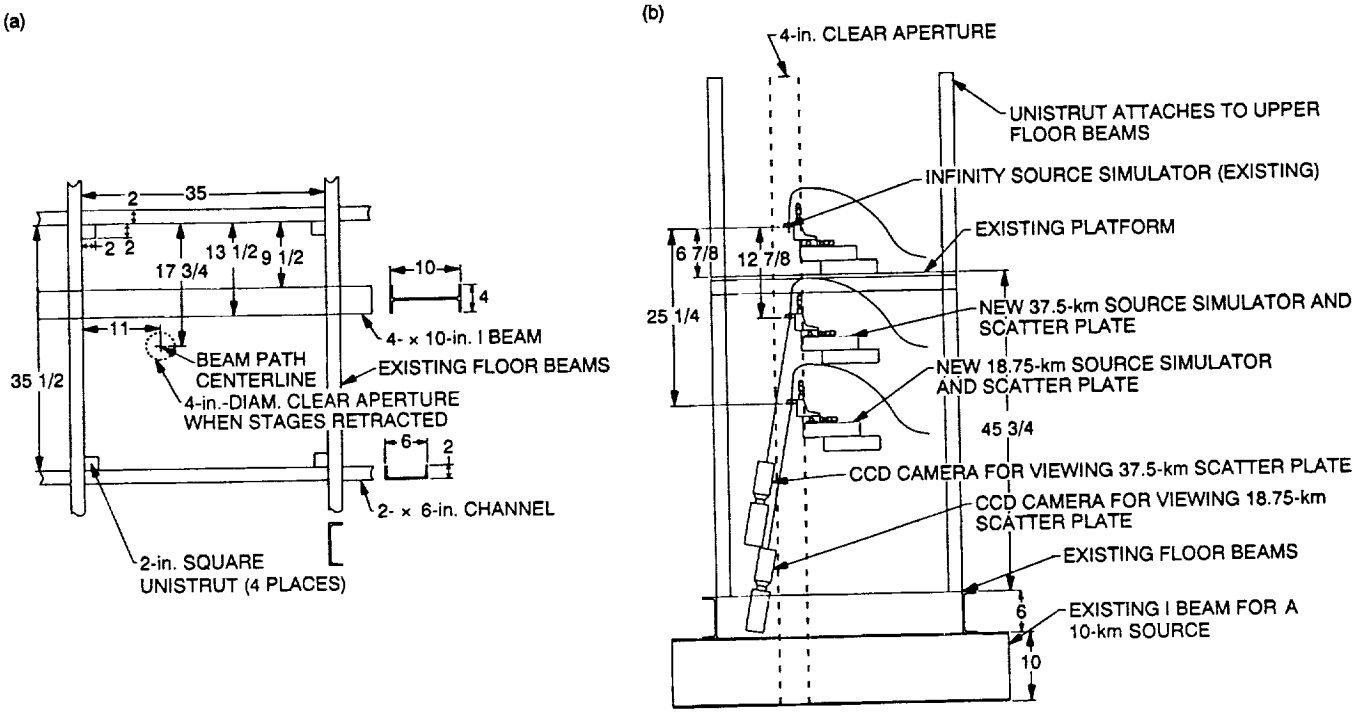


Fig. 5. 1.5-m pier area showing locations of new source simulators for GOPEX: (a) Top view and (b) side view.

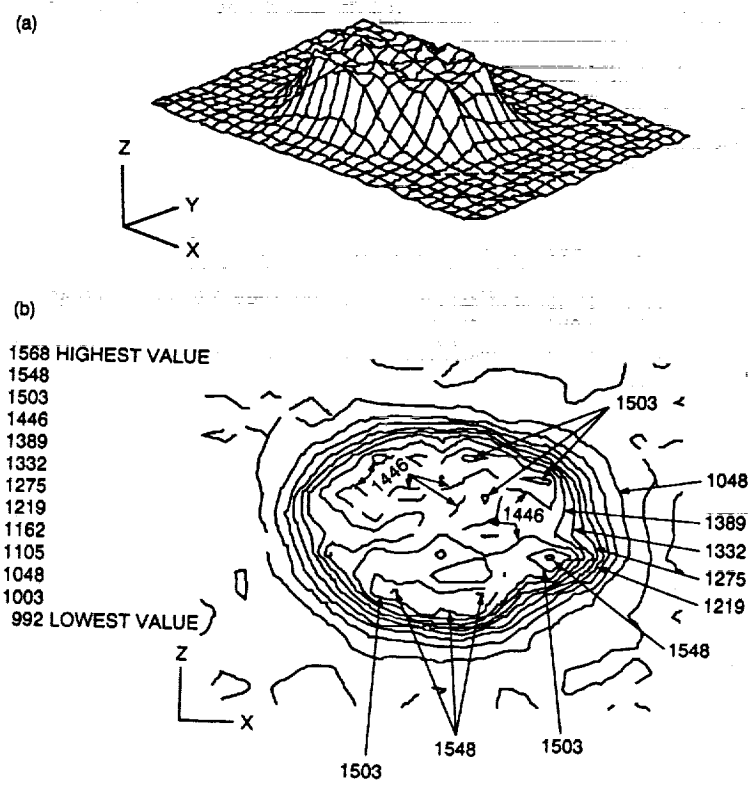


Fig. 6. Laser beam profile measured 127 cm in front of the laser. The beam is round, the distortion is due to the printer: (a) 3-D profile and (b) contour plot.

Appendix A
Test Director's Checklist

An example of the test director's checklist is shown on the following pages.

GOPEX at SOR						
Test Director's Check List						
Assigned Facility Operator (FO)	Fred Gallegos					
Test Day	344					
Beam Divergence	80					
First Propagation	11:13:34					
			Back-up people present			
Test Director	✓ Bob Fugate					
JPL Representative	✓ Hamid Hemmati					
Facility operator	Fred Gallegos					
Spotter 1	✓ Curt Batcheller					Told not to come in due to WX
Spotter 2	✓ Steve Torney					
Safety Officer	✓ Joe Lange					
Telescope Operator	✓ Rick Cleis					
Lidar Operator	✓ Phil Leatherman		Jack Drummond.			
r0 Operator	✓ Carolyn Morgenstern		Mina Brady			
Wavefront Sensor Operator	✓ Dave Swindle					
Photometerics Operator	✓ Gary Jones					
Optics and Laser Operator	✓ Jim Spinhirne					
Laser Diagnostics Operator	✓ Bruce Boeke					
Laser Maintenance	✓ Paul Stech					
Data Reduction	✓ Mike Oilker					
		Scheduled Start	To be completed	Completed by	Actual Completion	
Task	T-time	time (UTC)	by (init)	(initials)	Time	Comments
FACILITY PREPARATION						
Open domes	8:00	3:13	FO			
Open dome shutters	7:45	3:28	FO			Delayed due to FOG/High Humidity
Uncover 1.5 m tele	7:40	3:33	FO			Delayed due to FOG
Uncover r0 tele	7:35	3:38	FO			Delayed " " "
Turn on pier fan	7:30	3:43	FO			Delayed
Record site temps	7:25	3:48	FO			Delayed
Record wind data	7:20	3:53	FO	RQF	03:50	
Record RH data	7:15	3:58	FO	RQF	03:50	
Check all sky camera	7:10	4:03	FO	RQF	03:55	
Turn on WEFAX	7:05	4:08	FO	RQF	02:30	Turn on early for pics/prognosis
Turn on telescope control computer	7:30	3:43	WJL		04:15	Delayed
Turn on 1.5 m tele control console	7:20	3:53	WJL		05:20	Delayed
Turn on safety officer's console	7:10	4:03	WJL		04:15	Delayed
Turn on aircraft detection radar	7:05	4:08	WJL			Delayed
Perform radar check	7:00	4:13	WJL			Delayed.
Pre-test Briefing	6:30	4:43	RQF	RQF	05:20	Bruce Boeke absent
EQUIPMENT TURN-ON	6:00	5:13	TEAM			
Spotter comm and kill switch	6:00	5:13	CB	WJL	05:21	
r0 telescope and control computer	6:00	5:13	CM			
r0 instrumentation computer	5:50	5:23	CM			
LIDAR receiver electronics	6:00	5:13	PL	PL	05:25	
Annotation computer	6:00	5:13	RAC	—	already on	
Wavefront sensor camera and cooler	6:00	5:13	MDO	DWS	05:28	
Wavefront sensor control computer	5:55	5:18	MDO	DWS	05:28	
Real time digital reconstructor	5:50	5:23	MDO	DWS	05:30	
Digital reconstructor control computer	5:45	5:28	MDO	DWS	05:32	
Digital reconstructor diagnostic computers	5:38	5:35	MDO	DWS	05:35	
Photometrics camera and cooler	6:00	5:13	GJ	GJ	05:25	
Photometrics control computer	5:50	5:23	GJ	GJ	05:26	
Tilt mirror power supply and electronics	5:40	5:33	RAC	RAC	05:20	
Timing and tilt mirror control computer	5:35	5:38	RAC	RAC	05:22	
Tracker electronics	5:28	5:45	RAC	RAC	05:22	
Vacuum photodiode high voltage power supp	6:00	5:13	BRB	BRB	05:40	
Digital oscilloscope for pulse monitoring	5:55	5:18	BRB	BRB	05:40	
Laser diagnostics computer	5:50	5:23	BRB	BRB	05:40	
Laser water chiller and heat exchanger	6:00	5:13	RER	RER	05:50	
Laser power supply and control electronics	5:50	5:23	RER	RER	05:50	
FUNCTIONAL EQUIPMENT CHECK	5:30	5:43	RQF			
CB (spotter equipment)	5:30	5:43	RQF	RQF	05:35	
WJL (all safety equipment)	5:28	5:45	RQF	RQF	05:40	
CM (r0 instrumentation)	5:26	5:47	RQF	RQF	X	Delayed due to WX-A/D Failure
PL (LIDAR electronics)	5:25	5:48	RQF	RQF	X	Delayed " " " - Nopreclock possible
MDO (wavefront sensor and reconstructor)	5:24	5:49	RQF	DWS	05:40	
GJ (photometrics camera)	5:22	5:51	RQF	RQF	06:00	
BRB (laser diagnostic equipment)	5:20	5:53	RQF	RQF	06:15	
RAC (telescope and control electronics)	5:18	5:58	RQF	RQF	05:50	
RAC (tilt mirror and laser timing)	5:10	6:03	RQF	RQF	05:20	
RER (laser and cooling equipment)	5:05	6:08	RQF	RQF	06:10	
FO (met equipment)	5:03	6:10	RQF	RQF	06:05	
Load pointing files	5:00	6:13	RAC	RQF	02:10	

Fig. A-1. Test director's checklist for GOPEX, day 344, December 9, 1992.

Load timing files	4:45	6:28	RAC	RAC	02:10	No new data today
Generate timing scripts (only for new data)	5:00	6:13	RQF	RQF	—	
Validate pointing and timing files	4:30	6:43	RAC	RAC	02:00	
SET-UP DISK SPACE ON COMPUTERS						
Telescope control computer	4:15	6:58	RAC	RAC	05:30	
Timing and tilt mirror control computer	4:15	6:58	RAC	JMS	05:30	
Photometrics control computer	4:15	6:58	GJ	GJ	06:10	150 MB available.
Laser diagnostic computer	4:15	6:58	BRB	BRB	06:20	
Wavefront sensor control computer	4:15	6:58	MDO	MDO	06:20	
r0 instrumentation computer	4:15	6:58	CM	X	X	Delayed. - A/D Failure
LIDAR digital oscilloscope and computer	4:15	6:58	PL	PL	06:20	8MB available (2MB x 60 min)
PREPARE OPTICS AND LASER						
Check alignment of M4 source simulator	5:30	5:43	JMS	JMS	06:47	
Check alignment of infinity source simulator	5:00	6:13	JMS	↑	06:40	
Check pupil centration	4:45	6:28	JMS		↓	
Check alignment of phorometrics camera	4:00	7:13	JMS			
Check boresight of 18.75 km source slm	3:30	7:43	JMS			
Check and adjust laser boresight and cen.	3:15	7:58	JMS			
Check and adjust laser focus	3:30	7:43	JMS	↓		
Calibrate pulse width and energy monitor	3:00	8:13	BRB, RER	BRB	06:40	
Set telescope focus using wavefront sensor	4:00	7:13	MDO		X	Not fine due to wx
Perform integrated system checkout	2:00	9:13	RQF	RQF	10:45	Delayed due to high humidity
Proof read pointing and timing scripts	1:00	10:13	RQF, RAC	RQF, RAC	02:30	
Perform final system readiness check	0:45	10:28	RQF		10:50	
Perform final laser boresight and focus	0:40	10:33	JMS	JMS	10:48	
Remove 18.75 km source simulator	0:35	10:38	JMS	JMS	1053	
Adjust half wave plate for max output	0:33	10:40	JMS	JMS	1054	
Laser flashlamps to full power	0:30	10:43	JMS	JMS	10:13	part of system check
Start telescope control script	0:30	10:43	RAC	RAC	10:11	
Start timing control script	0:28	10:45	RAC	RAC	10:13	
Send photometrics images to telescope oper	0:25	10:48	GJ	RAC	10:50	
Start laser diagnostic computer program	0:30	10:43	BRB	RAC	10:13	
Verify all systems operational	0:20	10:53	team		10:55	

CONDUCT EXPERIMENT		11:13			✓	
Check for GO/NO-GO comm	0:03	11:10	HH		✓	
Monitor and check propagation times	0:00	11:13	RQF		✓	
Monitor comm lines for NO-GO command	0:00	11:13	HH		✓	
Monitor and record anomolies in scripts	0:00	11:13	TEAM		✓	
Record photometrics images	0:00	11:13	GJ		✓	
Record quad video of laser and photometrics	0:00	11:13	WJL		✓	
Last propagation		14:13				See propagation seq for details of sequences!
NOTES:						
(1) NOT ABLE to adjust Telescope focus using WFS - Focus set to best photometric image						
(2) Used a heat gun to blow warm air across Secondary mirror between propagations						
ΔT as low as ~1°C between secondary mirror Temp and dew point was recorded.						
	T+					
POST MISSION MAINTENANCE		14:13				
Calibrate pulse width and energy monitor	0:05	14:18	BRB	BRB	14:20	
Measure laser boresight and focus	0:10	14:23	JMS	JMS	14:25	~1 μrad max
Measure alignment of photometrics and slms	0:20	14:33	JMS	JMS	14:30	< 1/2 μrad
Consolidate and back-up data files	0:05	14:18	MDO	MDO	15:00	
Post experiment de-briefing	0:30	14:43	RQF	RQF	15:00	
Identify problems to be fixed	1:00	15:13	RQF	RQF	15:10	16 A/D board, Q-Switch glitch
Institute configuration control	1:00	15:13	RQF	RQF	15:15	
Generate quick look summary, fax to JPL	1:30	15:43	RQF	RQF	17:00	
Generate database summary	5:00	19:13	MDO	MDO	22:00	36 pages of spreadsheet.

Fig. A-1. (contd)

Appendix B

Day 344, December 9, 1992

This section contains summaries of the Galileo operations activities for day 344, the first test day.

	A	B	C	D	E	F	G	H	I	J	K	L
5												
6												
7	seq						prop time	GLL shutter	#center	#radius	radius	
8	number	day	hr	min	sec	ms	correct	time	shots	shots	size	Comments
9												
10												
11	001	344	11	13	35	264	0	400	4	0	0	Partly cloudy
12	002	344	11	16	37	259	0	400	0	4	30	No transmission
13	003	344	11	19	39	154	0	200	2	0	0	cloudy
14	004	344	11	22	41	449	0	800	1	7	60	Very cloudy - super dense, cannot see the guide star
15	005	344	11	25	43	242	0	400	4	0	0	same
16	006	344	11	28	45	237	0	400	0	4	30	same
17	007	344	11	31	47	131	0	200	2	0	0	same
18	008	344	11	34	49	426	0	800	1	7	60	same
19	009	344	11	37	51	221	0	400	4	0	0	same
20	010	344	11	40	53	215	0	400	0	4	30	same
21	011	344	11	43	55	110	0	200	2	0	0	same
22	012	344	11	46	57	405	0	800	1	7	60	same
23	013	344	11	49	59	200	0	400	4	0	0	cloudy
24	014	344	11	53	01	193	0	400	0	4	30	clearing
25	015	344	11	56	03	088	0	200	2	0	0	clearing
26	016	344	11	59	05	382	0	800	1	7	60	partly cloudy - but aborted part way through
27	017	344	12	02	07	177	0	400	4	0	0	partly cloudy
28	018	344	12	05	09	172	0	400	0	4	30	good
29	019	344	12	08	11	067	0	200	2	0	0	partly cloudy
30	020	344	12	11	13	361	0	800	1	7	60	mostly cloudy
31	021	344	12	14	15	056	0	200	2	0	0	mostly cloudy, can see the guide star
32	022	344	12	17	17	050	0	200	2	0	0	partly cloudy
33	023	344	12	20	19	011	0	133	1	0	0	partly cloudy
34	024	344	12	23	21	138	0	400	0	4	30	partly cloudy
35	025	344	12	26	23	033	0	200	2	0	0	partly cloudy
36	026	344	12	29	25	028	0	200	2	0	0	good
37	027	344	12	32	26	989	0	133	1	0	0	good
38	028	344	12	35	29	117	0	400	0	4	30	good
39	029	344	12	38	31	012	0	200	2	0	0	good
40	030	344	12	41	33	007	0	200	2	0	0	good
41	031	344	12	44	34	966	0	133	1	0	0	cloudy - no guide star
42	032	344	12	47	37	095	0	400	0	4	30	clear
43	033	344	12	50	38	990	0	200	2	0	0	partly cloudy
44	034	344	12	53	40	984	0	200	2	0	0	clear
45	035	344	12	56	42	945	0	133	1	0	0	partly cloudy
46	036	344	12	59	45	074	0	400	0	4	30	partly cloudy - beam was on late
47	037	344	13	02	46	967	0	200	2	0	0	no transmission due to technical difficulties
48	038	344	13	05	48	962	0	200	2	0	0	partly cloudy
49	039	344	13	08	50	924	0	133	1	0	0	partly cloudy
50	040	344	13	11	53	050	0	400	0	4	30	partly cloudy
51	041	344	13	14	54	945	0	200	2	0	0	very cloudy
52	042	344	13	17	56	940	0	200	2	0	0	very cloudy
53	043	344	13	20	58	901	0	133	1	0	0	very cloudy
54	044	344	13	24	01	029	0	400	0	4	30	very cloudy
55	045	344	13	27	02	923	0	200	2	0	0	semi clear
56	046	344	13	30	04	918	0	200	2	0	0	clear - stopped scanning starting with this sequence
57	047	344	13	33	06	879	0	133	1	0	0	semi clear
58	048	344	13	36	09	006	0	400	0	4	30	clearing
59	049	344	13	39	10	901	0	200	2	0	0	clearing but getting very light - think lasing delayed by 2 seconds not 1 sec asked for
60	050	344	13	42	12	896	0	200	2	0	0	clearing - propagation two seconds later than time at left
61	051	344	13	45	14	856	0	133	1	0	0	clearing - propagation 1 sec later than time at left
62	052	344	13	48	16	984	0	400	0	4	30	clear - all propagations from here on are one second later than published times at left
63	053	344	13	51	18	879	0	200	2	0	0	clear
64	054	344	13	54	20	873	0	200	2	0	0	clear
65	055	344	13	57	22	835	0	133	1	0	0	clear
66	056	344	14	00	24	962	0	400	0	4	30	clear
67	057	344	14	03	26	856	0	200	2	0	0	clear
68	058	344	14	06	28	851	0	200	2	0	0	clear
69	059	344	14	09	30	812	0	133	1	0	0	clear
70	060	344	14	12	32	940	0	400	0	4	30	clear

Fig. B-1. Propagation sequences for GOPEX, day 344, December 9, 1992.

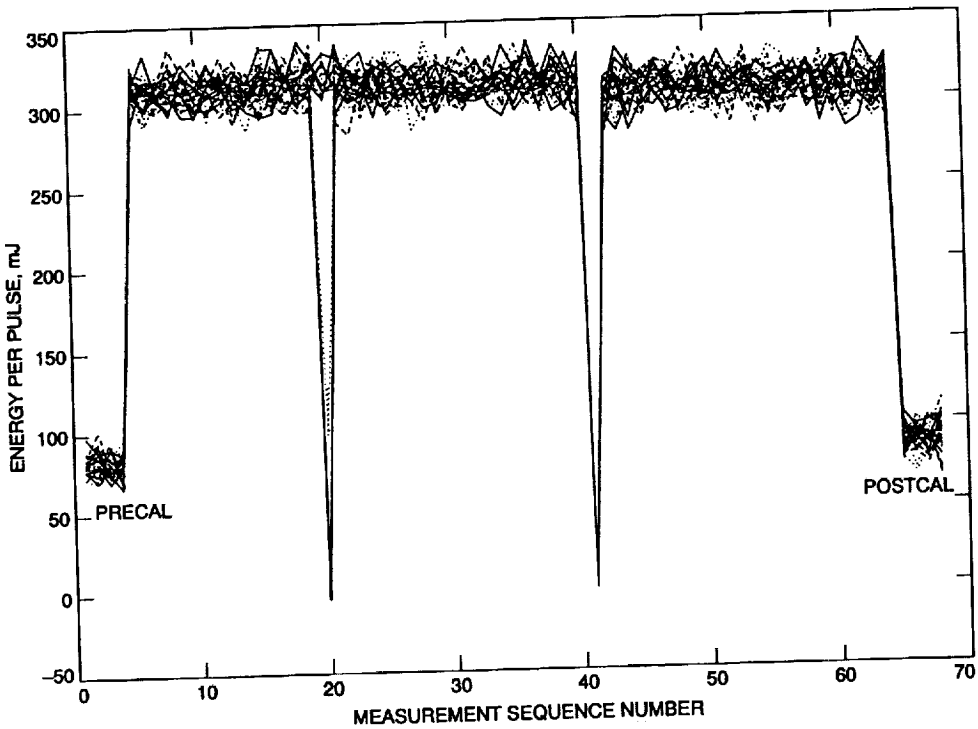


Fig. B-2. Plot of measured energy per pulse for each propagation sequence and for precalibration and postcalibration runs at lower power. The graph contains 30 points for each propagation sequence, corresponding to the 30 pulses propagated during each sequence. The atmospheric transmission sequences start at sequence number 5 and end at number 65. The drop-outs at number 20 and at number 41 were caused by laser Q-switch problems.

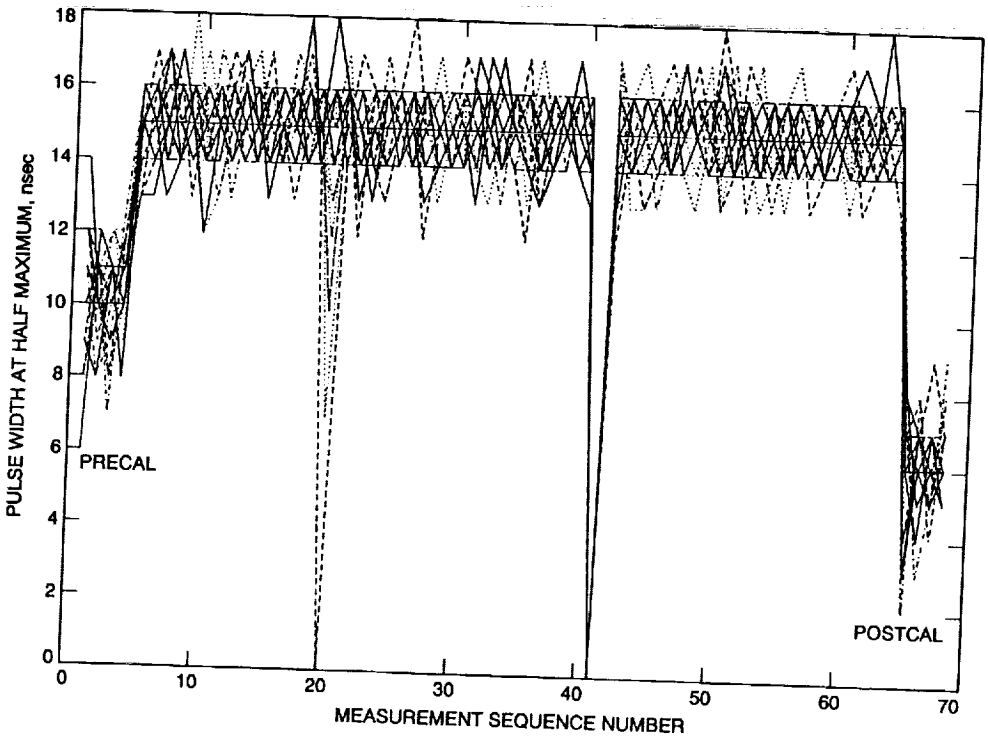


Fig. B-3. Plot of measured full-width half-maximum laser pulse widths for each of the atmospheric propagation sequences (numbers 5 through 65) and during pre- and postcalibration of the laser calorimeter (numbers 1–5 and 65–68). Thirty measurements (corresponding to 30 pulses) are plotted for each propagation sequence.

Seq Num	Day Num	Hrs	Mins	Secs	msecs	Act.-Plan msecs	Pulse Num	TIME days	ENERGY mJ	WIDTH ns	READ AZ degrees	READ EL degrees	ERROR AZ µrad	ERROR EL µrad	GOAL RA degrees	GOAL DEC degrees	OFFSET AZ µrad	OFFSET EL µrad	AN. ABER RA µrad	AN. ABER DEC µrad	REFRACT µrad	MIRROR X Pos µrad	MIRROR Y Pos µrad
1	344	11	13	35	261	-2.0	1	344.4677692361111	313	13	228.0409	38.0523	-2.1	-0.2	175.1628	-5.7346	80.0	26.6	0.0	0.0	287.9	0	0
							2	344.4677703935185	322	15	228.0419	38.0530	0.9	3.3	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.9	0	0
							3	344.4677715509259	310	16	228.0419	38.0530	0.9	3.3	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.9	0	0
							4	344.4677727083333	319	14	228.0428	38.0532	0.6	-0.8	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.9	0	0
							5	344.4677738657740	319	16	228.0428	38.0532	0.6	-0.8	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.9	0	0
							6	344.4677750231148	307	13	228.0436	38.0540	0.0	4.5	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.9	0	0
							7	344.4677761805555	320	16	228.0436	38.0540	0.0	4.5	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							8	344.4677773379962	309	16	228.0444	38.0543	-0.2	1.0	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							9	344.4677784953370	307	15	228.0444	38.0543	-0.2	1.0	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							10	344.4677796527777	322	16	228.0453	38.0548	1.0	0.1	175.1628	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							11	344.4677808101185	306	16	228.0453	38.0548	1.0	0.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							12	344.4677819675592	309	16	228.0462	38.0552	0.4	-0.9	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							13	344.4677831249999	306	15	228.0462	38.0552	0.4	-0.9	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							14	344.4677842824407	309	14	228.0471	38.0558	1.6	1.2	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							15	344.4677854398814	302	15	228.0471	38.0558	1.6	1.2	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							16	344.4677865972222	326	14	228.0479	38.0563	0.7	-0.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							17	344.4677877546629	312	15	228.0479	38.0563	0.7	-0.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							18	344.4677888912037	308	16	228.0486	38.0566	-0.9	-2.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							19	344.4677900069444	308	15	228.0486	38.0566	-0.9	-2.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							20	344.4677911226851	320	16	228.0495	38.0572	-0.8	-0.7	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							21	344.4677923384259	310	15	228.0495	38.0572	-0.8	-0.7	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							22	344.4677935541666	310	15	228.0504	38.0576	0.7	-3.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							23	344.4677946990074	312	16	228.0504	38.0576	0.7	-3.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							24	344.4677958456481	312	15	228.0513	38.0583	0.9	1.5	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							25	344.4677970013888	317	16	228.0513	38.0583	0.9	1.5	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							26	344.4677981171296	308	15	228.0521	38.0587	0.3	-0.2	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							27	344.4677993328703	302	14	228.0521	38.0587	0.3	-0.2	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							28	344.4678004861111	319	15	228.0529	38.0592	-1.3	-0.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							29	344.4678016435518	291	16	228.0529	38.0592	-1.3	-0.1	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
							30	344.4678028009925	318	13	228.0538	38.0597	-0.5	-0.4	175.1627	-5.7345	80.0	26.6	0.0	0.0	287.8	0	0
2	344	11	16	37	257	-1.0	1	344.4698756571296	333	13	228.8232	38.5027	-0.6	0.3	175.1554	-5.7326	80.0	26.6	0.0	0.0	283.2	0	0
							2	344.4698768028703	313	16	228.8241	38.5032	-0.8	1.2	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							3	344.4698779486111	318	15	228.8241	38.5032	-0.8	1.2	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							4	344.4698791143518	308	14	228.8250	38.5035	-0.6	-3.5	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							5	344.4698803009925	311	16	228.8250	38.5035	-0.6	-3.5	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							6	344.4698814583333	317	16	228.8257	38.5041	-2.5	-1.6	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							7	344.4698826157740	322	15	228.8257	38.5041	-2.5	-1.6	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							8	344.4698837731148	309	16	228.8267	38.5047	-0.3	0.3	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							9	344.4698849305555	315	15	228.8267	38.5047	-0.3	0.3	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							10	344.4698860879962	319	16	228.8275	38.5052	-1.4	1.7	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0
							11	344.4698872453370	293	15	228.8275	38.5052	-1.4	1.7	175.1554	-5.7325	80.0	26.6	0.0	0.0	283.2	0	0

Fig. B-4. First-night results for GOPEX, day 344, December 9, 1992.

Appendix C

Sample of environmental data collected during transmissions.

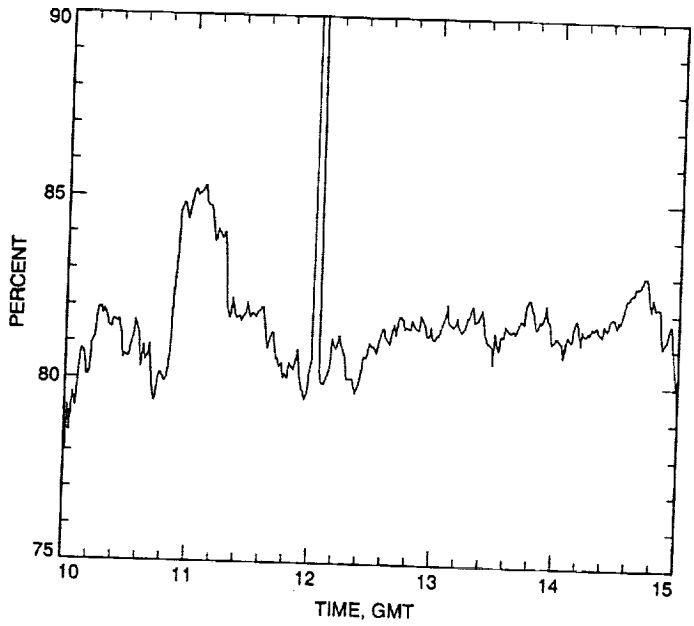


Fig. C-1. Computed relative humidity, day 344, December 9, 1992.

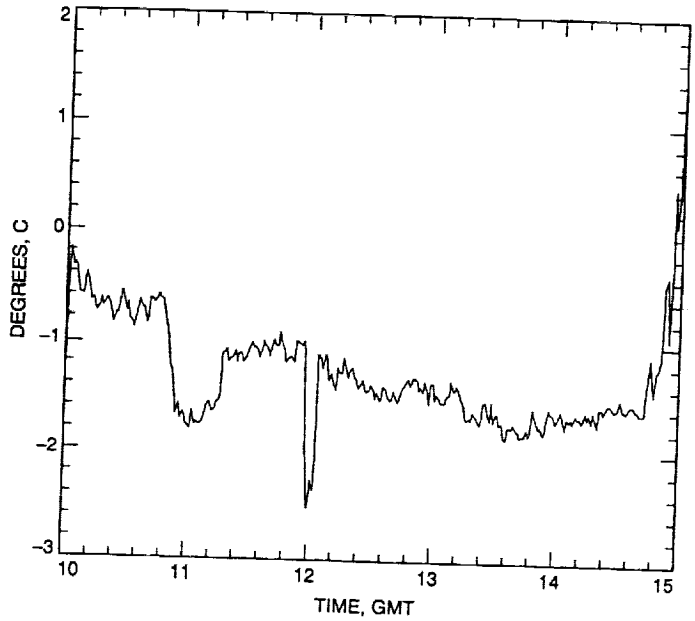


Fig. C-3. Ambient air temperature, Tower Number 1, day 344, December 9, 1992.

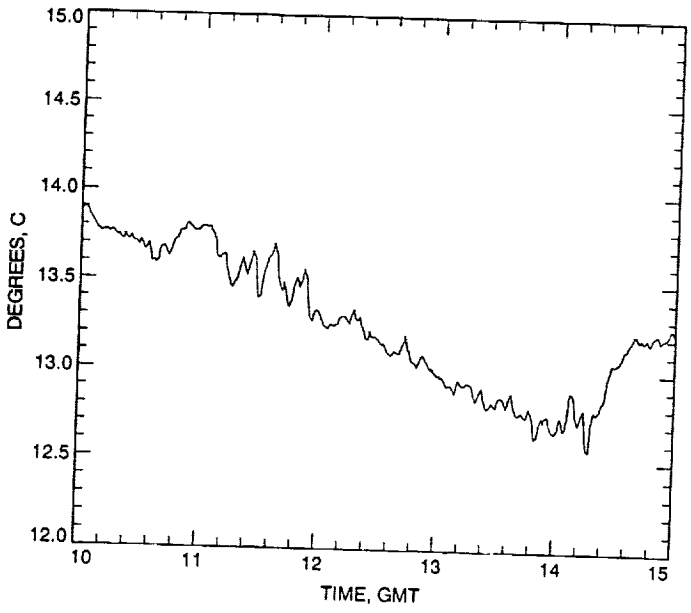


Fig. C-2. Optical path air temperature at the (source simulator) (see Fig. 2), day 344, December 9, 1992.

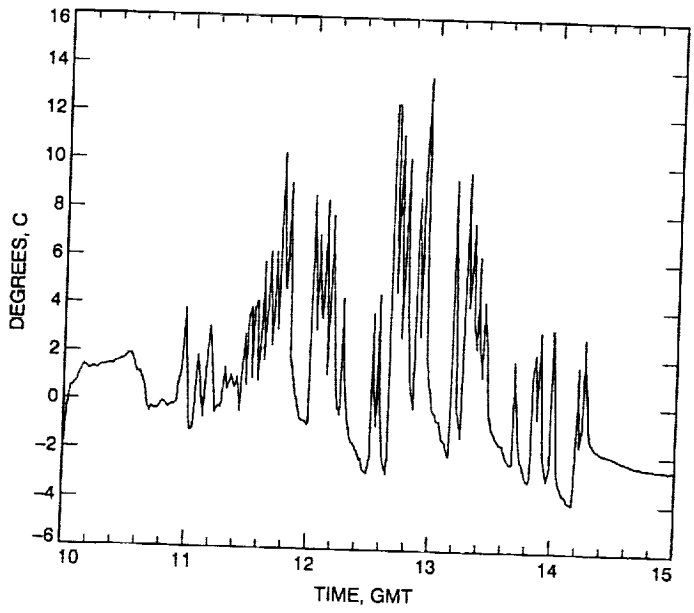


Fig. C-4. Secondary mirror temperature, day 344, December 9, 1992.

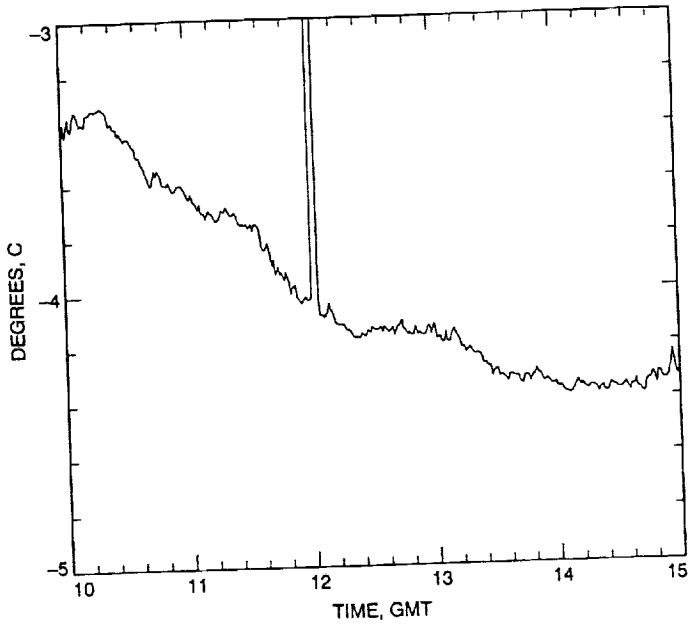


Fig. C-5. Ambient air dew point, Tower Number 2, day 344, December 9, 1992.

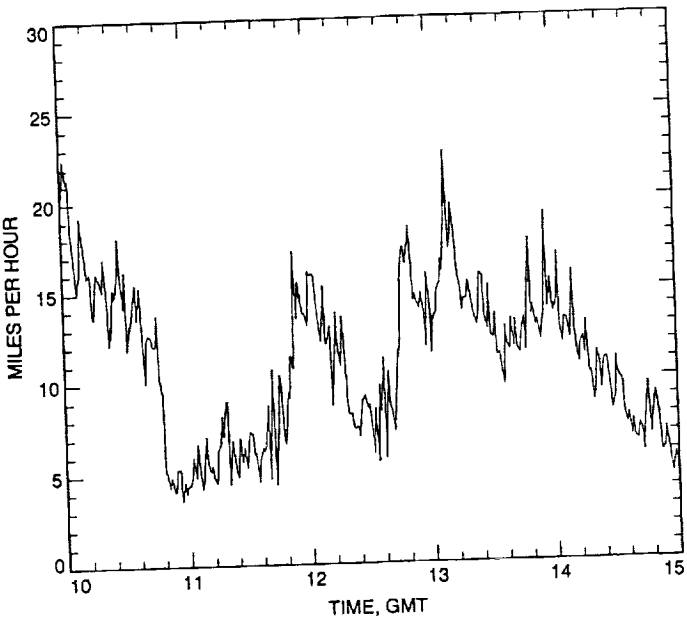


Fig. C-6. Tower Number 1 wind speed, top, day 344, December 9, 1992.

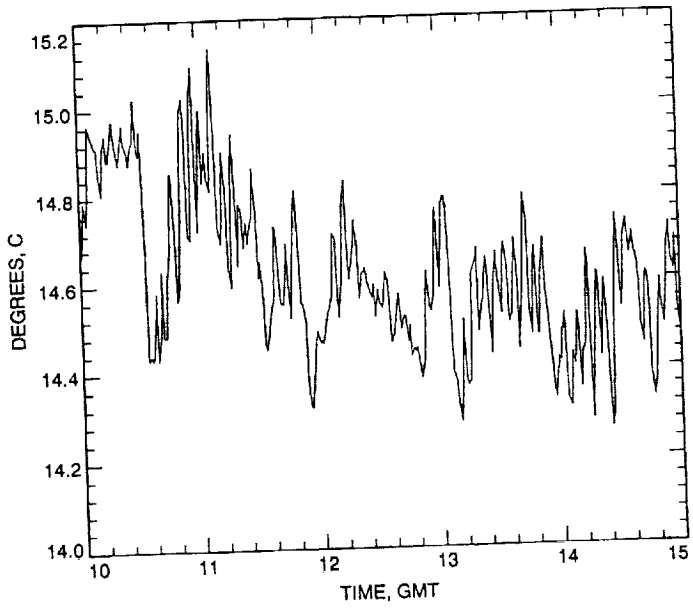


Fig. C-7. Optical path air temperature at M8 (see Fig. 2), day 344, December 9, 1992.

TDA Progress Report 42-114

August 15, 1993

Preliminary Design and Implementation of the Baseline Digital Baseband Architecture for Advanced Deep Space Transponders

T. M. Nguyen
Communications Systems Research Section

H.-G. Yeh
Spacecraft Telecommunications Equipment Section

This article investigates and identifies the baseline design and implementation of the digital baseband architecture for advanced deep space transponders. Trade studies on the selection of the number of bits for the analog-to-digital converter (ADC) and optimum sampling schemes are presented. In addition, the proposed optimum sampling scheme is analyzed in detail. Descriptions of possible implementations for the digital baseband (or digital front end) and digital phase-locked loop (DPLL) for carrier tracking are also described.

I. Introduction

Future NASA missions will require cheaper, smaller, and more energy-efficient spacecraft telecommunication equipment. These requirements motivated this study on advanced transponders for deep space applications. Recently, a study [1] has investigated various digital baseband architectures for future deep space transponders. Three different architectures were proposed for near-term, intermediate, and long-term solutions. The purpose of this article is to investigate and identify the baseline design and the conceptual implementation of the digital baseband architecture for a short-term solution.

The baseline architecture will use advanced digital technologies and signal-processing techniques for improved performance along with attractive functionality and adaptability to mission requirements. The identified architecture should also meet the interface constraint to minimize the cost of the design. The baseline architecture was developed based on the current configuration of the Cassini Deep Space Transponder (DST) [2]. The proposed architecture will maintain the analog IF section and the automatic gain control (AGC) loop at the first IF mixer identical to the current Cassini DST. However, the second IF will be redesigned to ease the digitization of baseband functions. In addition, the command detector unit (CDU) function, along with its modifications, will be included as a whole in the advanced DST. A description of the CDU and its modifications can be found in [1].

The simplified block diagram of the baseline architecture for the receiver of the DST is shown in Fig. 1. For this baseline architecture, the analog phase-locked loop

(APLL) for carrier tracking is replaced by a hybrid digital phase-locked loop (DPLL) and the ranging signal is extracted by filtering and turning the signal around without further signal processing (analog turnaround ranging). To simplify the hardware, the sampling frequency will be selected so that it is compatible with the sampling rate requirement of the CDU. A detailed description of the baseline architecture shown in Fig. 1 can also be found in [1].

This article begins with the trade studies for the selection of the number of bits for the ADC and the optimum sampling technique. Based on the selected optimum sampling scheme, the implementation of the baseband digital front end for simplified hardware was analyzed and proposed. Next, preliminary results related to the design and conceptual implementation of the DPLL for carrier tracking are presented and discussed. Finally, the article concludes with a summary of the salient features associated with this baseline design and direction of future work.

II. Selection of Number of Bits for the ADC

The number of bits required for the ADC at the second IF will determine the setting of the power for the AGC, the carrier signal-to-noise (SNR) degradation due to quantization, and the saturation noise. This section summarizes the results presented by Nguyen.[1] The carrier SNR degradation due to digitization, Δ, in the presence of Gaussian noise, is given by Nguyen[2] as

$$\Delta = \left(1 + \frac{P_N}{\sigma^2}\left(1 + \frac{P_S}{P_n}\right)\right)^{-1} \tag{1}$$

where

$$\frac{P_N}{\sigma^2} = 2\left(K^2+1\right)F(K) - \frac{2K}{\sqrt{2\pi}}e^{\left(-\frac{K^2}{2}\right)} + \frac{K^2}{12(M-1)^2}\left(1-2F(K)\right) \tag{2}$$

$$\frac{P_S}{P_n} = \left(\frac{P_S}{N_0}\right)\left(\frac{2}{F_S}\right) \tag{3}$$

$$K = \frac{1}{LF} \tag{4}$$

$$F(K) = \frac{1}{\sqrt{2\pi}}\int_K^\infty e^{\left(-\frac{u^2}{2}\right)}du \tag{5}$$

Note that (P_N/σ^2) denotes the quantization noise plus saturation noise-to-carrier signal power ratio; (P_S/P_n) denotes input carrier power-to-noise power ratio; N_0 is the one-sided input noise power density; F_S is the sampling frequency; $M = 2^{N-1}$, where N is the number of bits (including sign); and LF is the loading factor defined as follows:

$$LF = \frac{\text{rms amplitude of the total input signal}}{\text{ADC saturation voltage}} \tag{6}$$

The optimum values for K for various values of N have been calculated in [3], and the corresponding optimum LF as a function of N is depicted in Fig. 2. From the optimum values of N and LF found in Fig. 2, one can calculate the corresponding values of (P_N/σ^2) using Eq. (2). Using the calculated (P_N/σ^2) together with Eq. (1), one can calculate the carrier SNR degradation due to digitization. The results are plotted in Fig. 3 for 0 dB-Hz $< P_S/N_0 <$ 50 dB-Hz and 1 MHz $< F_S <$ 36 MHz. Note that for 0 dB-Hz $< P_S/N_0 <$ 50 dB-Hz and 1 MHz $< F_S <$ 36 MHz, one has: $1 + P_S/P_n \approx 1$. Using this approximation, the results are shown in Fig. 3.

Therefore, to achieve the digitization with a degradation in carrier SNR of less than 0.1 dB and to meet the required dynamic range of $6N$ dB for the input carrier signal, the required number of bits is $N \geq 4$ bits. In addition, the higher the number of bits that one selects, the less susceptible to interference the signal will be.[3] Consequently, the required number of bits for the ADC should be selected such that 4 bits $\leq N \leq$ 8 bits.

III. Optimum Sampling Scheme

A. Review of Current Sampling Techniques

Currently, there are several techniques for sampling the band-pass signals [4]: in-phase and quadrature (I&Q) baseband sampling with analog quadrature, I&Q sampling with analog Hilbert transform, band-pass sampling with

[1] T. M. Nguyen, "Selection of the A/D Sampling Frequency and Number of Bits for the Advanced Transponder," JPL Interoffice Memorandum 3313-92-024 (internal document), Jet Propulsion Laboratory, Pasadena, California, April 20, 1992, revised May 18, 1992.

[2] Ibid.

[3] J. Berner, "Number of Bits Required in Block-V ADC," JPL Interoffice Memorandum 3338-90-048 (internal document), Jet Propulsion Laboratory, Pasadena, California, March 26, 1990.

digital quadrature mixers, and band-pass sampling with digital Hilbert transform. Based upon the investigation in [4], the band-pass sampling technique with digital quadrature mixers is recommended (see Fig. 5) for the baseline design of the advanced transponder because of the following reasons:

(1) There are no phase and amplitude imbalances because the mixing is done in digital domain.

(2) The digital low-pass filter (LPF) using a finite-impulse response (FIR) filter provides constant group delay that is very important for ranging and Doppler information.

(3) Only one ADC is required.

(4) If the sampling period is exactly $1/4F_{IF}$, then the reference of I&Q components reduces to an alternating sequence.

It will be shown later that the hardware implementation can be simplified by using the last property with some modification. Before describing how to implement the "digital front end" for the advanced transponder, the digital front end needs to be defined.

The digital front end of the transponder (see Figs. 1 and 4) is designed to accept an IF analog signal and output digital baseband I&Q components for further processing by the remainder of transponder. The purpose of the digital front end is to provide the transponder with a demodulation capability from an IF-to-baseband digital signal.

B. Conceptual Implementation of the Digital Front End

To implement the digital front end (see Fig. 4), one must set up the criteria for selecting both the optimum sampling frequency and the analog IF. First, there are several criteria for selecting the optimum sampling frequency, namely,

(1) The hardware implementation should be simple.

(2) The sampling frequency should be sufficiently high to meet the required number of samples per symbol for the CDU and the carrier tracking loop.

(3) The sampling frequency should be sufficiently high to prevent aliasing of the baseband signal with the images that occur at the sampling rate.

(4) The sampling frequency selected should meet the current specification of the analog-to-digital (A/D) technology with reasonable cost.

Based on these criteria, the sampling frequency, F_S, must be selected to satisfy the following conditions [5–7]:

$$F_S \geq 2BW \tag{7}$$

$$\frac{2}{(l+1)}\left(F_{IF}-\frac{BW}{2}\right) \leq F_S \leq \frac{2}{(l+1)}\left(F_{IF}+\frac{BW}{2}\right) \tag{8}$$

where BW is the bandwidth of the band-pass signal in hertz, F_{IF} is the center of the IF band, and l is a positive integer. In order to simplify the hardware implementation, one chooses equality in Eq. (7) and an odd integer for l in Eq. (8):

$$F_S = 2pBW \tag{9}$$

$$nF_S = F_{IF} - \frac{BW}{2} \tag{10}$$

where $p \geq 1$ and $n = (l+1)/2$.

Solving for the sampling frequency in terms of F_{IF}, one gets

$$F_S = \frac{4pF_{IF}}{(1+4np)} \tag{11}$$

for $p = 1$, Eq. (11) reduces to

$$F_S = \frac{4F_{IF}}{(1+4n)} \tag{12}$$

where n satisfies the following inequality

$$n \leq \left[\frac{F_{IF}}{2BW} - \frac{1}{4}\right] \tag{13}$$

where $[x]$ is the smallest integer that is less than or equal to x. It should be mentioned that the sampling scheme proposed is known as the under-sampling scheme. Note that, in practice, to simplify the I&Q sampling technique using digital quadrature mixers (see Fig. 5) to the configuration shown in Fig. 7(a), the sampling frequency must

be chosen as $4F_{IF}$ [8–10].[4] However, using Eq. (12) one can avoid selecting a high sampling frequency (and hence achieve a more energy efficient spacecraft). As an example for the proposed under-sampling scheme, let the IF be 5 MHz and the bandwidth of the signal be 36 kHz; then using Eq. (13), one obtains $n \leq 69$. If one selects $n = 6$, then the sampling frequency required for this case is, from Eq. (12), $F_S = 0.8$ MHz. Figures 7(b) and (c) illustrate and compare the two sampling schemes discussed above.

Secondly, one must select the IF so that the analog circuitry in the transponder can be designed and built easily. There are several criteria for selecting the IF, namely,

(1) The quadrature sampling error caused by spectral bands overlapping [7] must be avoided by selecting the upper cutoff frequency of the band-pass filter (BPF) equal to an integer multiple of the bandwidth, i.e.,

$$F_{IF} + \frac{BW}{2} = cBW \tag{14}$$

where c is a positive integer.

(2) For minimum hardware implementation, F_{IF} and F_S should satisfy Eq. (12).

(3) F_{IF} must be chosen such that the associated band-pass filter in the analog-mixing and filtering circuitry is realizable. The passband of this filter must pass the required number of sidelobes of the command signal and possibly the highest ranging clock frequency.

(4) F_{IF} must be chosen by taking into consideration the throughput limitation of the digital filters of the digital front end.

(5) F_{IF} must be chosen to provide minimum carrier delay variation.

To show that Eq. (12) can be used to select the sampling frequency for hardware simplification, look at the mathematical model for the uplink signal, $S(t)$:

$$S(t) = \sqrt{2P}\sin\left((\omega + \omega_d)t + \Theta(t) + \varphi\right) \tag{15}$$

where

P = total received power

ω = angular carrier frequency

ω_d = Doppler angular frequency offset

$\Theta(t)$ = phase modulation $= md(t)\sin(\omega_{SC}t) + m_R R(t)$

φ = phase offset

m = command modulation index

$d(t)$ = command non-return-to-zero (NRZ) data

ω_{SC} = command subcarrier frequency

m_R = ranging modulation index

$R(t)$ = ranging signal

Without loss of generality, one can set $\omega = \omega_{IF} = 2\pi F_{IF}, \omega_d = 0, \varphi = 0$, and can expand Eq. (15) to get

$$S(t) = \sqrt{2P}\left[\cos(\Theta(t))\sin(2\pi F_{IF}t) + \sin(\Theta(t))\cos(2\pi F_{IF}t)\right] \tag{16}$$

The first term in Eq. (16) represents the carrier component, and the second is the command signal component.

Assume that the I&Q components of Eq. (15) are extracted by using analog quadrature mixers as shown in Fig. 6. If the cutoff frequency of the LPF is such that it rejects higher-order harmonics components and passes only the first harmonic component without distortion, then the output I&Q components are

$$I(t) = \frac{1}{2}\sin(\Theta(t)) \tag{17}$$

$$Q(t) = \frac{1}{2}\cos(\Theta(t)) \tag{18}$$

Note that Eqs. (17) and (18) have been normalized by $\sqrt{2P}$. If one assumes that the ranging signal is off, then the $I(t)$ and $Q(t)$ shown in Eqs. (17) and (18) become

$$I(t) = \frac{1}{2}d(t)\sin(m\sin(\omega_{SC}t)) \approx d(t)J_1(m)\sin(\omega_{SC}t) \tag{19}$$

$$Q(t) = \frac{1}{2}\cos(m\sin(\omega_{SC}t)) \approx \frac{1}{2}J_0(m) \tag{20}$$

[4] M. J. Agan and C. R. Pasqualino, "AMT Modem Digital Front End," JPL Interoffice Memorandum AMT:331-5-90-0 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 1990.

Note that the I-component contains the command information and the Q-component contains the amplitude of the carrier component.

A sample of the signal expressed in Eq. (16) is obtained by using the sampling frequency derived in Eq. (12). At this sampling frequency, one has

$$t = kT_S = \frac{k(4n+1)}{4F_{IF}},\ k = 0, 1, 2, 3, 4, \ \ldots \tag{21}$$

where T_S denotes the sampling period. Substituting Eq. (21) into Eq. (16) and evaluating it for $k = 0, 1, 2, 3, 4, \ldots$, one sees that Eq. (16) generates the following sequence:

$$\sqrt{2P}\sin(\Theta(0)), \sqrt{2P}\cos(\Theta(T_S)), -\sqrt{2P}\sin(\Theta(2T_S)),$$

$$-\sqrt{2P}\cos(\Theta(3T_S)), \sqrt{2P}\sin(\Theta(4T_S)),$$

$$\sqrt{2P}\cos(\Theta(5T_S)), -\sqrt{2P}\sin(\Theta(6T_S)),$$

$$-\sqrt{2P}\cos(\Theta(7T_S)), \sqrt{2P}\sin(\Theta(8T_S)),$$

$$\sqrt{2P}\cos(\Theta(9T_S)), \ \ldots$$

Taking the above sequence and multiplying by the $\{1, 1, -1, -1, 1, 1, -1, -1, 1, 1, \ldots\}$ sequence, one gets

$$\sqrt{2P}\sin(\Theta(0)), \sqrt{2P}\cos(\Theta(T_S)), \sqrt{2P}\sin(\Theta(2T_S)),$$

$$\sqrt{2P}\cos(\Theta(3T_S)), \sqrt{2P}\sin(\Theta(4T_S)),$$

$$\sqrt{2P}\cos(\Theta(5T_S)), \sqrt{2P}\sin(\Theta(6T_S)),$$

$$\sqrt{2P}\cos(\Theta(7T_S)), \sqrt{2P}\sin(\Theta(8T_S)),$$

$$\sqrt{2P}\cos(\Theta(9T_S)), \ \ldots$$

Note that this sequence alternates between samples of $I(t)$ and $Q(t)$ shown in Eqs. (18) and (19) with only a scaling factor difference. The above sequence can be simply expressed as

$$I(0), Q(T_S), I(2T_S), Q(3T_S), I(4T_S), Q(5T_S),$$

$$I(6T_S), Q(7T_S), I(8T_S), Q(9T_S), \ \ldots$$

Based on these results, the optimum implementation of the digital front end for the baseline design of the advanced transponder is shown in Fig. 7. As shown in this section, using the sampling frequency derived in Eq. (12), one can simplify the hardware. The hardware simplification is exactly the same as for the case when the sampling frequency is $4F_{IF}$ [9][5] except when using a lower F_S, and hence lower power consumption.

IV. Design and Implementation of the Carrier Tracking Loop

A. Description of the Carrier APLL and Transformation Techniques

The block diagram of the analog carrier tracking loop for the Cassini DST is depicted in Fig. 8. Based on the current design, an architecture of the DPLL for computer simulation is developed. Presently, the analog carrier tracking loop is Type I, second order PLL with the following characteristics:

$$AK = \text{loop gain} = 2.4 \times 10^7 \tag{22}$$

$$B(S) = \frac{1}{(1+\tau_{RC}S)},\ \tau_{RC} = 1.6 \times 10^{-5} \tag{23}$$

$$F(S) = \frac{1+\tau_2 S}{1+\tau_1 S},\ \tau_1 = 4707, \tau_2 = 0.0442 \tag{24}$$

$$V(S) = \frac{1}{(1+\tau_V S)},\ \tau_V = 1.0 \times 10^{-6} \tag{25}$$

$$K(S) = \frac{1}{S} \tag{26}$$

Note that $B(S)$ is the typical LPF, $F(S)$ is the loop filter, $V(S)$ is the rolloff filter of the voltage control oscillator (VCO), and $K(S)$ is the VCO integrator.

Let $G(S)$ be the transfer function of the analog loop defined as follows:

[5] Ibid.

$$G(S) = B(S)F(S)V(S) \tag{27}$$

From the analog characteristics of the loop, there are four different techniques (cases 1–4) to design the equivalent digital loop

(1) Bilinear Transformation Method. This method preserves the phase characteristics in the narrow passband when mapping the APLL into the digital domain with high sampling frequency. The mapping from analog (S-domain) to digital domain (Z-domain) can be achieved by direct substitution of the following equation into the analog transfer function [11–13]:

$$S = \frac{2}{T_S}\frac{(Z-1)}{(Z+1)} \tag{28}$$

(2) Impulse-Invariant Transformation Method. This mapping technique preserves the impulse response at the sampling points. The relationship between the S-variable and Z-variable is given by [12,13]

$$S = \frac{(Z-1)}{T_S Z} \tag{29}$$

However, the corresponding digital transfer function cannot be obtained by substituting Eq. (28) directly into the analog transfer function as in case (1) above. Let $g(t)$ be the impulse response of $G(S)$, i.e., $g(t) = L^{-1}\{G(S)\}$, where $L^{-1}\{.\}$ denotes the inverse Laplace transform of $\{.\}$. Thus, the digital approximation of the analog transfer function $G(S)$ is given by

$$G_D(Z) = T_S\left(z\{g(t)|_{t=nT_s}\}\right) \tag{30}$$

where $z\{.\}$ is the z-transform of $\{.\}$. Note that the analog transfer function $G(S)$ considered in this article is defined in Eq. (27).

(3) Step-Invariant Transformation Method. This method preserves the step response at the sampling points when mapping S-domain to Z-domain. The relationship between S- and Z-variables is given in [12]:

$$S = \frac{(Z-1)}{T_S} \tag{31}$$

Similar to case (2), the equivalent digital transfer function of the open loop cannot be found by substituting Eq. (31) directly into the analog transfer function. The relationship between the analog and digital transfer function is [12,13]

$$G_D(Z) = \frac{Z-1}{Z}\left(z\left\{\left\lfloor L^{-1}\left[\frac{G(S)}{S}\right]\right\rfloor_{t=nT_S}\right\}\right) \tag{32}$$

where $z\{.\}$ and $G(S)$ are defined the same as above.

(4) Rational Transformation Method. This mapping technique is identical to the impulse-invariant transformation technique [11,12].

B. Recursive Implementation of the Carrier DPLL

To obtain the digital approximation of the carrier APLL described in Section IV.A, each functional block in the analog loop, i.e., $B(S)$, $F(S)$, $V(S)$, and $K(S)$, can be mapped directly into the Z-domain using bilinear transformation or the composite function $B(S)F(S)V(S)$ using impulse-invariant (or step-invariant) transformation. These mappings are accomplished by applying Eqs. (28), (30), and (32), depending on the type of transformation used. Following are the recursive implementations of the digital transfer functions.

1. Recursive Implementation of $B(S)$, $F(S)$, $V(S)$, and $K(S)$ Using Bilinear Transformation. To obtain the digital approximation of the analog loop using bilinear transformation, one substitutes Eq. (28) into Eqs. (23)–(26) to get the Z-domain representations for the LPF $B(S)$, loop filter $F(S)$, VCO rolloff filter $V(S)$, and the integrator $K(S)$. The results are

$$B(Z) = \frac{(1+Z^{-1})}{(A_{00}Z^{-1}+A_{11})} \tag{33}$$

$$F(Z) = \frac{(A_0 Z - B_0)}{(A_1 Z - B_1)} \tag{34}$$

$$K(Z) = \frac{T_S(Z+1)}{2(Z-1)} \tag{35}$$

where

$$A_{00} = 1 - C_0,\ A_{11} = 1 + C_0 \tag{36}$$

$$A_0 = 1 + a_o,\ A_1 = 1 + b_0,\ B_0 = a_0 - 1,\ B_1 = b_0 - 1 \tag{37}$$

and

$$C_0 = \frac{2\tau_{RC}}{T_S},\ a_0 = \frac{2\tau_2}{T_S},\ b_0 = \frac{2\tau_1}{T_S} \tag{38}$$

Note that the Z-domain representation for $V(S)$ is exactly the same as Eq. (33), except that C_0 is replaced by

$$C_0 = \frac{2\tau_V}{T_S} \tag{39}$$

The digital closed-loop transfer function, $H(Z)$, for this case is given by

$$H(Z) = \frac{AK(B(Z)F(Z)V(Z)K(Z))}{[1 + AK(B(Z)F(Z)V(Z)K(Z))]} \tag{40}$$

Plots of the analog and digital closed-loop phase and magnitude responses are shown in Figs. 9(a) and 9(b). These figures show that for sampling frequencies below 100 kHz, distortions in phase and magnitude can occur for the digital approximation loop. In addition, the figures show that for sampling frequencies greater than or equal to 100 kHz the response of the digital loop approaches that of the analog counterpart. Hence, the minimum sampling frequency for this case is 100 kHz. Figures 10(a), (b), and (c) show the recursive implementation of the LPF $B(Z)$, integrator $K(Z)$, and loop filter $F(Z)$.

The implementation of the rolloff filter $V(Z)$ is similar to that of the LPF $B(Z)$.

2. Recursive Implementation of the Analog Transfer Function $G(S)$ and $K(S)$ Using Impulse-Invariant Transformation. To obtain the equivalent digital approximation for the integrator $K(S)$, one substitutes Eq. (29) into Eq. (26) to get

$$K(Z) = \frac{ZT_S}{(Z-1)} \tag{41}$$

The digital approximation for the analog transfer function $G(S)$ (see Eq. (27)) is obtained by finding the inverse Laplace transform of $G(S)$ and then substituting the result into Eq. (30). Evaluating Eq. (30), one has

$$G_D(Z) = T_S \left[\frac{\alpha_0}{1 - Z^{-1}e^{-aT_S}} + \frac{\alpha_1}{1 - Z^{-1}e^{-bT_S}} + \frac{\alpha_2}{1 - Z^{-1}e^{-cT_S}} \right] \tag{42}$$

where

$$\alpha_0 = \frac{\tau_1 - \tau_2}{(\tau_1 - \tau_{RC})(\tau_1 - \tau_V)} \tag{43}$$

$$\alpha_1 = \frac{\tau_{RC} - \tau_2}{(\tau_{RC} - \tau_1)(\tau_{RC} - \tau_V)} \tag{44}$$

$$\alpha_2 = \frac{\tau_V - \tau_2}{(\tau_2 - \tau_1)(\tau_V - \tau_{RC})} \tag{45}$$

and

$$a = \frac{1}{\tau_1},\ b = \frac{1}{\tau_{RC}},\ c = \frac{1}{\tau_V} \tag{46}$$

The digital closed-loop transfer function for this case is given by

$$H(Z) = \frac{AK(G_D(Z)K(Z))}{[1 + AK(G_D(Z)K(Z))]} \tag{47}$$

From Eq. (47), the plots of the phase and magnitude responses can be obtained for the digital approximation loop. Figures 11(a) and 11(b) illustrate the closed-loop phase and magnitude responses for both analog and digital loops. The figures show that the response of the digital loop approximated by using impulse-invariant transformation is the same as the analog loop when the sampling frequency is higher than or equal to 100 kHz. When the sampling frequency is less than 100 kHz, the digital loop can encounter serious distortion in both phase and amplitude. The recursive implementations $G_D(Z)$ and $K(Z)$ using impulse-invariant transformation are shown in Figs. 12(a) and 12(b).

3. Recursive Implementation of the Analog Transfer Function $G(S)$ and $K(S)$ Using Step-Invariant Transformation. Digital approximations $K(Z)$ and $G_D(Z)$ for $K(S)$ and $G(S)$ using step invariant transformation can be obtained by using Eqs. (31) and (32). The results are

$$K(Z) = \frac{T_S}{(Z-1)} \tag{48}$$

$$G_D(Z) = \beta_0 + \beta_1 \left[\frac{1 - Z^{-1}}{1 - Z^{-1}e^{-aT_S}}\right] + \beta_2 \left[\frac{1 - Z^{-1}}{1 - Z^{-1}e^{-bT_S}}\right] + \beta_3 \left[\frac{1 - Z^{-1}}{1 - Z^{-1}e^{-cT_S}}\right] \tag{49}$$

where

$$\beta_0 = \frac{\alpha_0}{a} + \frac{\alpha_1}{b} + \frac{\alpha_2}{c} \tag{50}$$

$$\beta_1 = -\frac{\alpha_0}{a},\ \beta_2 = -\frac{\alpha_1}{b},\ \beta_3 = -\frac{\alpha_2}{c} \tag{51}$$

The parameters α_0, α_1, a, b, and c are defined in Eqs. (43)–(46), respectively. Again, Eq. (47) can be used to evaluate the closed-loop transfer function for this case. The plots of the closed-loop transfer functions for both analog and digital loops are shown in Figs. 13(a) and 13(b). The figures show that the magnitude response approaches the analog response when the sampling frequency is higher than or equal to 100 kHz. However, the phase response suffers serious distortion when the sampling frequency is less than 1 MHz. Thus, in order to achieve the same response as the analog loop, the digital approximation loop using step-invariant transformation must be sampled at least at 1 MHz, i.e., this method requires a 10 times higher sampling frequency than the previous methods. The recursive implementations of $G_D(Z)$ and $K(Z)$ using step-invariant transformation are shown in Figs. 14(a) and 14(b).

V. Conclusions and Future Work

This article presented preliminary results on the design and implementation of the baseline digital baseband architecture for future deep space transponders, and also presented trade studies on: (1) the number of bits required by the ADC, (2) the sampling and IF for hardware simplification, and (3) the optimum sampling technique. A conceptual implementation of the proposed optimum sampling technique was presented and discussed. In addition, the phase and amplitude responses of digital approximations of the analog loop were briefly investigated. It was found that in order to achieve the same closed-loop responses as the analog counterpart, the step-invariant transformation method requires a higher sampling frequency than the other methods.

Furthermore, this article identified various architectures for possible implementation of the digital carrier tracking loop. The architecture that was found to provide the smallest phase jitter and fastest response is appropriate for the advanced transponder.

Acknowledgments

The authors thank S. Kayalar for his useful comments and suggestions, C. Kyriacou and A. Kermode for their constant support, and S. Hinedi for many useful discussions on the implementation and design of the digital phase-locked loop.

References

[1] T. M. Nguyen, S. Kayalar, H.-G. Yeh, and C. Kyriacou, "Advanced Transponders for Deep Space Applications," *Proceedings of IEEE Conference on Aerospace Applications*, Steamboat Springs, Colorado, January 30–February 3, 1993.

[2] N. Mysoor, J. Perret, and A. Kermode, "An X-Band Spacecraft Transponder for Deep Space Applications—Design Concepts and Breadboard Performance," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, no. 6, pp. 1192–1198, June 1992.

[3] G. A. Gray and G. W. Zeoli, "Quantization and Saturation Noise Due to Analog-to-Digital Conversion," *IEEE Transactions on Aerospace and Electronics Systems*, vol. 7, no. 1, pp. 222–223, January 1971.

[4] R. Sadr and M. Shahshahani, "On Sampling Band-Pass Signal," *The Telecommunications and Data Acquisition Progress Report 42-96, vol. October–December 1988*, Jet Propulsion Laboratory, Pasadena, California, pp. 14–20, February 15, 1989.

[5] A. Kohlenberg, "Exact Interpolation of Bandlimited Functions," *J. Appl. Phys.*, vol. 24, no. 12, p. 1432, December 1953.

[6] J. L. Brown, Jr., "On Quadrature Sampling of Bandpass Signals," *IEEE Trans, Aer. Elec. Sys.*, vol. AES-15, no. 3, pp. 366–371, May 1979.

[7] C. E. Persons, "Quadrature Sampling Error Formula," *J. Acoust. Soc. Am.*, vol. 57, no. 2, p. 511, February 1975.

[8] T. Hack, "I/Q Sampling Yields Flexible Demodulators," *RF Design*, vol. 14, no. 4, pp. 40–43, April 1991.

[9] R. Cali and G. Ferrari, "Algorithms for Computing Phase and AGC in Digital PLL Receivers," *RF Design*, vol. 15, no. 11, pp. 33–40, October 1992.

[10] R. Sfeir, S. Aguirre, and W. J. Hurd, "Coherent Digital Demodulation of a Residual Carrier Signal Using IF Sampling," *The Telecommunications and Data Acquisition Progress Report 42-78, vol. April-June 1984*, Jet Propulsion Laboratory, Pasadena, California, pp. 135–142, August 15, 1984.

[11] L. R. Rabiner and B. Gold, *Theory and Applications of Digital Signal Processing*, Englewood Cliffs, New Jersey: Prentice-Hall, 1975.

[12] W. D. Stanley, *Digital Signal Processing*, Reston, Virginia: Reston Publishing Co., 1975.

[13] K. Ogata, *Discrete Time Control Systems*, Englewood Cliffs, New Jersey: Prentice-Hall, 1987.

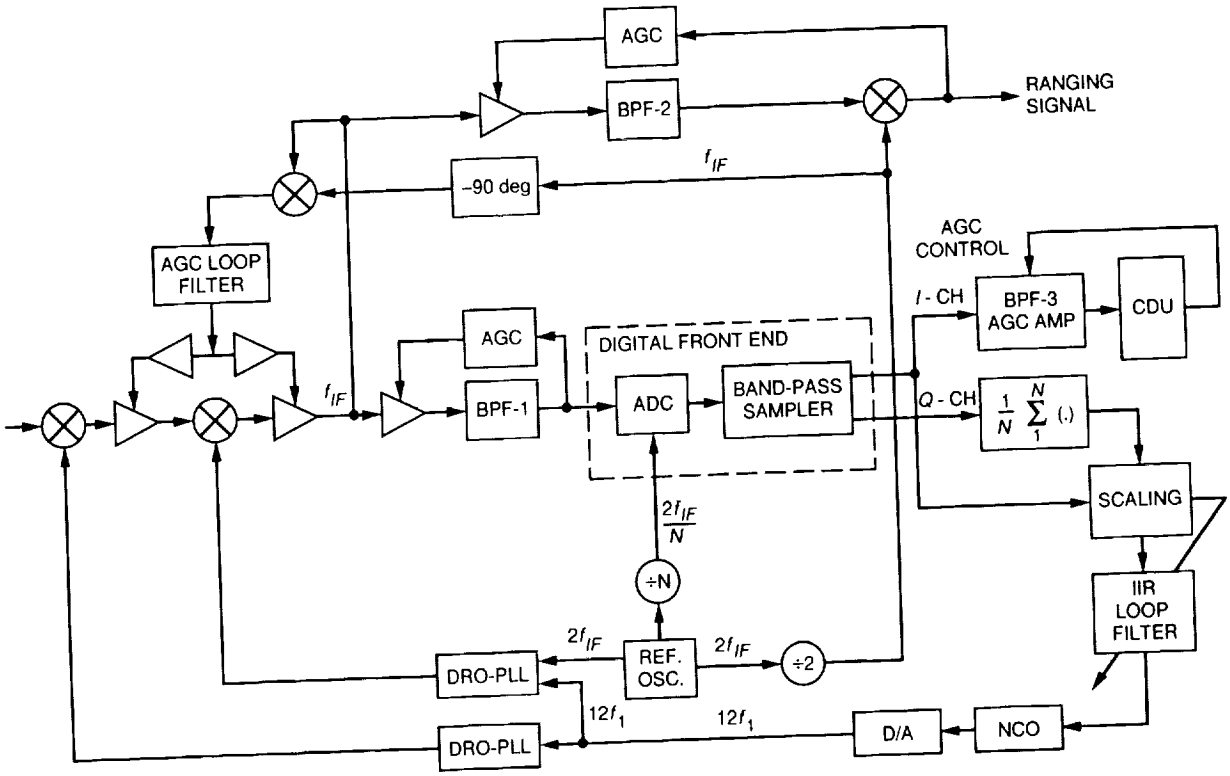


Fig. 1. Proposed baseline architecture for the advanced deep space transponder—a basic digital structure with analog turnaround ranging.

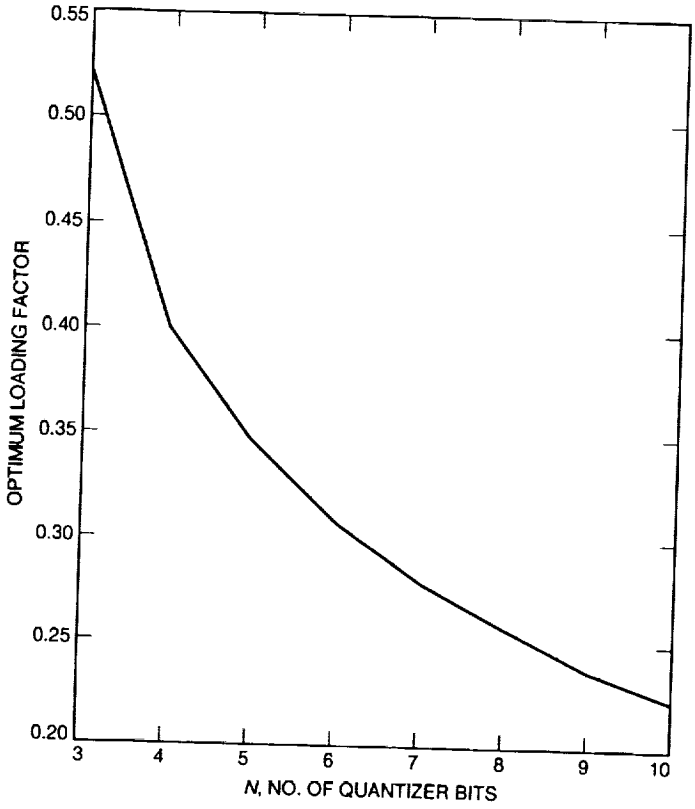


Fig. 2. Optimum loading factor as a function of *N*.

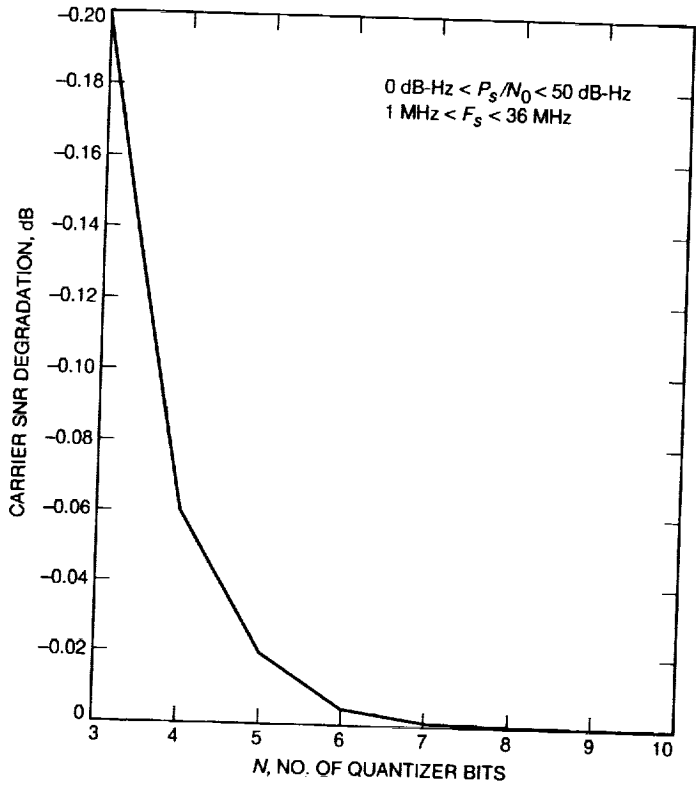


Fig. 3. Carrier SNR degradation as a function of *N*.

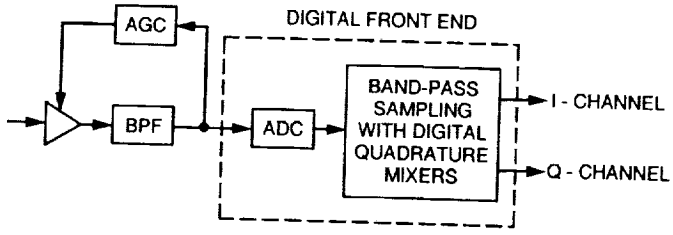


Fig. 4. A simplified block diagram of the proposed configuration for the digital front end of the transponder.

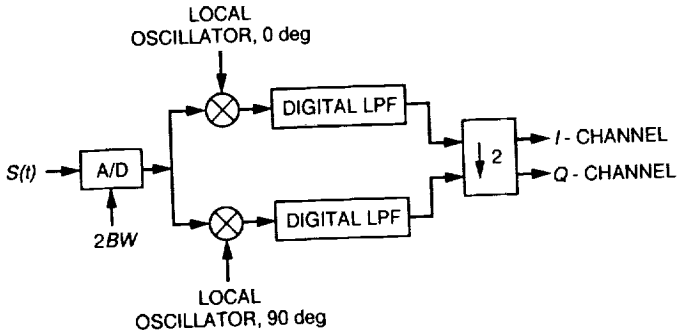


Fig. 5. I&Q sampling technique with digital quadrature mixers.

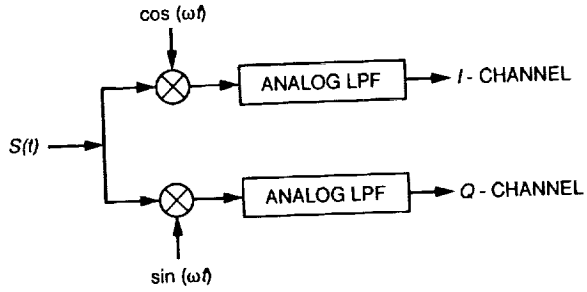
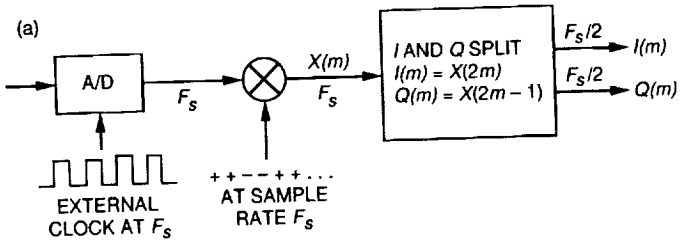


Fig. 6. I&Q extraction using the analog mixer approach.



$$F_S = \frac{4F_{IF}}{(4n+1)}, \; n = 0,1,2,3, \ldots$$

$$n \leq \left[\frac{F_{IF}}{2BW} - \frac{1}{4}\right]$$

$[X]$ = SMALLEST INTEGER $\leq X$

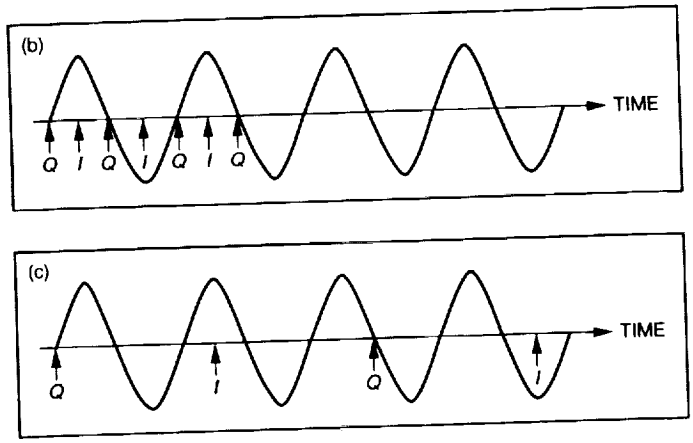


Fig. 7. Implementation and illustrations of the I&Q sampling technique: (a) proposed implementation using digital quadrature mixers; (b) illustration of over-sampling scheme, $F_C = 4F_{IF}$; and (c) illustration of under-sampling scheme, $F_S = \frac{4F_{IF}}{(4n+1)}$, $n = 6$.

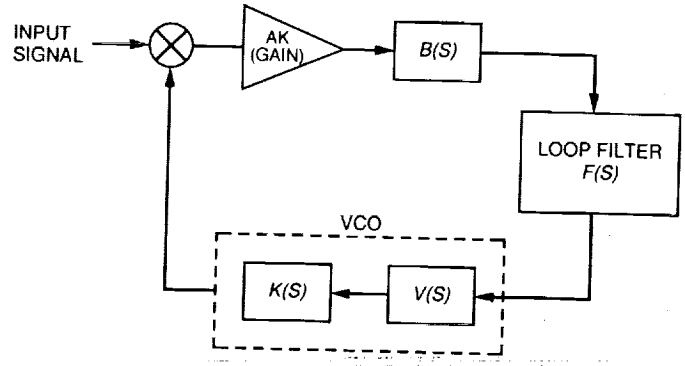


Fig. 8. Analog model of the Cassini carrier-tracking phase-locked loop—a short-loop version.

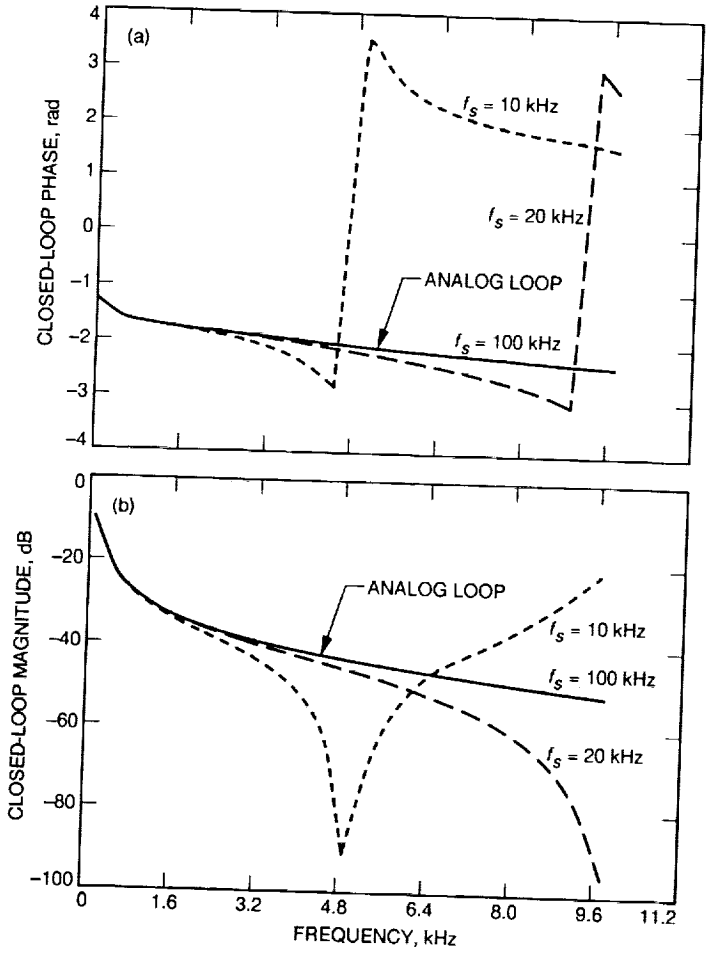


Fig. 9. Plots of two closed-loop responses for the digital loop using the bilinear transformation method: (a) phase response and (b) magnitude response.

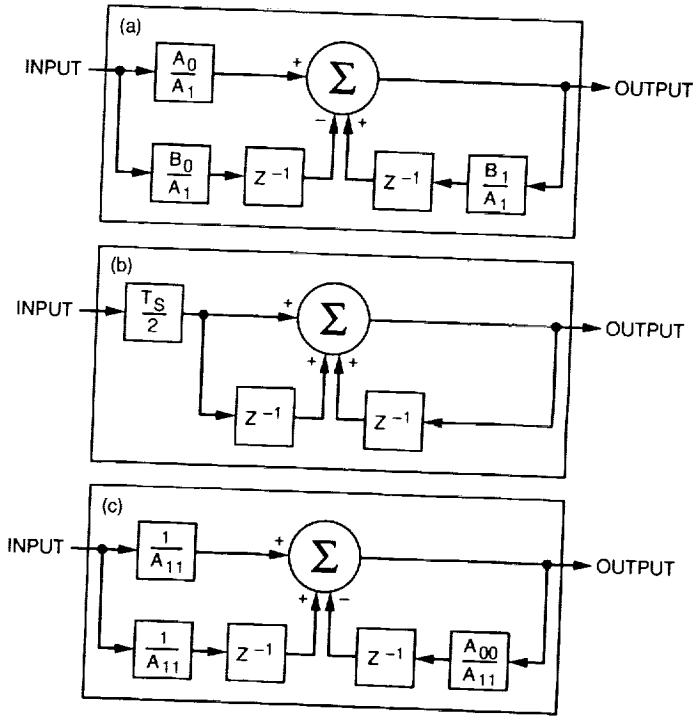


Fig. 10. Recursive implementation of the digital loop using the bilinear transformation method: (a) low-pass filter $B(Z)$, (b) integrator $K(Z)$, and (c) loop filter $F(Z)$.

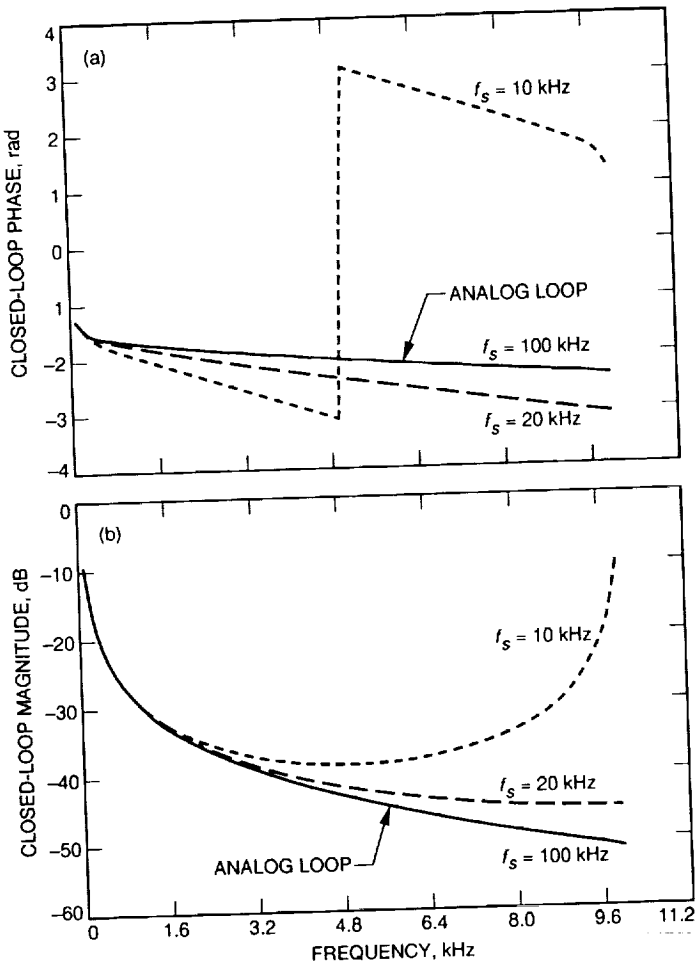


Fig. 11. Plots of two closed-loop responses for the digital loop using the impulse-invariant transformation method: (a) phase response and (b) magnitude response.

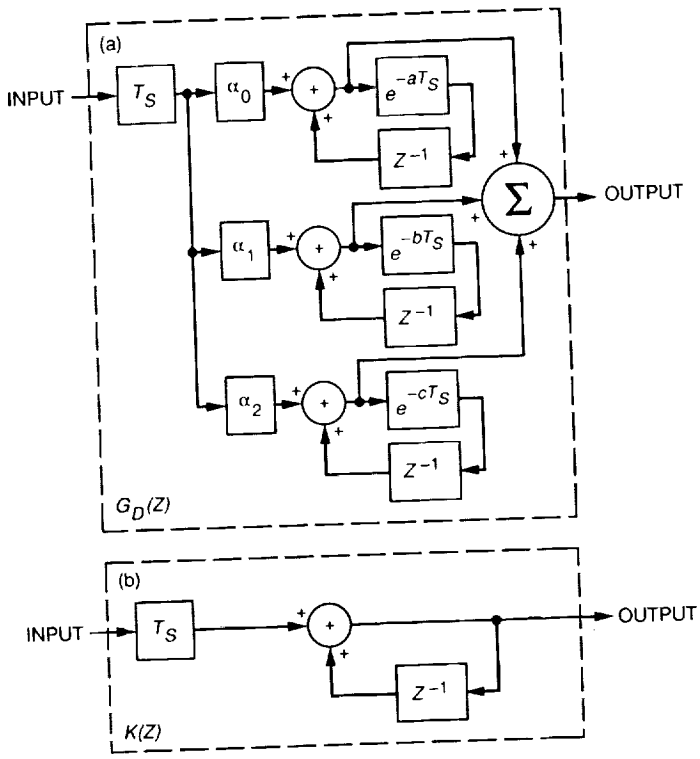


Fig. 12. Recursive implementation of the digital loop using the impulse-invariant transformation method: (a) open-loop transfer function $G_D(Z)$ and (b) integrator $K(Z)$.

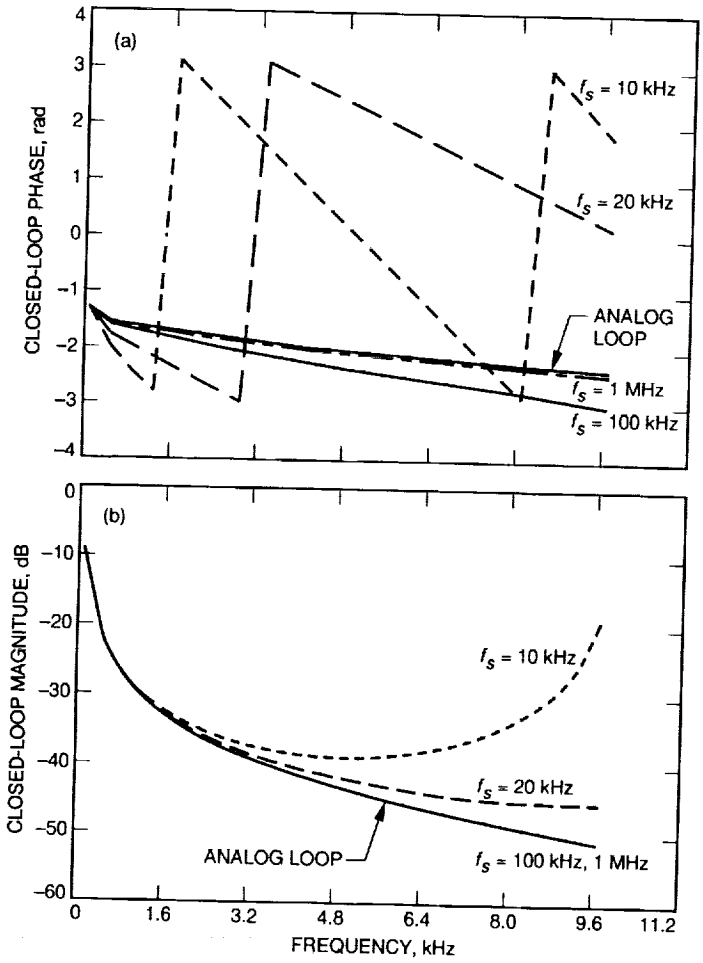


Fig. 13. Plots of two closed-loop responses for the digital loop using the step-invariant transformation method: (a) phase response and (b) magnitude response.

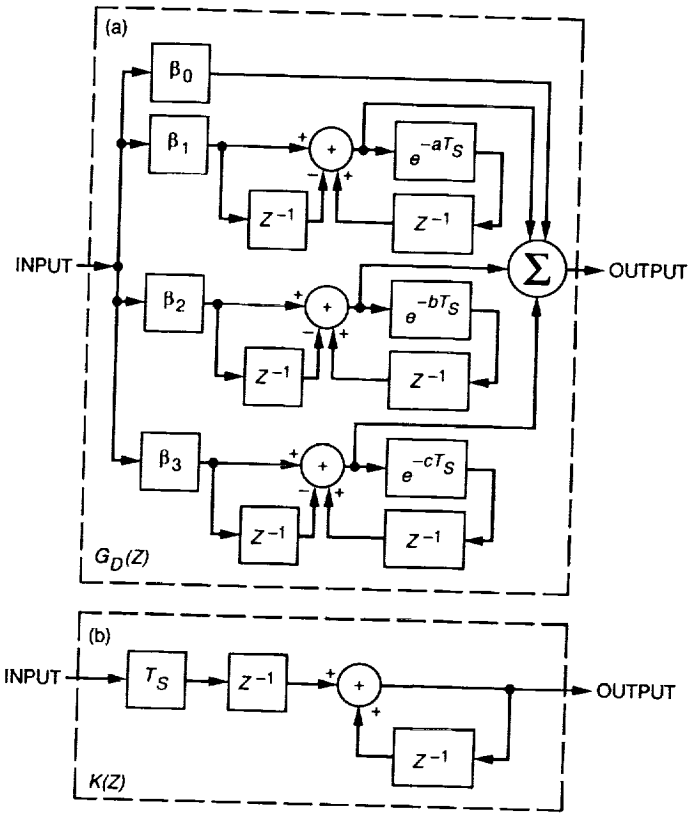


Fig. 14. Recursive implementation of the digital loop using the step-invariant transformation method: (a) open-loop transfer function $G_D(Z)$ and (b) integrator $K(Z)$.

August 15, 1993

TDA Progress Report 42-114

Hardware Design for the Autonomous Visibility Monitoring (AVM) Observatory

K. Cowles
Communications Systems Research Section

The hardware for the three Autonomous Visibility Monitoring (AVM) observatories has been redesigned. Changes in hardware design include electronics components, weather sensors, and the telescope drive system. Operation of the new hardware is discussed, as well as some of its features. The redesign will allow reliable automated operation.

I. Introduction

Extensive testing of original hardware for the engineering model of the Autonomous Visibility Monitoring (AVM) observatories has provided new specifications for the field hardware. The revised task allows improved capabilities and better specification of procured and manufactured components of the AVM hardware. Some of the original equipment has been retained. This article will outline the specifications for the hardware and describe the operation of the new system.

II. Objectives and Specifications

Several problem areas were identified with the hardware used in the engineering model. Based on experience operating the equipment, new specifications were developed for the observatory system. The objective of the AVM task is to develop an atmospheric transmission model for optical communications through autonomous visibility measurements of known stellar objects from several locations in the southwestern United States, and to update this model on a quarterly basis. In order to meet this objective, autonomous observatories will be deployed to continuously monitor star transmission intensities. Atmospheric attenuation probabilities will be determined from these measurements.

To achieve this objective, the system must observe weather conditions and sky transparency for several years. Weather sensors must be able to withstand the environmental extremes of remote sites, and the system must respond to indications from these sensors in case of inclement weather. Sky transparency sensors must be able to discern cloud/no cloud conditions and measure atmospheric attenuation. Attenuation from sensor components must be calibrated to ensure accurate calculation of attenuation from the atmosphere only. The system must be designed to measure atmospheric attenuation to within ten percent error.

The sensor chosen for the sky transparency measurements is a charge-coupled device (CCD) camera mounted to a ten-inch Schmidt–Cassegrain telescope. To provide

data on the spatial correlation properties of the atmosphere, three identical autonomous sensor systems will be developed and deployed. Operations of these items and the associated control systems are discussed in the following paragraphs. A hardware block diagram is shown in Fig. 1.

A. Telescope Mount

The telescope mounts are the original mounts designed by Autoscope Corporation, with several modifications that have been made in the past few years. Open loop control of the mount's equatorial friction drive system was tested and found to be accurate to within 2.3 mrad. The redesigned system has an error of less than 0.4 mrad to match the camera's field of view the through the telescope. In order to achieve this accuracy, encoders are used on the equatorial drives to provide closed loop feedback. Total error from the encoders, encoder mounting, and mount mechanical error is less than 0.4 mrad. A mount model may be used to further reduce the mount mechanical error. The encoders will improve system accuracy by removing errors associated with slippage, backlash, balance, and improper stepping from the motors. All components on the mount are rated to operate in temperatures from −20 to 40 deg C.

B. Optical System

A Meade ten-inch Schmidt–Cassegrain telescope focuses a star image onto a CCD camera, providing a field of view of 1.1 mrad. A filter wheel in the path allows selection of one of eight positions: three standard astronomical filters, V, R, and I; three narrow (10- or 20-nm) filters corresponding to laser wavelengths, 532 nm, 860 nm, and 1064 nm; plus a block position and a clear (no filter) position. Filter transmission curves are given in Fig. 2.

A telecompressor lens, which was used in the past, can increase the field of view up to three times. However, with a smaller field of view, daytime observations will provide better resolution of attenuation because of the decreased background sky per pixel observed. Longer exposure times are possible before saturation of the pixels. Increased pointing accuracy, achieved by using encoders, enables operation without the telecompressor.

The optical path will be calibrated before operations begin, so that losses through each of the optical elements will be well known. In addition, a calibration lamp will be placed in the dome to provide a reference source for changes in telescope losses. Observations of this source several times daily will indicate any changes in optics throughput due to dust accumulation or condensation on the Schmidt plate of the telescope.

C. CCD Camera

The CCD camera is a Spectra Source Lynxx PC camera with 192 × 165 pixels in an area of 2.64 × 2.64 mm and 12-bit resolution. It is responsive over the 500- to 1064-nm wavelength range, and has a pixel area small enough to discern stars from the background sky during the day. The pixels are rectangular, 13.75 μm (192 dimension) by 16 μm (165 dimension). Combining this CCD with the telescope gives an image size of 1.1 mrad with a pixel resolution of 0.005 × 0.006 mrad. The detector is cooled to −10 deg C inside a sealed housing. Quantum efficiency and responsivity are indicated in Fig. 3.

D. Control Computer

The observatory is controlled by a 386-based computer. This computer is placed inside an environmentally controlled electronics cart and is tested to withstand operating conditions at altitudes of 3000 meters, operating temperatures of 10 to 30 deg C, and storage temperatures of −20 to 40 deg C.

Control boards inside the chassis include CCD control, telescope and observatory control, communications ports, clock control, and a monitor driver board. The CCD board is matched for each camera. The telescope and observatory control board is capable of driving the telescope stepper drivers, encoders, and filter wheel stepper driver; monitoring mount limit switches and joystick control; and driving or monitoring various input/output (I/O) modules. This board runs with a C program and will control sidereal rate, ramp, and slew operations. It has a run-time library of commands for various interfaces with the board, a joystick matched for operations with the board, and adjustable control parameters to compensate for differences between the three telescope mounts.

Three communications ports are used for connections to the uninterruptible power supply (UPS), modem, and weather station. A clock board regulates the DOS clock to within 0.5 seconds in order to properly track sidereal motion. All control boards are rated to operate at altitudes of 3000 meters, operating temperatures of 10 to 30 deg C, and storage temperatures of −20 to 40 deg C.

E. Observatory and Telescope Integrated System (OTIS)

OTIS controls telescope movement, filter control, and various other observatory operations. A block diagram of its components and connections is shown in Fig. 4. Major components are the microstepper controllers for the mount stepper motors, a stepper controller for the filter

wheel, CCD thermoelectric cooler control, I/O modules, watchdog relay, and various cable interfaces. OTIS will be placed inside the electronics cart, so its components are rated for the same altitudes, operating temperatures, and storage temperatures as the computer.

The microsteppers that control the telescope mount motors are adjustable for fine tuning from 1/256 step to 1 step. Optimum adjustment is at 1/16 step. They efficiently transfer power to the motors to prevent excessive heating of the system, and supply a high voltage to allow for high-speed operation of the mount axes. Voltage supplied to the motors can be up to 60 V. Previous drivers used a maximum of only 24 V.

The filter stepper driver is a single-step 12-V driver compatible with the PC digital-signal-processing controller. The CCD thermoelectric cooler (TEC) is compatible with the existing sensor in the camera (AD590) and controls at -10 to within ± 0.05 deg C. The cooler has an analog output so it can be monitored by the computer.

The analog and digital I/O control modules receive inputs from devices needing computer monitoring (24 V ac or 5 V dc) and provide outputs for computer control (dry contact output, 117 V ac, 24 V ac). Items controlled by the I/O modules are the calibration lamp, roof open, roof close, watchdog alarm, roof alarms, UPS alarm, high cart temperature alarm, watchdog notification, roof limit switch status, rain alarm, and wind speed alarm. The watchdog relay uses a 24 V ac input to switch 24 V ac to inform the roof controller, alarm system, and computer if the timer has not been updated within ten minutes.

The OTIS power supply provides ± 12 V at 2 A each, $+5$ V at 10 A, and $+48$ V at 5 A. Various connector interfaces are also provided within OTIS.

F. Weather Station

The weather station monitors wind speed and direction, inside (cart) and outside (ambient) temperatures, and relative humidity. Sensors located on the weather mast are capable of surviving 90-m/sec gusts and temperatures from -20 to 40 deg C. Alarms are available to shut down the observatory in case of precipitation or wind speeds greater that 13 m/sec. The weather station allows for three additional 0–5 V analog signals to be monitored. The signals monitored through the weather station are the CCD temperature (from the CCD TEC), the calibration lamp intensity, and precipitation. Alarms for precipitation and high wind speed drive the alarm system and roof controller, and are input to the I/O modules in OTIS. Weather conditions and voltage inputs are recorded via an RS232 connection to the computer. The new weather station was recommended by the National Oceanic and Atmospheric Administration (NOAA).

G. Enclosure Control

Observatory enclosures, including the building, roll-off roof structures, and electronics carts, must operate reliably through adverse weather conditions. New actuators were purchased to reliably control operation of the south walls of the buildings. They allow the walls to travel between fully upright and horizontal positions to allow full visibility to the south, and are sealed to prevent water contamination of the drive systems. The roofs must reliably close and open when instructed by the computer or alarm systems to allow quality observations and equipment safety. The original enclosure was not designed to operate under these expected adverse conditions. However, the second two enclosures, which are of a different design, have been operating reliably at high altitude sites. If the third observatory is moved from the JPL mesa to a new site, an improved enclosure will be required.

The electronics carts house all electronics that can be placed remotely from the sensors. The carts have been modified to allow temperature control by an industrial air conditioner, and have been strengthened to support more weight than originally designed.

H. Environmental Controls

An air conditioning and heating system controls the cart temperature, keeping it between 15 and 27 deg C. Normal operating temperature is 20 deg C. The air conditioner provides temperature control and dehumidifying of the cart using a closed-loop system, which recycles the air inside the cart to ensure a clean operating environment. The cart is vented between sections to improve air circulation. The heater is thermostatically controlled and turns on if temperatures fall below 10 deg C. It is rated to turn on in temperatures as low as -20 deg C. The observatory enclosure is vented to allow heat generated by the air conditioner to escape while the building is closed.

I. Power Conditioning

Power requirements for the observatory are 500 W for the main roof, 330 W for the south wall actuator, 1640 W for the air conditioner, and approximately 600 W for the electronics cart. A 1-kVA UPS provides backup power and surge protection for the instruments in the electronics cart. The roof motors and air conditioner have no backup power, but sites selected for the observatories should have

a backup generator for these items. The main roof can be closed manually in case of power failure.

J. Alarm System

The alarm system monitors eight alarm conditions plus power failures. It is a dialing alarm system, which will call a series of telephone numbers to report alarm conditions until an acknowledgment is received. The system can be programmed to respond differently according to the conditions of the individual alarms. It has a battery backup to retain memory in case of power failure. Inputs to the alarm system are dry contact closures.

III. Operations

The automated observatory will continuously monitor sky conditions by observing bright stars using the CCD camera and telescope optics. Observations through each of the filters will indicate atmospheric transmission at different wavelengths. The system is unattended, so all components must operate reliably without supervision for weeks at a time. Maintenance visits will be scheduled periodically for system upkeep. The system must be properly calibrated to accurately determine site conditions.

The system is controlled by an observing program written in C++. The program will schedule operational events based on a priority queue, and respond to varying environmental conditions. Data from daily observations are transferred via modem to JPL, where they will be analyzed. Data returned will consist of observatory and weather status and CCD observations in a flexible image transport system (FITS) format.

IV. Conclusions

The redesign of the automated observatories will provide a reliable system capable of collecting all required data. The engineering model that was developed over the past several years identified several components of the system that required redesign in order to ensure reliable operation. The hardware described in this article is a result of the lessons learned through development of the system. A hardware reference book is being kept that contains all specifications and manuals for hardware contained in the system.

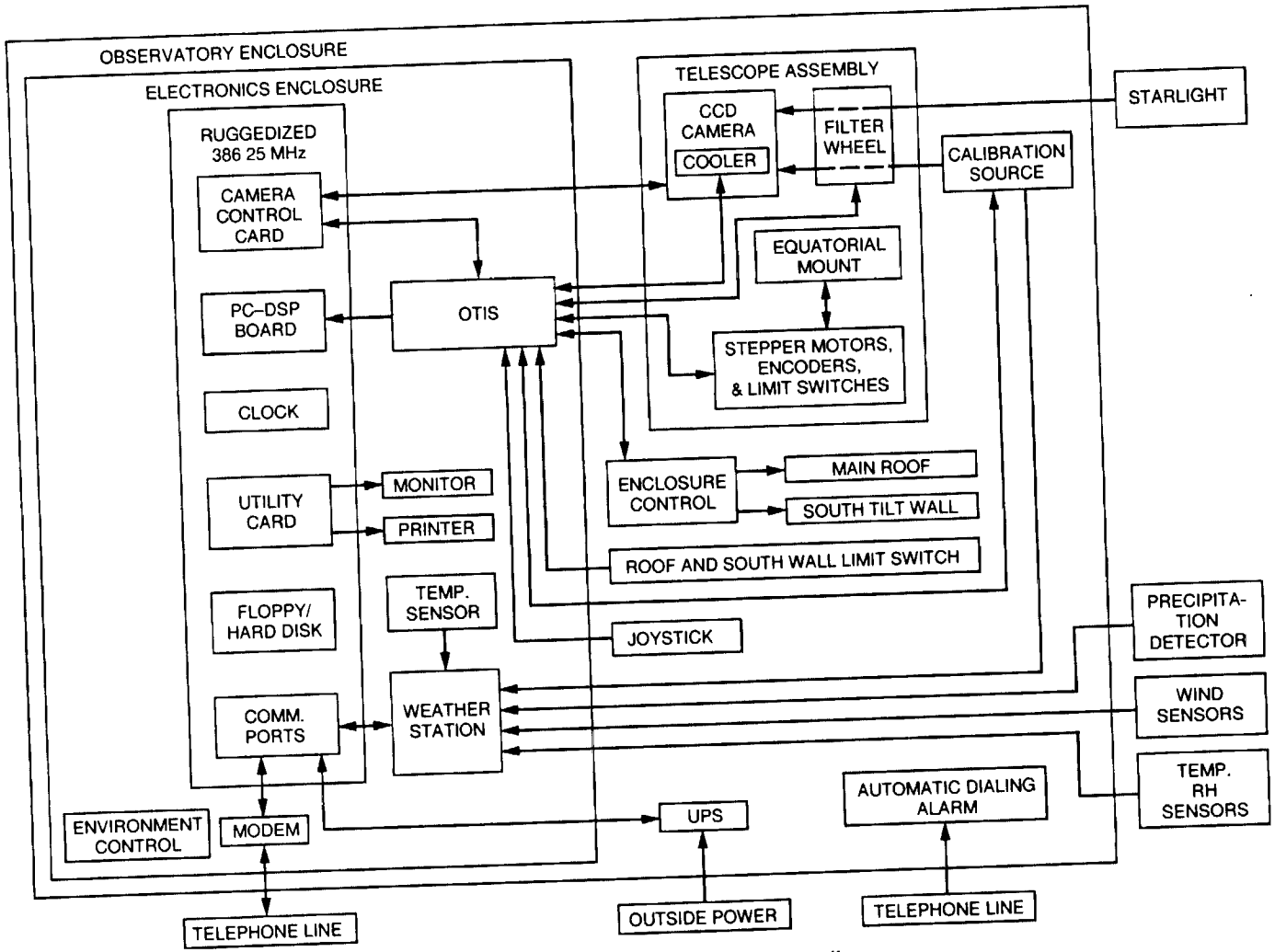


Fig. 1. Automated observatory hardware flow diagram.

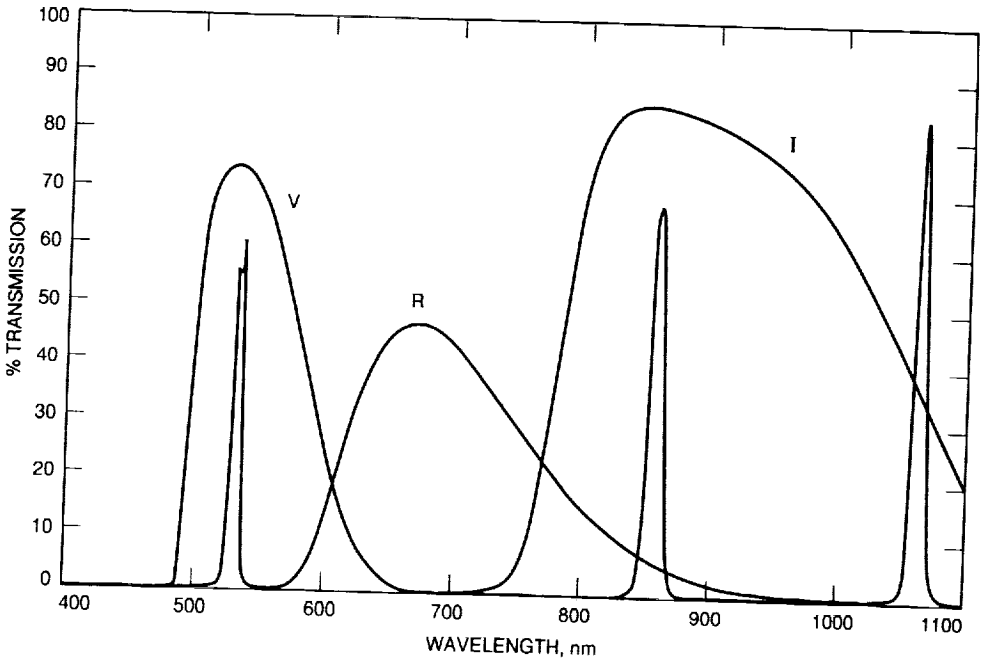


Fig. 2. Filter characterization of 1-nm resolution.

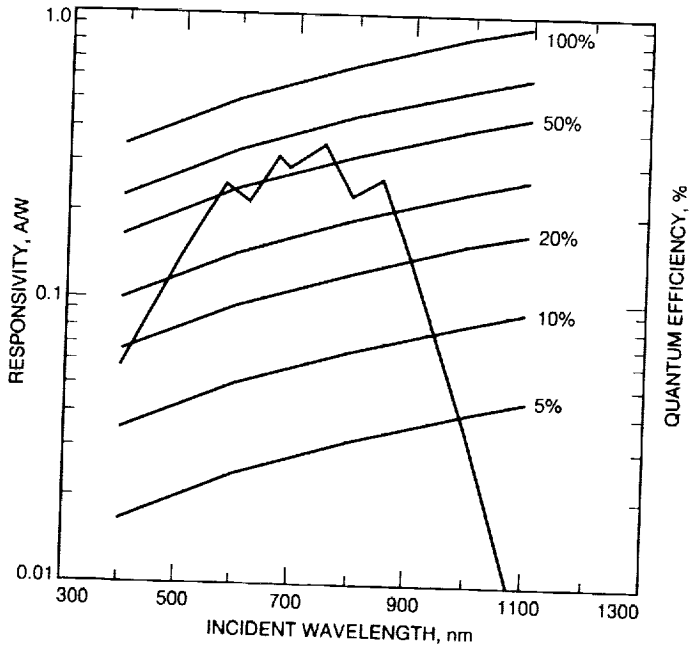


Fig. 3. Lynxx CCD sensor spectral characteristics.

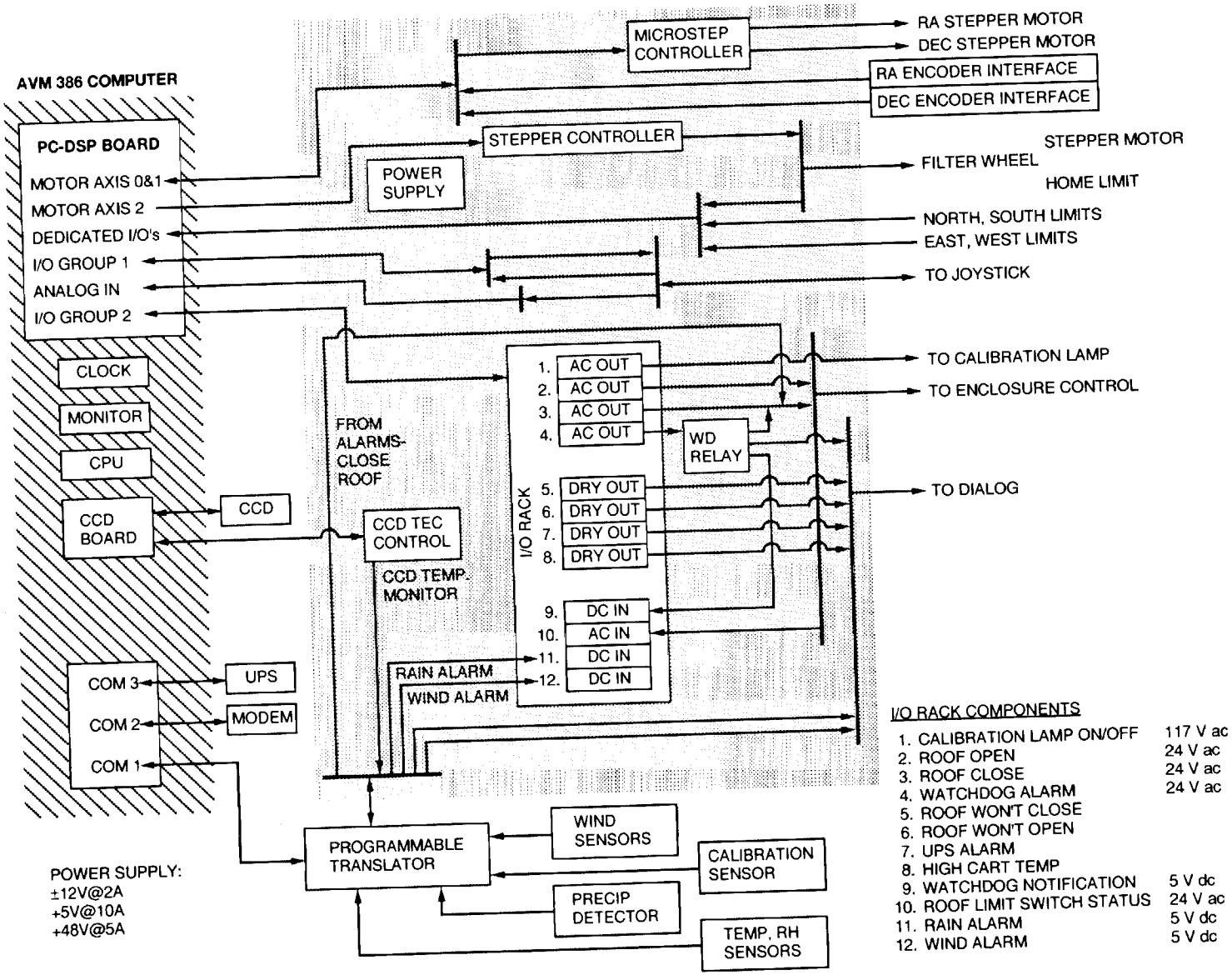


Fig. 4. Observatory and telescope integration system (OTIS).

TDA Progress Report 42-114

August 15, 1993

Layout and Cabling Considerations for a Large Communications Antenna Array

R. T. Logan, Jr.
Communications Systems Research Section

In this article, layout considerations for a large deep space communications antenna array are discussed. A novel fractal geometry for the antenna layout is described that provides optimal packing of antenna elements, efficient cable routing, and logical division of the array into identical sub-arrays.

I. Introduction

Phased arrays of many small antennas have been proposed as alternatives to large antenna structures for communications and navigation support of deep-space missions. Large arrays of antennas have been constructed previously for radio astronomy applications, but the considerations pertaining to a telemetry receiving array are fundamentally different. In a radio-astronomy array, the requirement for high angular resolution motivates a large inter-element spacing. However, the design goal for a communications array is to maximize the total gain, which increases with the number and size of the elements, but does not depend on the physical dimensions of the array. To keep construction costs low, the land area and cable requirements should be minimized. Therefore, the densest packing of the antenna elements is desired for a communications array.

In this article, the design constraints for a hypothetical communications array will be defined. A novel fractal array geometry that satisfies all of the design constraints will be presented and then compared to a conventional square-grid layout geometry. The fractal layout geometry may also be advantageous to other types of phased-array antenna systems, including printed-circuit microwave and millimeter-wave antenna arrays.

II. Design Constraints

Throughout this article, the following design constraints will be assumed to illustrate the layout and cabling considerations of a large communications array:

(1) No shadowing of adjacent antenna elements is permitted over the entire tracking range of 360 deg in azimuth, to an angle $\alpha = 10$ deg above horizontal.

(2) The most dense packing of antenna elements without shadowing is desired to minimize the physical size and cabling requirements of the array.

(3) The whole array should be divisible into independent, identically shaped subarrays, so that the correlator/combiner hardware and software are not subarray-dependent.

(4) The array will be operated from a single, centralized control facility, so that all antenna cables must be routed to this facility.

The analysis presented below addresses each of these questions in turn.

III. Shadowing

It is assumed that the antennas are all pointed toward a distant source such as a spacecraft, and that the received wave front is nearly planar when it reaches the antenna array. The shadowing of adjacent antennas will determine the lowest tracking elevation angle, as sketched in Fig. 1.

When all antennas point in the same direction, the minimum distance required between them is determined by the diameter, D, of the antennas, and the minimum required tracking elevation angle, α, from horizontal: $L_{min} = D/\sin\alpha$.

IV. Packing

From the results of the previous section, each antenna must be located a distance L_{min} away from any other antenna to avoid shadowing. This is guaranteed if each antenna element occupies an effective circular area of diameter $L_{eff} = L_{min}$ that does not overlap any other element. For the telemetry-receiving application, the gain of the array depends only on the sum of the collecting areas of the individual antenna elements. Another consideration is to minimize the amount of cable needed to connect the antennas to the central control facility. From these considerations, the optimum antenna layout is simply the most dense packing arrangement for circles of area $A_{eff} = \pi\left(\frac{L_{eff}}{2}\right)^2$

From mathematics or solid-state physics, it is known that the most dense packing arrangement of spheres is the hexagonal close-packed (hcp) arrangement, illustrated in Fig. 2. In the hcp arrangement, the land area required by each antenna is equal to the area of the hexagon inscribed around each circle, as shown in Fig. 3. The area of this hexagon is $0.866L_{eff}^2$. If a square-packing arrangement is used, each antenna requires a square area of L_{eff}^2. Obviously, other less-dense packing arrangements require yet more area. Since the hcp packing arrangement is the densest possible, the cable lengths between antennas will be minimized in this arrangement.

V. Layout

For maximum configuration flexibility, it is desirable to break the large array into smaller sub-arrays that may be used independently or arrayed together to synthesize various size apertures. The sub-arrays should have equal numbers of elements and be identically shaped to allow maximum flexibility and compatibility of the sub-arrays with the correlator/combiner, monitor and control, and receiver subsystems. To conserve the amount of land and cabling required for the array, the sub-arrays should interlock perfectly so that no space is wasted between sub-arrays.

It was noted in the previous section that the closest arrangement of circles is the hcp, so that each antenna occupies a hexagonal unit cell of area $0.866L_{eff}^2$. Given the sub-array considerations outlined above, the layout of the array can be cast as a tiling problem, in which the hexagonal unit cells must be arranged in some larger sub-array shape that is perfectly interlocking with itself. At first glance, a larger hexagonal shape is a possible candidate for the sub-arrays, since the hexagon interlocks perfectly with other adjacent hexagons, as seen from Fig. 3. However, "it is a widespread source of irritation that hexagons put together do not quite make up a bigger hexagon" [1], i.e., it is impossible to construct identical hexagonal sub-arrays from the single-antenna hexagonal unit cells.

A fractal object called the Gosper snowflake [2,3] solves the problem of interlocking sub-arrays while maintaining the hcp arrangement. The Gosper snowflake is formed from a recursive tiling of hexagons. Starting with a single hexagon, the first-order Gosper snowflake is created by breaking up each face into segments of equal length, such that the original area of the hexagon is preserved. The first-order transformation is equivalent to grouping seven hexagons together, as depicted in Fig. 3. The second-order transformation groups seven of the first-order objects together, as in Fig. 4. The third-order transformation groups seven second-order objects together as in Fig. 5.

The second-order Gosper snowflake is a fractal object that retains its basic shape after subsequent transformations. Note that the general shapes of the second-order and third-order snowflakes in Figs. 4 and 5 are similar, although the perimeter becomes more complex. This is a distinguishing feature of a fractal object.

Clearly, in the Gosper snowflake, the hcp structure is maintained, so that the closest possible packing of antennas is achieved. But more importantly, Nth-order Gosper snowflakes interlock perfectly to form another Gosper snowflake of order $N+1$. These two properties together solve the problems of most efficient packing and division of the array into identical sub-arrays.

VI. Cable Layout Considerations

In this large array structure composed of hundreds, or even thousands, of individual antenna elements, the cable routing scheme deserves careful consideration, since it determines not only the amount of cable required, but the installation technique. To motivate the discussion of the cabling scheme, a set of reasonable requirements on the signal distribution scheme is assumed. The details of the signal distribution and combining hardware are not the subject of this article, but for the purposes of this analysis, it is assumed that each antenna will require (1) a fiber-optic cable, containing multiple fibers for RF signals, frequency references, and monitor/control data and (2) a power cable. Optical fiber for RF and digital signals and wires for ac power can be contained in the same cable, since the optical fiber is immune to electromagnetic interference.

The basic requirements for the cabling scheme are assumed to be

(1) Minimize the cable lengths.

(2) Minimize the number of cable varieties used throughout the array.

(3) Route all individual fiber-optic antenna cables to a central signal-processing location at the center of the array.

(4) Minimize cable installation and maintenance costs.

(5) Minimize environmental perturbations experienced by the cables.

Costs, attenuation, and differential phases between elements are minimized when the lengths of the individual cable runs are minimized. It is desirable to use as few types of cable as possible throughout the array, since the unit cost of cable is lower when purchased in large quantities. Routing all antenna cables directly to the central processing facility saves construction costs and facilitates operation of the array. Finally, the cables should be protected from extreme environmental perturbations, to minimize phase shifts due to thermal expansion.

The least expensive installation technique for fiber-optic cable is direct burial with a tractor-pulled cable plow. Also, the thermal stability of direct-buried cable is excellent, with peak-to-peak variations in the ground temperature at DSS 15 at Goldstone at a depth of 1.5 m measured to be less than 0.1 deg C daily, over several months,[1] compared to variations on the order of 40 deg C at the surface. Also, commercial direct burial fiber-optic cable has a mean-time-to-failure (MTTF) of approximately 100 yr. For these reasons, direct-burial of the fiber-optic cable is far superior to above-ground cable trays or underground conduits in terms of installation cost, stability, and MTTF. However, direct burial by plowing typically requires that cables not cross each other, to avoid snagging a previously buried cable with the plow. Therefore, in a large antenna array, the cable layout must be carefully planned to avoid crossing of cables.

Applying the reasoning used in the packing scheme to the cabling problem, the Gosper snowflake layout offers a naturally centralized cabling scheme with no crossing of cables for arrays as large as 343 elements. The basic element of the cabling scheme is a star that connects six outer elements of the first-order Gosper snowflake to the central element, as shown in Fig. 6. The seven cables from the individual antenna elements are spliced to a larger trunk cable at the central antenna site of the first-order snowflake. Then, each of the six first-order elements in the second-order snowflake is connected using a similar star-shaped structure, as illustrated in Fig. 7. This scheme can be repeated for the third-order snowflake, as in Fig. 8, in which 343 elements are connected to the center of the array without crossing of cables. For a snowflake of order 4 or larger, the trunk cables connecting the third-order snowflakes to the center cannot be straight runs if cable crossing is to be avoided.

Each of the fiber cables in Fig. 6 contains N fibers, where N is determined by the signal and monitor and control requirements for a single antenna. The 49-element second-order sub-array cabling scheme proceeds in a similar fashion, as in Fig. 7, but each of these fibers contains $7N$ fibers. Finally, for the 343-element array of Fig. 8, the largest cables contain $49N$ fibers each. An attractive feature of this cabling scheme over a daisy-chaining scheme is that only three types of fiber are required for the entire array: N-fiber cable, $7N$-fiber cable, and $49N$-fiber cable.

Note that none of the cables in any of these arrays crosses another. This is a subtle, yet extremely important, characteristic of this cable-routing scheme, since it provides for the possibility of direct burial of the cables using a tractor-pulled cable plow. The installation could proceed with the tractor plowing in all the N-fiber cable, then switching reels to the $7N$-fiber cable, and finally the $49N$ cables—never crossing trenches.

The distance between centers of adjacent $(n-1)$st-order elements composing an nth-order transformation is

[1] M. Calhoun, P. Kuhnle, and J. Law, "Environmental Effects on the Stability of Optical Fibers Used for Reference Frequency Distribution," presented at the 39th Annual Meeting of the Institute of Environmental Sciences, Las Vegas, Nevada, May 2–7, 1993.

$7^{(n-1)/2}\ L_{eff}$. Thus, the centers of adjacent 7-element first-order snowflakes in the second-order snowflake of Fig. 7 are separated by $7^{1/2}\ L_{eff}$, and the centers of adjacent 49-element second-order snowflake sub-arrays of Fig. 8 are separated by $7L_{eff} = 7^{(3-1)/2}\ L_{eff}$.

With this formula, the length of any of the fiber cable runs can be calculated as a function of L_{eff}, which depends on the diameter of a single antenna element, as discussed previously. Let the N-fiber cable used within a first-order snowflake be defined as type 1, the $7N$-fiber cable used in the second-order array as type 2, etc. By inspection of Fig. 8, for a 343-element array there are

(1) 49 first-order (7-element) snowflakes, each requiring 6 type-1 cables of length L_{eff}, for a total of $294L_{eff}$ m of type-1 cable

(2) 7 second-order (49-element) snowflakes, each requiring 6 type-2 cables of length $7^{1/2}\ L_{eff}$, for a total of $111.12L_{eff}$ m of type-2 cable

(3) 1 third-order (343-element) snowflake, requiring 6 type-3 cables, each of length $7L_{eff}$, for a total of $42L_{eff}$ m of type-3 cable.

From these numbers, the total length of the buried cable required is $L_{cable} = 447.12L_{eff}$ m. The total fiber length contained in those cables is $L_{fiber} = (294L_{eff}N) + (111.12L_{eff}7N) + (42L_{eff}49N) = 3129.84NL_{eff}$ m.

For comparison, consider a hypothetical array of $M = 343$ antennas laid out in a regular square grid pattern, separated by the minimum distance derived previously: $L_{eff} = D/\sin\alpha$, where D is the diameter of the antennas, and α is the minimum elevation angle to be tracked above the horizon. The grid will have sides of length $L_{side} = (M^{1/2} - 1)L_{eff}$. Assume that the cables are buried in trenches, with a central cable trench running the length of the square array through the center, and horizontal trenches perpendicular to the central trench for each row. There will be $M^{1/2} + 1$ trenches, each of length $M^{1/2}\ L_{eff}$. The total trench length is thus $L_{trench} = (M + M^{1/2})L_{eff} = 361.52L_{eff}$ m. The total length of cable required to run individual cables in the trenches from each antenna to the center of this square array is well approximated by $L_{cable} = \left((M^{3/2})/2\right) L_{eff}$ for $M > 50$. Thus, for the square grid array with $M = 343$, $L_{cable} = 3176.22L_{eff}$ m. In this scheme, all of the cables contain N fibers, so that $L_{fiber} = 3176.22NL_{eff}$ m.

By this estimate, the fractal array uses less optical fiber to connect all of the antennas to the center. But the more significant feature of the fractal array layout is that the optical fibers from many antennas are combined into larger cable assemblies, like the branching of a tree. Thus, by the above estimate, the fractal array layout requires substantially less total *cable* length than the square array, for which each antenna requires a separate cable assembly that may only contain a few optical fibers. This can significantly influence the cost of a large array, since much of the cost of an optical fiber cable assembly is for the outer jacket materials, and the costs of installing a 100-fiber cable assembly or a 1-fiber cable assembly are essentially equal. Finally, the fractal geometry permits installation of the cables using a plow pulled by a tractor, at roughly 1/100 the cost of digging trenches, laying conduits, and pulling the cables through them.

VII. Conclusions

The layout and cabling problems of a large deep space telemetry-receiving antenna array were investigated. It is concluded that the antenna elements should be arranged in the hexagonal-close-pack configuration, and that the sub-arrays should have the shape of a Gosper snowflake. This arrangement provides the most dense packing without shadowing, so it requires the least amount of land and shortest cable runs. Also, the Gosper snowflake provides perfectly interlocking, identically shaped sub-arrays and enables a cabling scheme that does not require any crossing of cable trenches. This greatly simplifies construction, since all cables may be installed at a uniform depth using a tractor-pulled cable plow. This antenna layout and cabling scheme may also be generally useful for other types of phased-array antenna systems, including integrated printed-circuit microwave and millimeter-wave arrays.

This work was performed while the author was a member of the design team for the JPL Small Aperture Arraying task, headed by George M. Resch. The team produced a design and cost model for the synthesis of a large communications aperture using small antennas for the DSN.

References

[1] B. Mandelbrot, *Fractal Geometry of Nature*, chap. 6, New York: Freeman, 1977.

[2] M. Gardner, "Mathematical Games," *Scientific American*, vol. 235, pp. 124–133, 1976.

[3] L. Cook, "Fractal Fiber Optics," *Applied Optics*, vol. 30, no. 36, pp. 5220–5222, December 20, 1991.

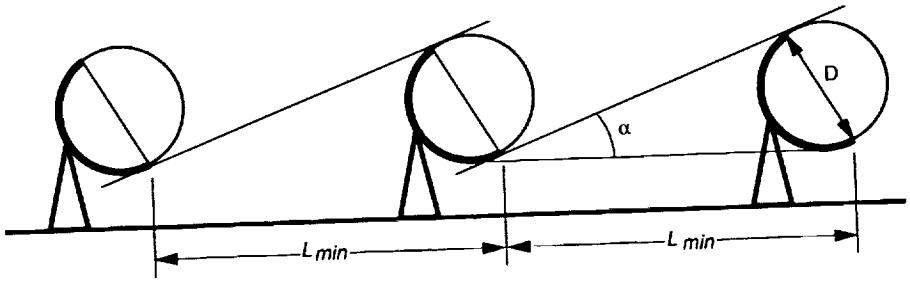


Fig. 1. Shadowing of adjacent antennas.

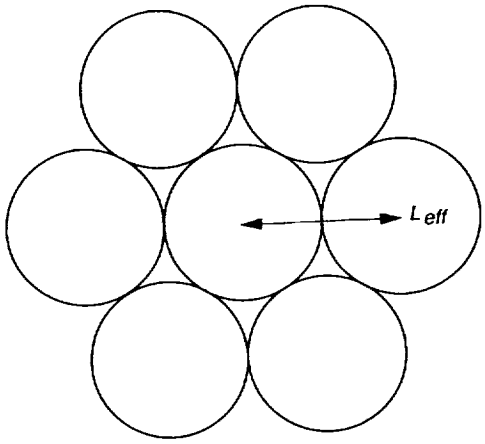


Fig. 2. Hexagonal close-packed structure.

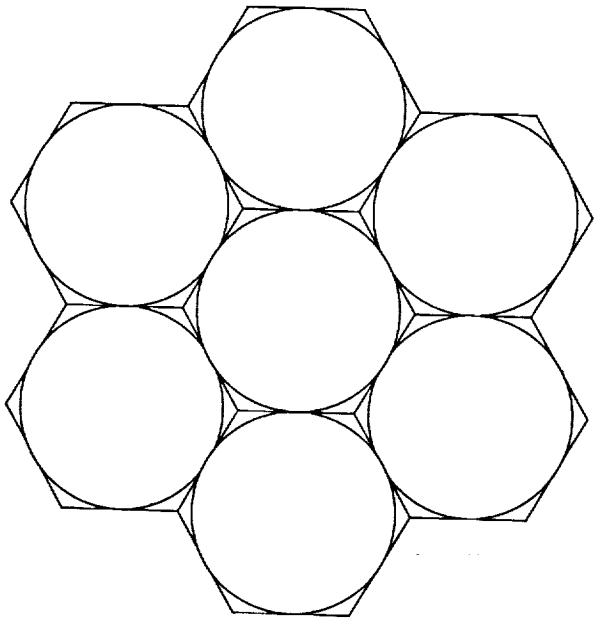


Fig. 3. Area required for hexagonal close-packed array.

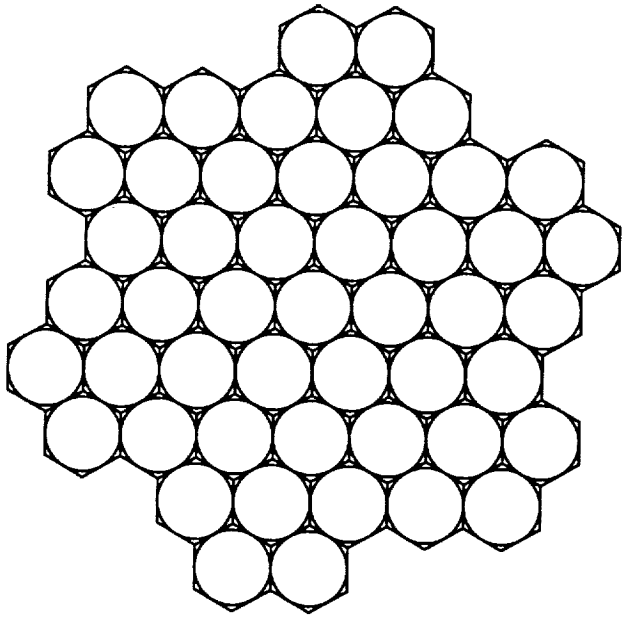


Fig. 4. Second-order Gosper snowflake, containing 49 elements.

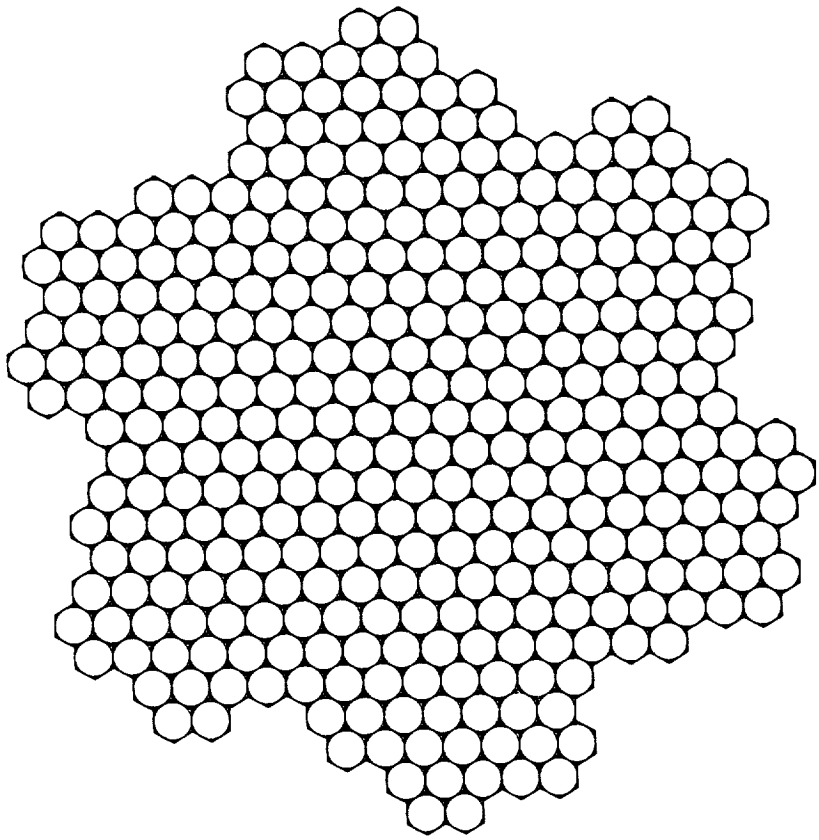


Fig. 5. Third-order Gosper snowflake, containing 343 elements.

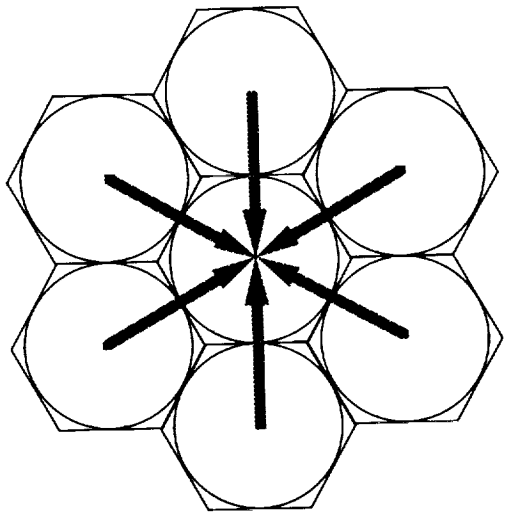


Fig. 6. Cabling scheme for first-order Gosper snowflake array.

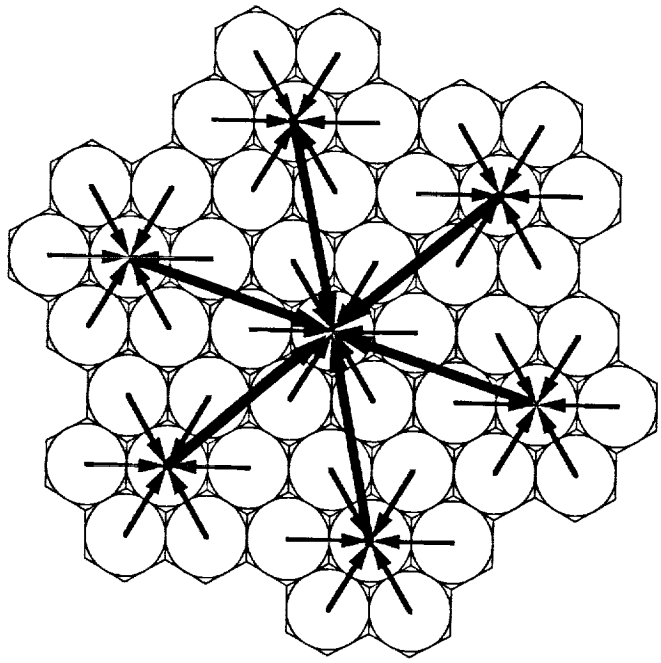


Fig. 7. Cabling scheme for second-order Gosper snowflake array.

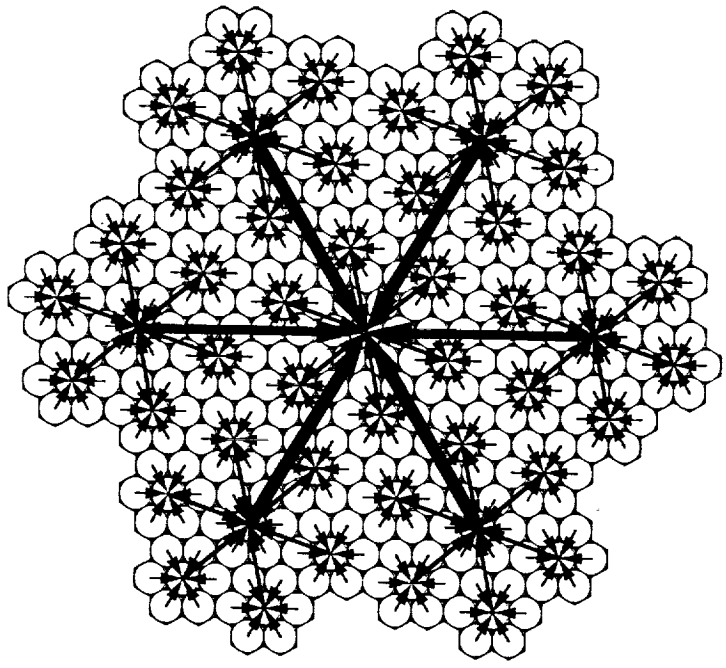


Fig. 8. Cabling scheme for third-order Gosper snowflake array.

TDA Progress Report 42-114

August 15, 1993

An Initial Study of Using the 34-m Antenna for Lunar Mission Support

A. Cha and P. Cramer
Ground Antennas and Facilities Engineering Section

As a part of the long-range planning of future Mars and lunar missions, a feasibility study has been made of a 34-m antenna system with differentially pointed multiple beams. The performance loss mechanisms of the differentially pointed multiple-beam systems were identified and quantified. Techniques that can significantly improve the multiple-beam system performance are identified. The goal is to determine the feasibility of using the 34-m antenna to support widely separated elements associated with lunar missions.

I. Introduction

Telecommunication plans for most future planetary missions include the use of 34-m antennas. For lunar missions, smaller antenna aperture sizes are also being considered, since the communications distance is short and the relatively large angle subtended by the Moon at the Earth cannot be covered by the beamwidth of a 34-m antenna at X-band (8.45 GHz) or Ka-band (32 GHz). Nevertheless, there is a good reason to consider using the 34-m antennas for lunar mission support, i.e., the high interest within the DSN in standardizing the 34-m beam-waveguide (BWG) antenna as a platform for cross-mission support. Indeed, the report on *America's Space Exploration Initiative* [1], assumed a lunar/Martian DSN cross-support strategy using 34-m BWG antennas with multiple differential pointing for lunar coverage.

To provide simultaneous coverage of several mission elements on the lunar surface, a number of feeds may be used on the 34-m antenna to realize a multiple-beam system. The Earth–Moon geometry is shown in Fig. 1. It is seen that the Moon subtends a half angle of 0.26 deg at the Earth. As shown in the tabulation that accompanies the figure, this is equal to about 15 beamwidths (BW) at Ka-band and 4 BW at X-band. (Note that the Ka-band frequency of 32 GHz is used expediently in the computations of this study in the interest of obtaining quick approximate results, although this is not the frequency assigned for lunar mission support.)

The large number of beams needed to cover the angle space from the antenna boresight to the lunar limb leads one to expect substantial performance losses, especially at Ka-band, of the 34-m multiple-beam antenna system. One well-known performance loss is scan loss, which refers to the loss of gain of the scanned beams as compared with that of the boresight beam. For a reflector-antenna type multiple-beam system, the severity of the scan loss is gen-

erally proportional to the number of beamwidths scanned, although not generally in a simple linear relationship.

The fact that the Ka-band beam which covers the lunar limb would be scanned 15 BW raises the question of the feasibility of using 34-m antennas for lunar mission support. On the positive side, the 34-m antenna starts with a much higher boresight gain as compared with a smaller antenna.

The purpose of the present study is thus to approximately quantify the performance of a 34-m multiple-beam antenna system in order to aid in planning the next phase of studies. It is noted that no serious study has been made of the DSN 34-m antenna scanning characteristics and no current database exists on the 34-m antenna scanning characteristics.

II. The Analytical Model

The analytical model is shown in Fig. 2. The antenna model is the 34-m Cassegrain dual-shaped reflector design of DSS 13. A number of feeds, each generating a beam in the far field, are assumed to be in the focal plane at the Cassegrain focus. It is assumed that the antenna is boresighted at the center of the Moon and the feeds are movable by mechanical means in order to cover moving mission elements on the Moon.

Clearly this is a rudimentary multiple-beam system whose performance can be significantly improved by various means, given more study time. However, this first study has the limited goal of establishing ballpark performance data for a simple 34-m multiple-beam antenna. More complex systems and more elaborate analysis efforts can be undertaken in the next phase of study.

III. Best and Worst Cases in Multiple-Beam Antenna Gain Losses

The best- and worst-case scenarios in using a reflector antenna-based multiple-beam system to support lunar missions are shown in Fig. 3. In the best-case scenario, all mission elements on the lunar surface are widely separated in angle space, see Fig. 3(a). Each mission element can then be supported at the peak gain point of a beam of the antenna. While in the worst-case scenario, three or more mission elements are closely spaced and can only be supported by two beams from two side-by-side horns, as shown in Fig. 3(b). In this case, it is not possible to support the three mission elements by using three beams because there is a minimum amount of spacing between adjacent beams, imposed by the finite physical size of the feedhorns. In the very worst-case scenario, shown in Fig. 3(b), one mission element is located in the direction where the two beams cross over. This scenario would sustain a somewhat higher loss than it would if it could be supported by the peak gain of some beam.

Figure 4 shows the radiation patterns due to four Ka-band side-by-side horns in the 34-m antenna. The first horn is positioned at the Cassegrain focus and gives the boresight beam of the antenna. The best-case antenna gain-loss characteristic for supporting a lunar mission is the curve connecting the peak gain point of the individual beams in Fig. 4. The curve of the peak gain of the beams gives the gain- versus scan-angle curve, defined as the scan loss in reflector antenna literature. The worst-case gain loss is seen to be the locus connecting the crossover point of the main beams generated by side-by-side horns. It is seen that the difference between the best- and worst-case gain loss is at a maximum near the antenna boresight. This is an aspect of multiple-beam antennas that may not be as well recognized as the scan loss, but it is one that certainly needs some attention in future studies and designs.

The crossover level of the first two beams is approximately −15 to −16 dB down from the peak gain of the boresight beam. This is a larger loss than the scan loss from boresight to lunar limb in the X-band case, and thus can lead to potentially bad surprises for lunar missions if not attended to properly.

In Fig. 5, the antenna gain loss versus the scan angle was plotted after a large number of 34-m antenna far-field patterns at Ka- and X-band were computed and data-processed. The Ka-band boresight beam gain is assumed to be 10 dB higher than the X-band boresight beam gain. This difference is based on DSS-13 gain accounting at the Cassegrain focus [2]. It is seen that Ka-band antenna gain is higher than that of X-band for much of the lunar surface. The crossover point of Ka- and X-band gain- versus scan-angle curves is 0.17 to 0.19 deg from the boresight. The X-band gain is, at most, 3 dB higher than that of Ka-band at large scan angles but has a large drop at 0.06 deg, the angle location where the first two X-band beams cross over. This observation may be of some interest as, in the early years of lunar missions, exploration activities may likely be confined to a portion of the lunar surface. It is also of interest to note that the 34-m antenna X-band multiple-beam system may be viewed alternatively as a 10-m antenna Ka-band multiple-beam system. Taking this view, one might conclude that the 34-m antenna gain is higher than that of a 10-m antenna at Ka-band for a partial lunar coverage from 0 to 0.17 deg. It would

be interesting and useful to determine the trade-off with other aperture sizes.

A final remark is that the analysis made is applicable at the Cassegrain focus of a BWG antenna. The feasibility of using the intermediate focus and the pedestal room focus of a DSS-13-type BWG antenna to implement a multiple-beam system is not clearly understood at this time.

IV. Summary

The plot of 34-m antenna performance loss versus scan angle characteristics has been established using a simple multiple-beam system model. In addition to quantifying the scan loss curves at X- and Ka-band, the beam-cross levels at small angles are identified as a potential problem that needs some design attention. For a 34-m antenna operating at Ka-band, the worst-case performance loss is determined by the scan loss from the center of the Moon to the lunar limb. For a 10-m antenna operating at Ka-band, the worst-case performance loss is determined by the crossover level of the first two beams rather than the scan loss from Moon center (antenna boresight) to the lunar limb.

It should be stressed that this represents the preliminary results of a study aimed at obtaining order-of-magnitude estimates of 34-m antenna scanning characteristics. In all likelihood, significantly better 34-m antenna scanning performance can be achieved with more design and some technology development efforts. The following elements could contribute to increased performance and implementation feasibility:

(1) Performance loss reduction techniques for 34-m antennas:

   (a) Computer search of the best focal surface (Petzval surface) for offset feed locations.

   (b) Arraying of adjacent feeds for a single beam.

   (c) Under-illuminating the main reflector aperture to reduce scan loss.

(2) Implementation options for 34-m BWG antennas:

   (a) Locations of the feeds on a 34-m BWG antenna.

   (b) Scanned beam losses in the beam-waveguide shroud.

(3) Performance loss versus antenna diameter trade-offs for lunar mission support.

(4) New reflector designs with improved scanning characteristics.

References

[1] T. P. Stafford, *America at the Threshold: Report of the Synthesis Group on America's Space Exploration Initiative*, Superintendent of Documents, no. 033-000-01097-4, Washington, D.C.: U.S. Government Printing Office, June 1991.

[2] D. A. Bathker, W. Veruttipong, T. Otoshi, and P. W. Cramer, "Beam Waveguide Antenna Performance Predictions with Comparisons to Experimental Results," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, no. 6, pp. 1274–1285, June 1992.

FREQUENCY	Ka-BAND, 32 GHz	X-BAND, 8.45 GHz
BEAMWIDTH OF BORESIGHT BEAM, deg	0.016	0.063
SCAN RANGE TO LUNAR LIMB, BW[a]	15.3	4.1

[a]ASSUMING ANTENNA BORESIGHTED ON CENTER OF MOON.

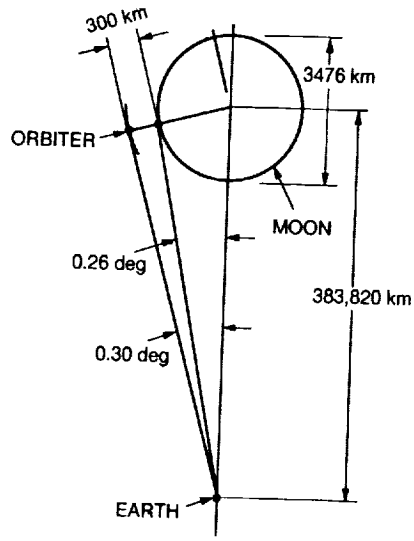


Fig. 1. Earth–Moon geometry.

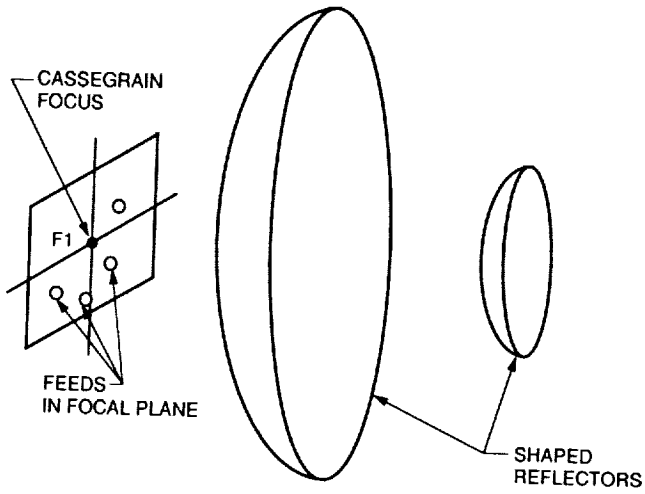


Fig. 2. Analytical model of the DSS-13 34-m dual-shaped reflector antenna, single feed per beam. The calculation matrix includes frequencies of 32 GHz and 8.45 GHz with the feed position offset in the focal plane.

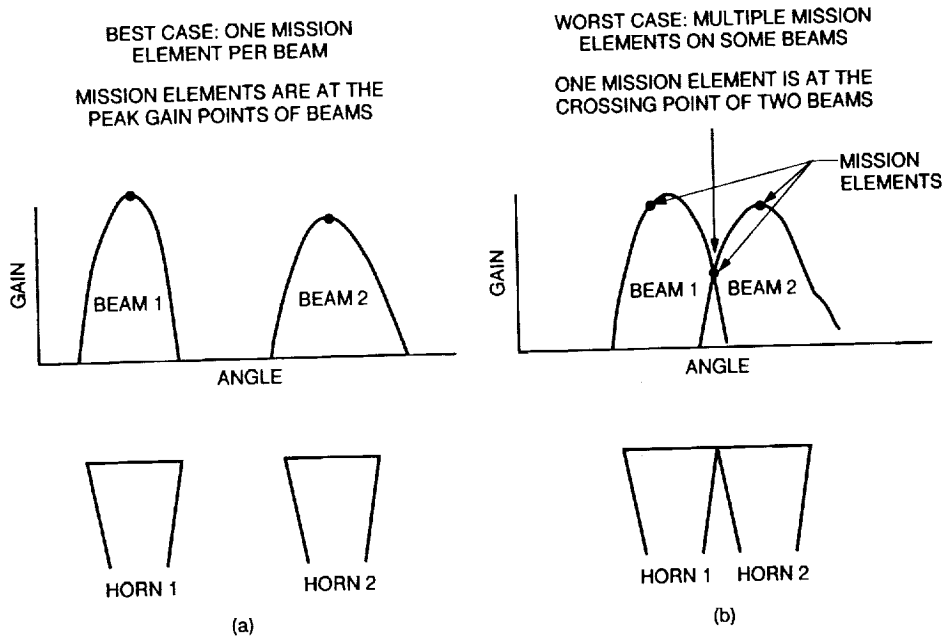


Fig. 3. Best and worst cases of 34-m antenna performance losses for: (a) separated horns and (b) side-by-side horns (minimum spacing).

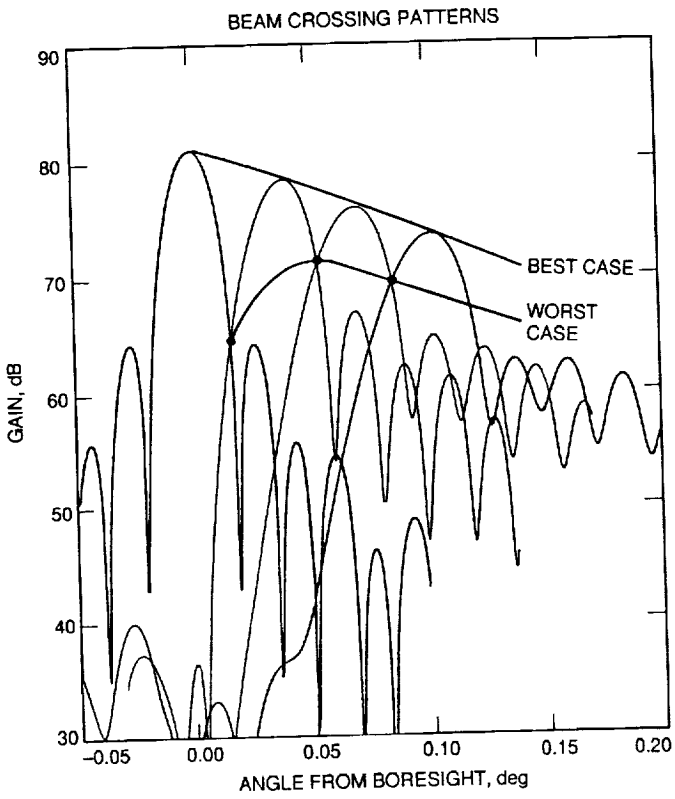


Fig. 4. 34-m antenna scan patterns, Ka-band (32 GHz).

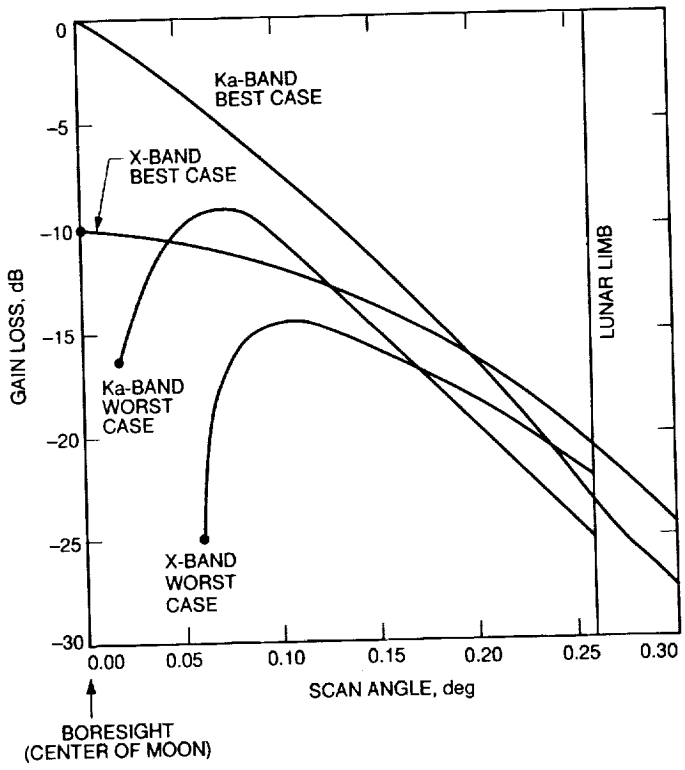


Fig. 5. Performance loss versus scan angle.

TDA Progress Report 42-114

August 15, 1993

HRMS Sky Survey Wideband Feed System Design for DSS 24 Beam Waveguide Antenna

P. H. Stanton, P. R. Lee, and H. F. Reilly
Ground Antennas and Facilities Engineering Section

The High-Resolution Microwave Survey (HRMS) Sky Survey project will be implemented on the DSS 24 beam waveguide (BWG) antenna over the frequency range of 2.86 to 10 GHz. Two wideband, ring-loaded, corrugated feed horns were designed to cover this range. The horns match the frequency-dependent gain requirements for the DSS 24 BWG system. The performance of the feed horns and the calculated system performance of DSS 24 are presented.

I. Introduction

The High-Resolution Microwave Survey (HRMS) project includes a sky survey over the frequency range of 1 to 10 GHz. The DSS 24 34-m beam waveguide (BWG) antenna will be used from 2.86 to 10 GHz; for frequencies below this, DSS 24 is unsuitable and another antenna will be utilized. To minimize complexity and expense, two wideband feed horns were designed to cover this broad frequency range on DSS 24. The performance of the horns was analyzed, and the upper-frequency horn was built and tested. The system performance of the horns installed on DSS 24 was also analyzed using computer software.

II. Feed-Horn Design

Two wideband conical feed horns were designed covering the frequency ranges of 2.86 to 5.35 GHz and 5.34 to 10 GHz. These ranges were chosen to match those of the upper-band HRMS orthomode transitions. A wideband feed horn for the DSS 24 BWG antenna needs to track specific gains and phase-center locations versus frequency to obtain high antenna efficiency and low spillover noise. Curves were fitted to the existing DSN narrowband designs for DSS 24 at 2.1625 GHz, 7.8725 GHz, and 33.25 GHz, and used as initial design values for the HRMS horns. Figure 1 shows these curves. The wideband (approximately 1:1.9) HRMS feeds for the DSS 24 antenna were matched to the fitted gain curve using the Gaussian beam model in [1] by adjusting the horn output flare angle and the aperture diameter. Fortunately, these wideband HRMS BWG feed designs adequately track the desired phase-center location versus frequency curve.

Once the horn output flare angle and the aperture diameter were determined, the detailed design followed the one presented in [2]. The horn parameters used in the design of the HRMS BWG feed horns are listed in Table 1. Figure 2 shows the four sections of the horn design: ring-loaded mode converter, frequency transition, angle transition, and output flare. Figures 3(a) and 3(b) show profiles

of the 2.86- to 5.35-GHz horn and the 5.34- to 10-GHz horn, respectively.

III. Horn Fabrication

The corrugated horn beyond the mode converter was fabricated using computer numerical control (CNC) lathe techniques. This procedure uses a computer-generated program written and checked on plotters before fabrication begins. It reduces the possibility of human error caused by repetitious movement.

The ring-loaded slot mode converter section at the horn input (see Fig. 2) presented a different problem. With the ring-loaded grooves, the mode converter section is impossible to fabricate as a single unit on conventional equipment. Several fabrication techniques were investigated. The cost, tolerance, and ability to make minor changes were the drivers in choosing the technique used. Each ring-loaded slot was split down the center, and half of each slot was machined into adjacent disks (Fig. 4). All of the disks had the same precise outside diameter. A holder with a precision bore and length was fabricated to align and compress these disks. Interfaces between the horn, the holder, and the disks were treated like typical waveguide joints. This design allows changes to individual ring-loaded slots by changing disks.

IV. Horn Performance

The horns were analyzed with a JPL computer program.[1] The program uses field matching techniques to determine the scattering matrices of the transverse electric (TE_{mn}) and transverse magnetic (TM_{mn}) modes of the horns [3]. From the scattering matrices, the aperture mode fields and the return loss of the horns are known. The radiation patterns of the horns can be calculated from the aperture mode fields by using the radiation integral.

The calculated gain and near-field phase-center location of the horns versus the frequency are shown in Figs. 5 and 6, along with the initial design gain and phase-center location. The calculated gain is somewhat higher than the initial design gain because of the approximation of the Gaussian beam model used in [1]. Computer analyses on DSS 24 revealed that the higher gain provided better noise temperature performance than the initial design gain but still had a comparable gain-to-noise temperature ratio (G/T).

[1] D. J. Hoppe, *Scattering Matrix Program for Ring-Loaded Circular Waveguide Junctions* (unnumbered, internal document), Jet Propulsion Laboratory, Pasadena, California, August 3, 1987.

The upper-frequency horn has been built and tested. The lower-frequency horn has been fabricated, but has not been fully tested. The return loss of the upper-frequency horn was measured on a Hewlett Packard 8510C network analyzer. The measured return loss, along with the calculated return loss of both horns, is shown in Fig. 7. Far-field radiation patterns of the upper-frequency horn were measured on JPL's Mesa Antenna Range at several frequencies from 5.34 to 11 GHz. The copolarization measurements matched the calculated patterns very well. Figure 8 compares one typical measurement with the calculated pattern. The maximum measured cross-polarization signal is 32.6 dB below the peak copolarization signal within the operating band of 5.34 to 10 GHz. The maximum error between the measured and the calculated values is 4.4 dB. Figure 9 shows the measured values for the 5.34- to 10-GHz horn, and the calculated values for both horns.

V. DSS 24 Performance

The performance of the DSS 24 BWG antenna with the HRMS feeds was predicted by using JPL computer programs[2] based on physical optics (PO) and physical optics/Jacobi-Bessel (POJB) techniques [4,5]. These programs are numerically intensive and consume large amounts of computer time. They were run on the JPL Cray Y-MP2E supercomputer, Voyager.

With the PO program results, it is possible to calculate the total spillover power of the BWG reflectors, the subreflector, and the main reflector. For the spillover power of the BWG mirrors, an effective noise temperature value of 280K was used in the pedestal room, and a value of 230K was used in the shroud above the pedestal room.[3] The effective noise temperature values for the spillovers past the main and subreflectors were taken to be 240K and 5K, respectively, for the antenna at zenith.

To find the best physical positions of the feeds, several feed locations were analyzed until the peak G/T was found. Five frequencies in each feed operating band were examined. The optimal position for each feed varies less than 23 cm over its frequency band. By locating the feeds in the center of their respective optimal position ranges,

[2] R. E. Hodges and W. A. Imbriale, *Computer Program POMESH for Diffraction Analysis of Reflector Antennas* (unnumbered, internal document), Jet Propulsion Laboratory, Pasadena, California, February 1992.

[3] W. Veruttipong and M. Franco, "A Technique for Computation of Noise Temperature Due to a BWG Shroud," JPL Interoffice Memorandum 3328-92-0149, Rev. A (internal document), Jet Propulsion Laboratory, Pasadena, California, November 3, 1992.

less than 0.05 dB/K is lost and the feed horns do not need to be relocated during operation.

To place the feeds in the pedestal room of DSS 24, a flat plate must be used with each feed. The 5.34- to 10-GHz horn will fit in the available space with its axis vertical and a flat plate above it (Fig. 10). However, the 2.86- to 5.35-GHz horn will not fit so easily. The horn will be placed with its axis horizontal, which, unfortunately, causes extra wear on the low-noise-amplifier refrigeration vacuum pump. Figure 11 shows the configuration of the horn and the mirror.

Beam squint was examined at the highest and lowest frequencies in both horn bands by running both right-circular and left-circular polarizations (RCP and LCP) of the horn patterns through the PO programs. The tilt angle varied from 6 millidegrees at 2.86 GHz to 2 millidegrees at 10 GHz. The tilt angle to half-power beamwidth ratio remained fairly constant, varying from 0.031 to 0.034. The loss due to squint ranged from 0.011 dB at 2.86 GHz to 0.013 dB at 10 GHz.

Major RF system requirements[4] are listed in Table 2. Table 3 shows the calculated system performance for the antenna at zenith, including noise temperature,[5] aperture efficiency, beam efficiency, gain,[6] G/T, and target G/T.[7]

VI. Conclusions

Two HRMS Sky Survey wideband (approximately 1:1.9) feed horns were designed to be compatible with the DSS 24 BWG antenna, covering a combined frequency range of 2.86 to 10 GHz. The upper-frequency horn has been fabricated and tested. The RF measurements were in close agreement with the calculated values in both the radiation patterns and return loss. Based on the results of the upper-frequency horn, the performance of the lower-frequency horn is expected to closely match the calculated performance.

The predicted DSS 24 aperture efficiency is greater than 0.747 over this frequency range. At the lower end of each of the two feed-horn frequency ranges, the noise temperature is higher than the 25 K maximum required. The predicted G/T, however, is at least 0.3 dB greater than is required over the entire frequency range. The beam efficiency is in the range of 0.772 to 0.881, short of the $\geq$0.9 requirement. The beam efficiency could be improved at the expense of the aperture efficiency.

A DSN frequency in the HRMS 2.86- to 10-GHz range available for comparison is 8.45 GHz (X-band). The design expectations of DSS 24 for DSN operation at 8.45 GHz, excluding the S-/X-band dichroic plate, are 0.766 aperture efficiency and 24.20 K noise temperature.[8] For the HRMS design at 8.45 GHz, substituting the DSN estimated values for the DSN preamplifier assembly for the HRMS preamplifier assembly, an aperture efficiency of 0.732 and a noise temperature of 23.16 K are predicted. The HRMS design has a resulting G/T of 54.56 dB/K compared to 54.57 dB/K for the DSN design.

[4] G. A. Zimmerman, *Search for Extra-Terrestrial Intelligence — Microwave Observing Project Sky Survey Element*, JPL 1720-4100 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 12, 1991.

[5] The noise temperature includes contributions from the sky and atmosphere, reflector spillover, reflector ohmic losses, main reflector leakage, horn ohmic losses, scattering from the subreflector support structure, and the low-noise amplifier (LNA) assembly. Sky and atmosphere and LNA assembly values from: JPL-ARC Front-End Design Team, *NASA SETI Common Radio Frequency System Design Team Report* (internal report), Appendix D, p. 6, NASA, Washington, DC, August 1, 1991. Reflector ohmic losses, main reflector leakage, and scattering from the subreflector support structure based on estimated values at 2.295 GHz and 8.45 GHz from W. Veruttipong and D. Bathker, JPL Interoffice Memorandum (unnumbered, internal document), Jet Propulsion Laboratory, Pasadena, California, February 7, 1992.

[6] The gain calculations include the following: illumination efficiency, reflector spillover loss, reflector ohmic loss, reflector surface RMS loss, main reflector leakage, BWG return loss, subreflector support structure blockage, reflector alignment loss, horn ohmic loss, horn return loss, and pointing squint. The values for all losses other than illumination efficiency, spillover losses, and horn losses are based on estimated values at 2.295 GHz and 8.45 GHz from W. Veruttipong and D. Bathker, JPL Interoffice Memorandum (unnumbered, internal document), Jet Propulsion Laboratory, Pasadena, California, February 7, 1992.

[7] The target G/T is for 65-percent aperture efficiency and 25 K noise temperature.

[8] W. Veruttipong and D. Bathker, JPL Interoffice Memorandum (unnumbered, internal document), Jet Propulsion Laboratory, Pasadena, California, February 7, 1992.

References

[1] P.-S. Kildal, "Gaussian Beam Model for Aperture-Controlled and Flareangle-Controlled Corrugated Horn Antennas," *IEE Proceedings*, vol. 135, no. 4, part H, pp. 237–240, August 1988.

[2] B. M. Thomas, G. L. James, and K. L. Greene, "Design of Wide-Band Corrugated Conical Horns for Cassegrain Antennas," *IEEE Transactions on Antennas and Propagation*, vol. AP-34, no. 6, pp. 750–757, June 1986.

[3] G. L. James and B. M. Thomas, "TE_{11} to HE_{11} Cylindrical Waveguide Mode Converters Using Ring-Loaded Slots," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-30, no. 3, pp. 278–285, March 1982.

[4] V. Galindo-Israel and R. Mittra, "A New Series Representation for the Radiation Integral with Application to Reflector Antennas," *IEEE Transactions on Antennas and Propagation*, vol. AP-25, no. 5, pp. 631–641, September 1977.

[5] Y. Rahmat-Samii and V. Galindo-Israel, "Shaped Reflector Antenna Analysis Using the Jacobi-Bessel Series," *IEEE Transactions on Antennas and Propagation*, vol. AP-28, no. 4, pp. 425–435, July 1980.

Table 1. Feed-horn design parameters.

Parameter	2.86- to 5.35-GHz horn	5.34- to 10-GHz horn
Mode converter section		
Low-frequency limit, f_L, GHz	2.86	5.159
Output design frequency f_i $(= 1.15 f_L)$, GHz	3.289	5.933
Normalized input radius, $k_i a_1$	2.9	2.9
Normalized converter length, z_i/λ_i	1	1
Conical semiangle, θ_i, deg	6	6
Slot width-to-pitch ratio, δ_i	0.75	0.75
Normalized slot pitch, p_i/λ_i	0.1	0.1
Frequency transition section		
Normalized length, z_t'/λ_i	3	3
Angle transition section		
Normalized length, z_t'/λ_i	1	1
Output flare section		
Conical semiangle, θ_i, deg	11	9.85
Normalized aperture radius, a_o/λ_L	3.44	5.33
Slot width-to-pitch ratio, δ_o	0.75	0.75
Pitch, p_o	0.1	0.1

Table 2. Major RF system requirements.

Parameter	Required value
Noise temperature, K	≤ 25
Polarization	RCP[a] and LCP[b] (simultaneous)
Instantaneous bandwidth, MHz	≥ 360
Aperture efficiency, %	≥ 65
Beam efficiency, %	≥ 90

[a] Right-circular polarization.
[b] Left-circular polarization.

Table 3. DSS 24 system performance at zenith.

Frequency, GHz	Noise temperature, K	Beam efficiency	Aperture efficiency	Gain, dB	G/T, dB/K	Target G/T, dB/K
2.86- to 5.35-GHz feed horn						
2.86	27.64	0.857	0.774	59.04	44.62	44.31
3.289	25.88	0.877	0.783	60.31	46.18	45.53
3.87	24.21	0.879	0.783	61.72	47.88	46.94
4.54	23.64	0.861	0.788	63.14	49.40	48.33
5.34	23.61	0.881	0.750	64.33	50.60	49.74
5.34- to 10-GHz feed horn						
5.34	25.56	0.810	0.787	64.54	50.46	49.74
5.933	25.03	0.803	0.786	65.45	51.47	50.65
7.1	24.60	0.792	0.767	66.90	52.99	52.21
8.45	22.27	0.803	0.757	68.36	54.88	53.72
10.0	23.30	0.772	0.747	69.76	56.09	55.19

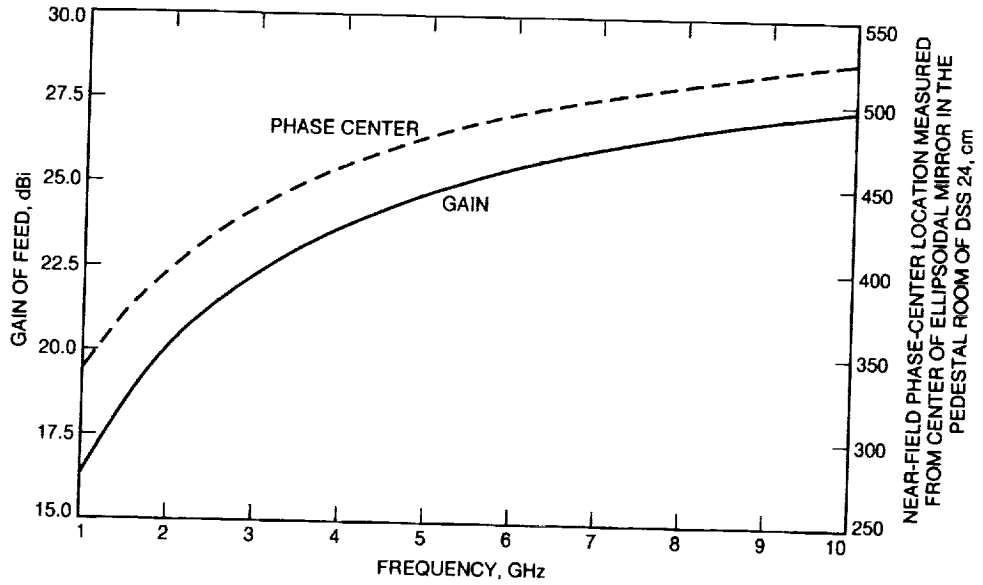


Fig. 1. Initial design gain and near-field phase-center location of feeds for DSS 24.

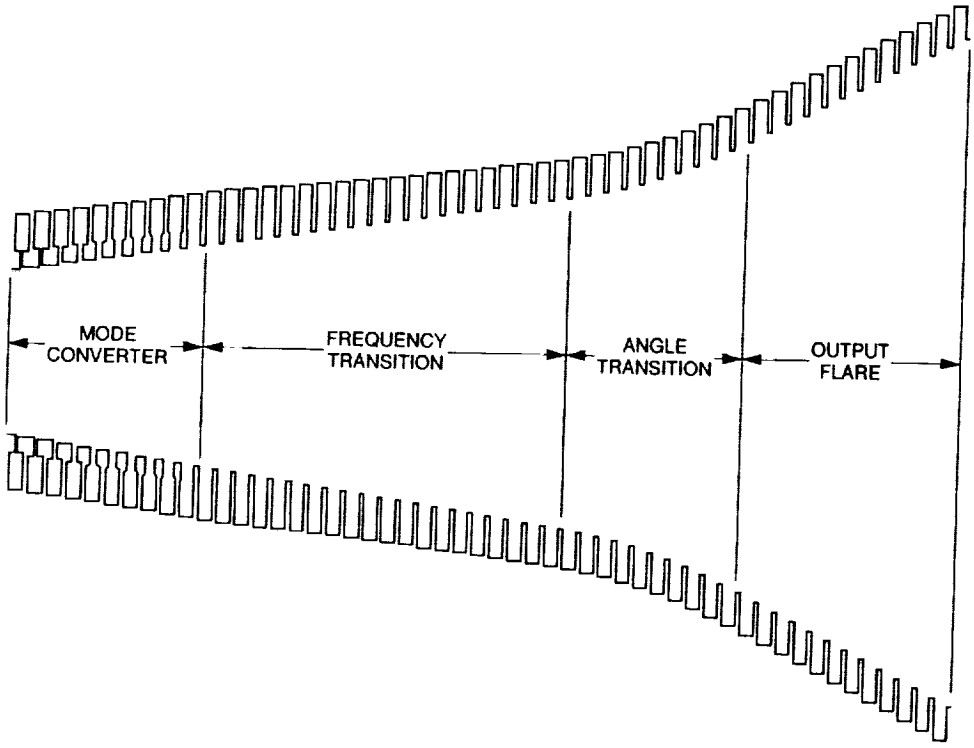


Fig. 2. Basic wideband horn configuration.

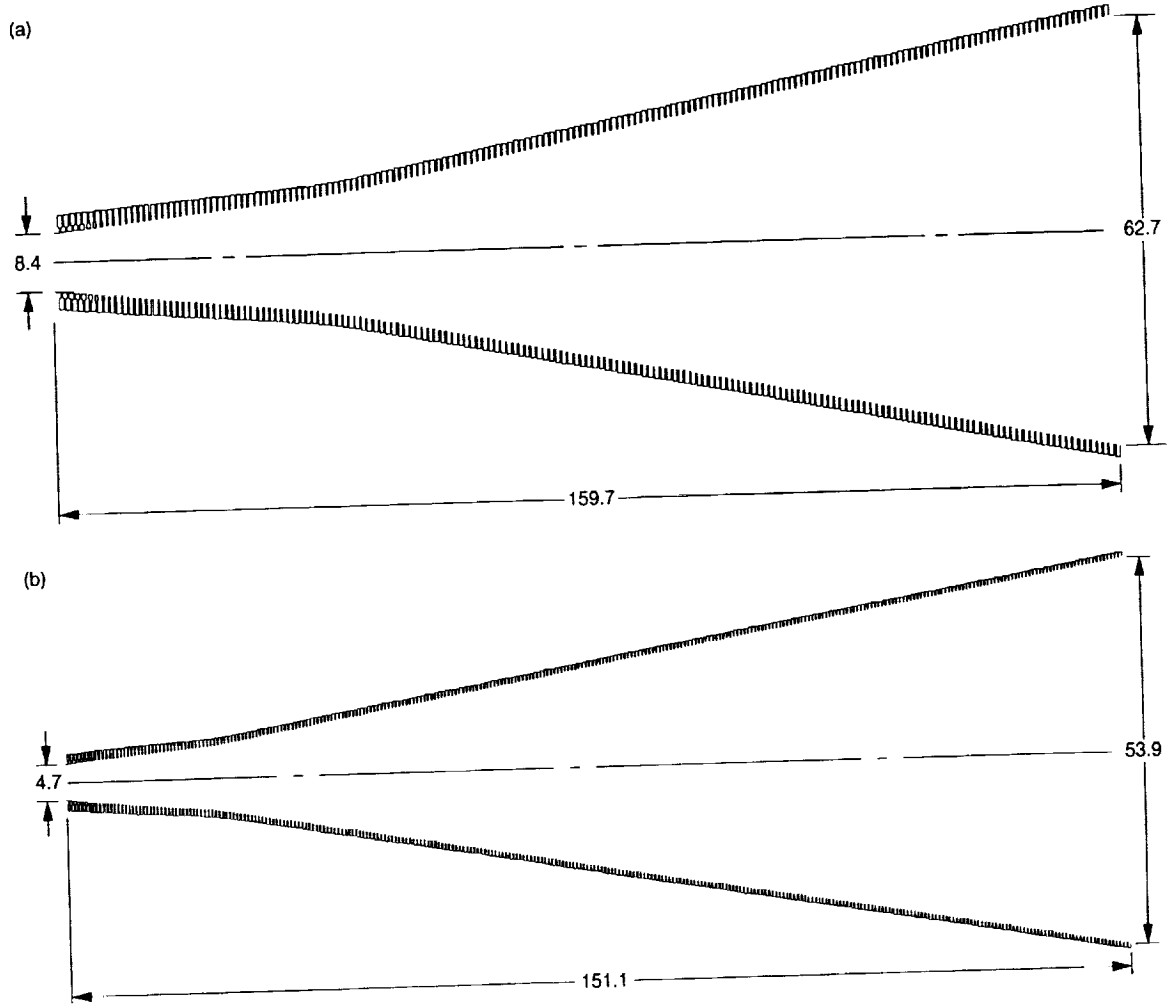


Fig. 3. Profiles of: (a) 2.86- to 5.35-GHz horn and (b) 5.34- to 10-GHz horn.

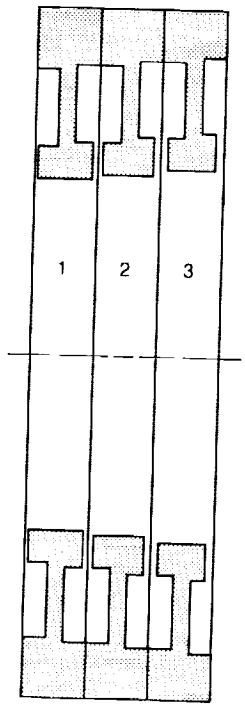


Fig. 4. Cross section of three adjacent mode converter disks.

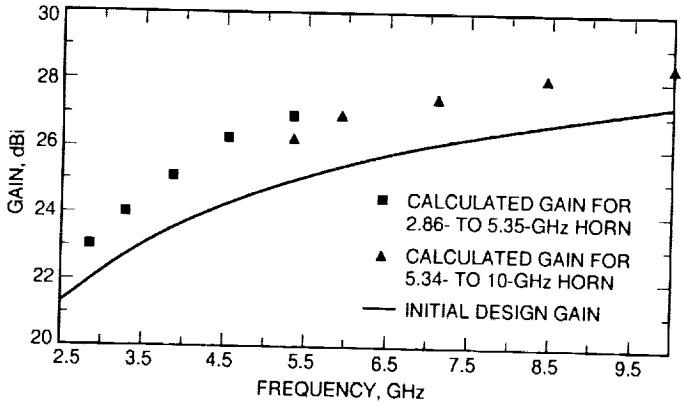


Fig. 5. Calculated horn gain.

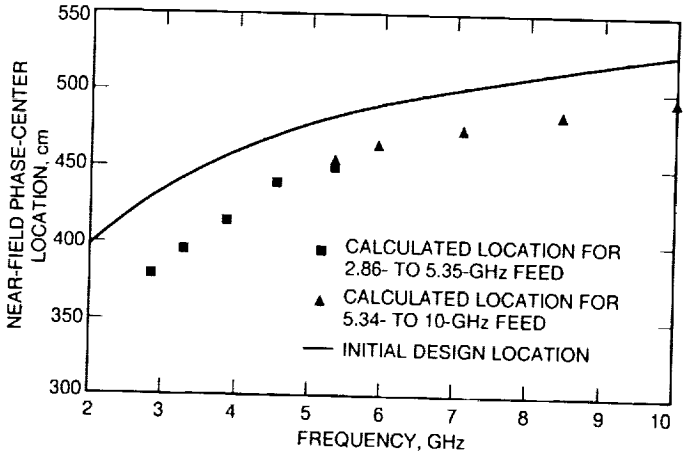


Fig. 6. Calculated near-field phase-center location.

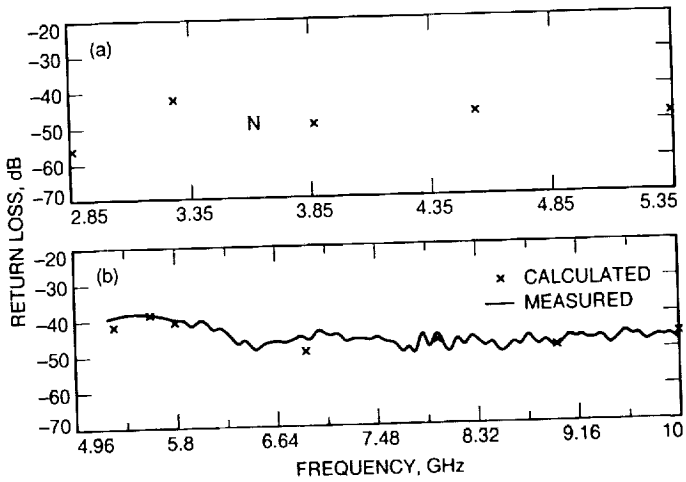


Fig. 7. Return loss of: (a) 2.86- to 5.35-GHz horn and (b) 5.34- to 10-GHz horn.

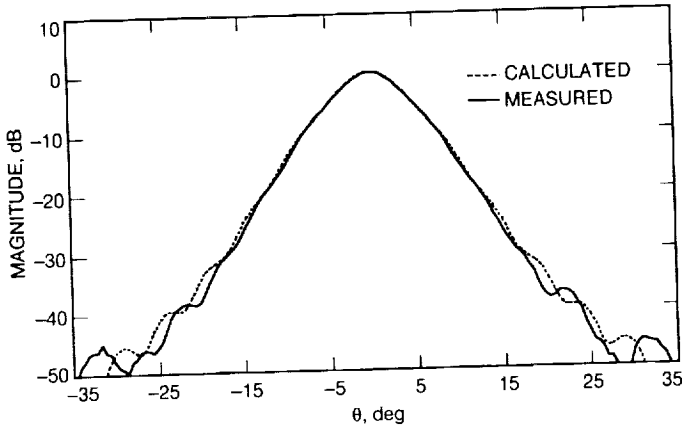


Fig. 8. Far-field copolarization radiation pattern at 7.1 GHz.

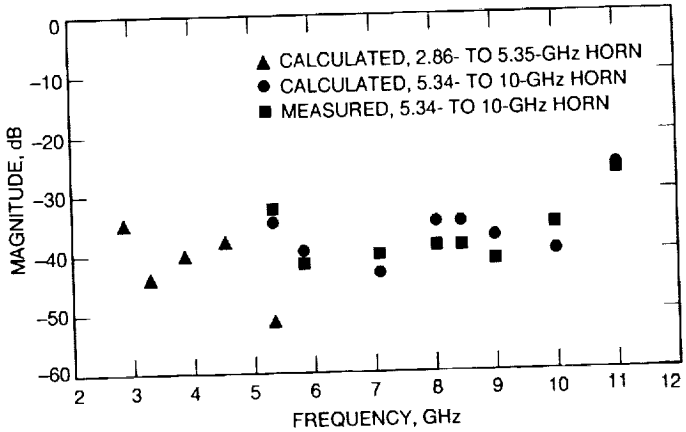


Fig. 9. Maximum cross-polarization of feeds relative to peak copolarization.

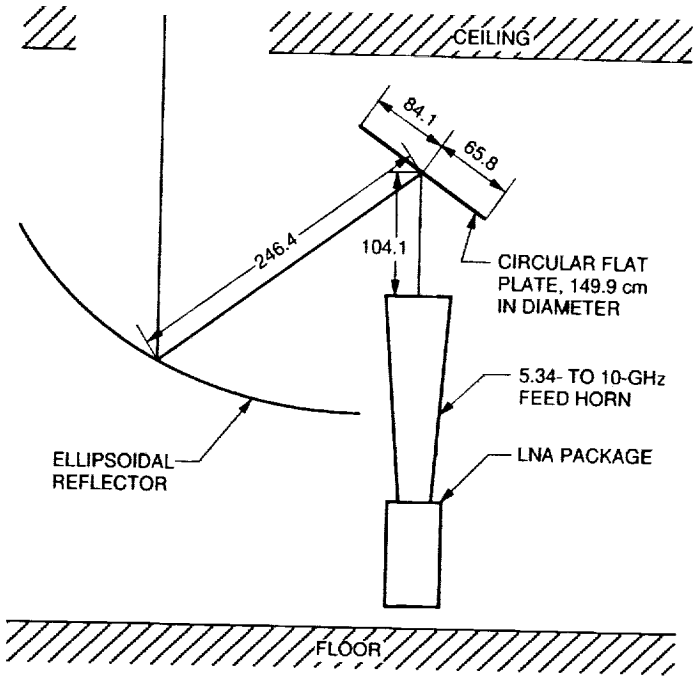


Fig. 10. Flat-plate design for 5.34- to 10-GHz horn.

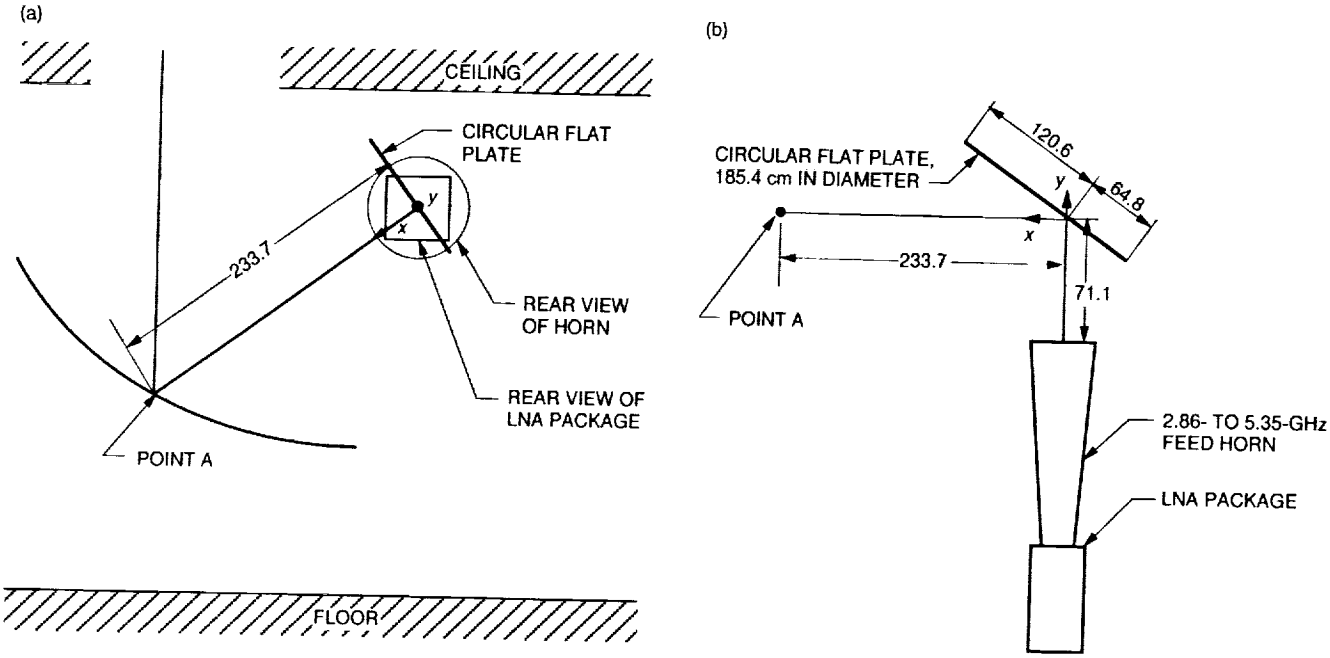


Fig. 11. Flat-plate design for 2.86- to 5.35-GHz horn: (a) rear view and (b) view in *x*, *y* plane.

The Flight Performance of the Galileo Orbiter USO[1]

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 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 se-

[1]Due to an error in the production of *The Telecommunications and Data Acquisition Progress Report 42-113*, this article is being reprinted in its entirety.

quence 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

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

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 spacecraft which introduced dynamic motion. Stator prepositioning 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.

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

[1] R. Woo, personal communication, Radio Frequency and Microwave Subsystems Section, Jet Propulsion Laboratory, Pasadena, California, August 1992.

[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.

(Fig. 26). An aging model was removed from the estimated spacecraft-transmitted frequencies for this period so that the resulting residuals could be analyzed and remaining error sources identified. The model removed was of a combination logarithmic curve and a linear aging drift:

$$\hat{f}_{s/c} = C_1 \log_e [C_2 (t - t_0) + 1] + C_3 + C_4 (t - t_0) \quad \text{for } t < t_b$$

$$\hat{f}_{s/c} = C_5 + C_6 (t - t_b) \quad \text{for } t \geq t_b$$

where t_0 is the time tag of the first data point, and t_b is the time tag of the first data point of where the linear aging realm begins. C_1 through C_6 are constant coefficients, which were estimated by least-squares analysis.

The initial attempt to fit this model over the data acquired during the first USO-on cycle resulted in the residuals displayed in Fig. 27. A large systematic variation is evident for passes occurring shortly after the USO was powered on, suggesting that the model is insufficient for this period. The rms scatter of the residuals is about 48 mHz with a χ_n^2 of 272. Since the behavior of the USO is known to be complex during this period and not easily modeled (initial stabilization phenomena mask the aging behavior), the first 12 passes were deleted, and the model was fit to the remaining 64 passes (see Fig. 28). The corresponding post-fit residual plot of Fig. 29 displays significantly less scatter and no significant systematic variation. The rms residual scatter of this fit was 17 mHz with a χ_n^2 of 35. It is believed that the 17-mHz variations are due to the random walk of the USO or mismodeling. In an attempt to identify any mismodeling, the post-fit residuals were examined against troposphere correction, spacecraft–Earth–Sun angle, elevation angle, spacecraft range, signal level (AGC), gravitational redshift correction, station identification, USO oven current (from telemetry), and spacecraft temperatures (from telemetry). No apparent correlations or trends were detected. The observed 17-mHz scatter is consistent with random walk behavior, as discussed previously (see Section VII.D).

The logarithmic time constant C_2^{-1} was estimated to be about 71 days. The turnover of the curve occurred 259 days after initial turn-on. The slope of the linear aging region was estimated to be -1.50×10^{-7} Hz/sec, which is in agreement with the measured slope of the Voyager 2 USO (-1.28×10^{-7} Hz/sec).[6] This translates to an aging rate over a 1-day period of -5.6×10^{-12}, which shows that the USO is drifting well below its specification of 5×10^{-11}/day.[7]

In an attempt to verify when the USO had reached its linear aging realm, different subsets of the last several passes in Fig. 28 were subjected to linear fits of frequency versus time. Initial fits of the last 31 passes prior to first turn-off (90-262 to 91-154), and the last 21 passes (90-360 to 91-154) displayed a signature in the residuals with significant curvature. A linear fit of the last 19 passes (91-006 to 91-154) showed no significant residual signature and yielded results consistent with the logarithmic and linear combination fit discussed previously. This implies that the Galileo USO required a little over a year of operation time prior to reaching its linear aging realm.

It is preferred that once the USO is switched on, it is left on so that it is allowed to reach the linear aging realm where it should remain for the duration of the Galileo activities. The linear aging realm allows one-way Radio Science experiments, such as the gravitational redshift experiment and planetary occultations, to be accurately calibrated. Since August 1991, the USO was switched off and back on in support of cooling turns as part of the effort to free the HGA stuck ribs (in conjunction with corresponding warming turns). The USO and its heater were powered off on August 5, 1991 (91-217). There were 6 passes conducted after the USO was powered back on August 16, 1991 (91-228). The estimated spacecraft-transmitted frequencies for these passes display a smooth continuity in time (see Fig. 30). A simple four-parameter aging model was fit to these frequencies yielding a post-fit rms scatter of 3 mHz. Although the number of points and the time period were insufficient to infer an accurate linear aging rate, this result suggests that the USO is behaving in a reasonable manner during this period. The USO was again turned off after this time period. If the USO had been allowed to remain on, it was expected that not as long of a period would be required for it to reach its linear aging realm as it did during the first on–off data set (89-341 to 91-154). The Allan deviations appear consistent with those of the first on–off data set, although two of the passes have somewhat higher values.

G. Retrace

If the USO and its heater are turned off, and then turned back on, the crystal will oscillate at a different frequency which is difficult to predict. This phenomenon

[6] S. W. Asmar, and P. M. Eshe, "Evaluation of the USO Performance - Final Report," JPL IOM Voyager-RSST-90-121 (internal document), Jet Propulsion Laboratory, Pasadena, California, January 17, 1990.

[7] A. Gussner, "Summary of Galileo USO Testing," JPL IOM 3364-80-080 (internal document), Jet Propulsion Laboratory, Pasadena, California, August 20, 1980.

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 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.

[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.

[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.

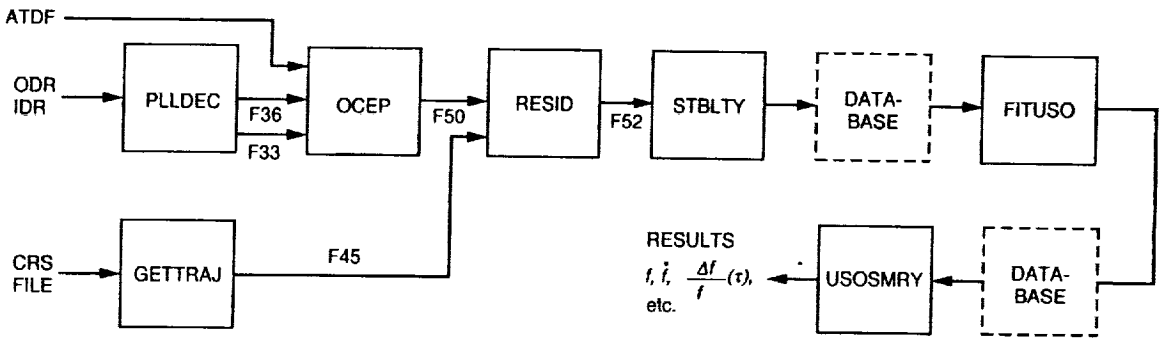


Fig. 1. STBLTY program-set block diagram.

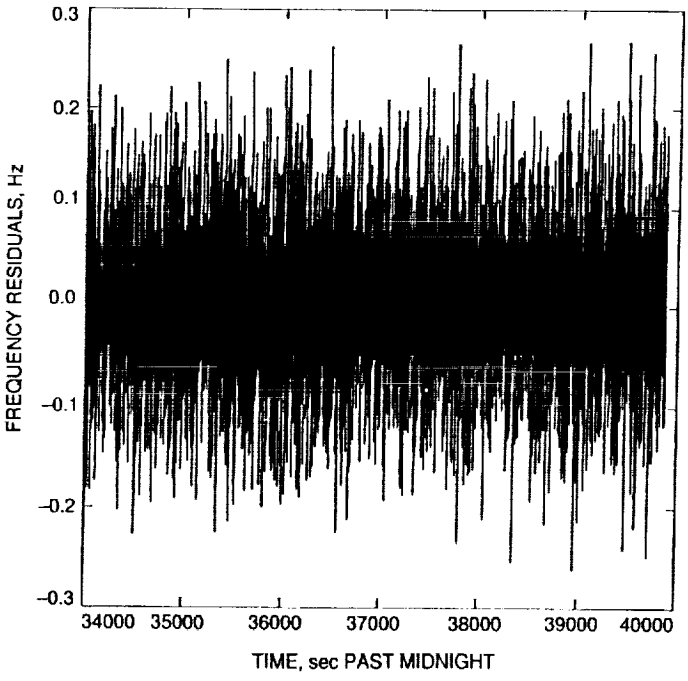


Fig. 2. Frequency residuals of sampled 1/sec Doppler for the USO pass of May 8, 1990.

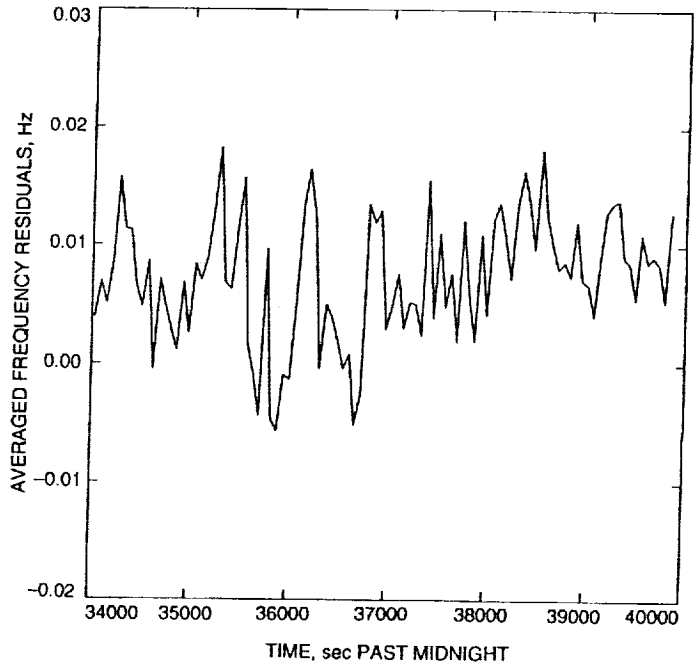


Fig. 3. Frequency residuals averaged every 60 sec for the USO pass of May 8, 1990.

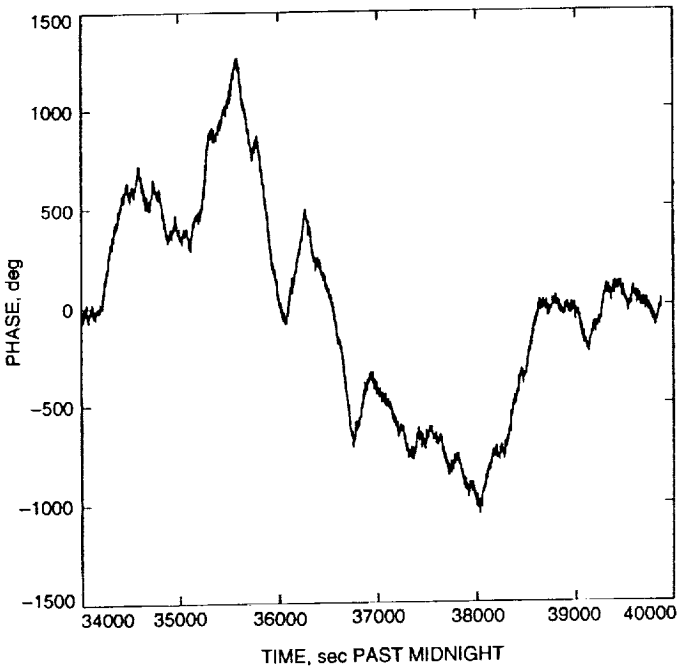


Fig. 4. Phase reconstructed from frequency residuals for the USO pass of May 8, 1990.

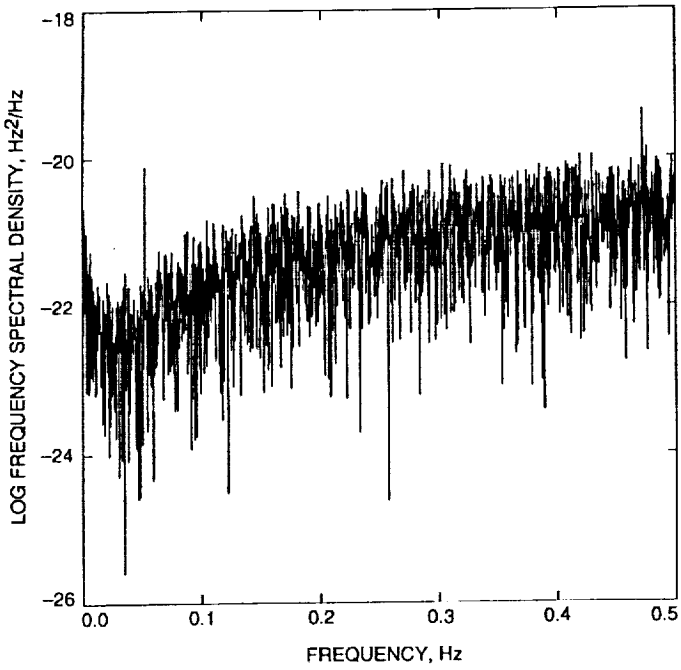


Fig. 6. Log of frequency spectral density of frequency residuals for the USO pass of May 8, 1990.

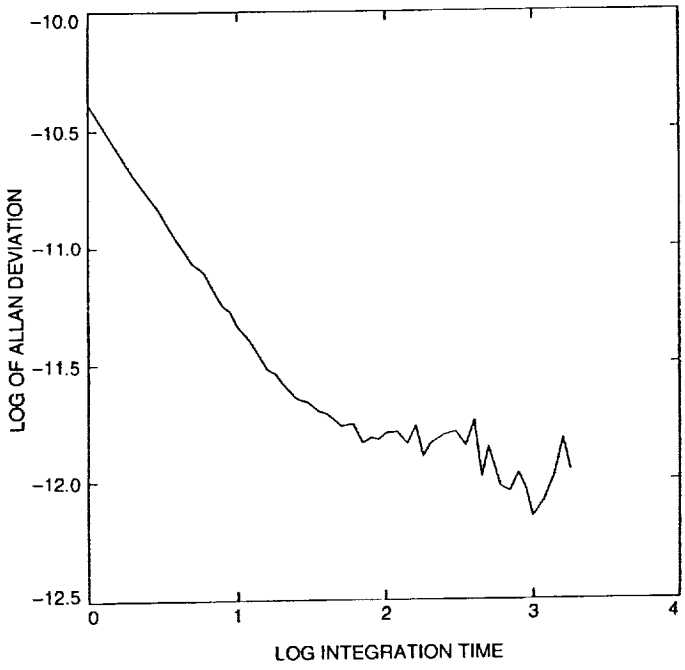


Fig. 5. Log of Allan deviation of frequency residuals for the USO pass of May 8, 1990.

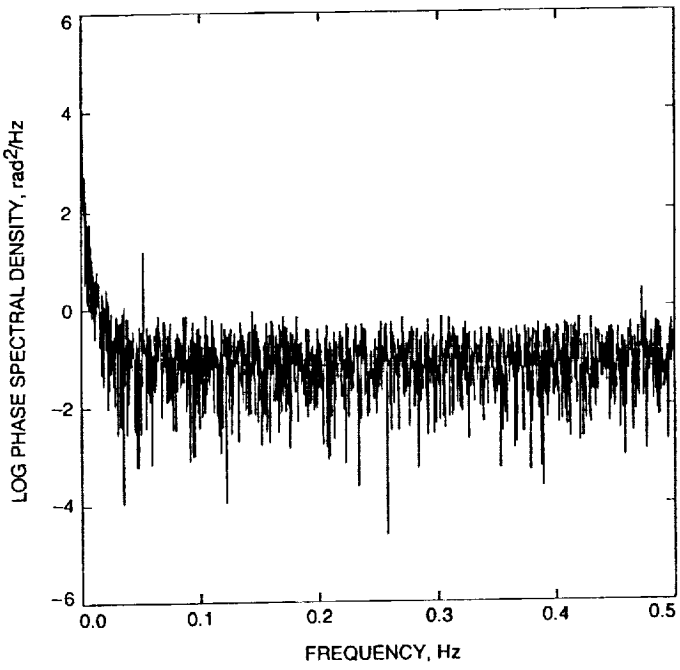


Fig. 7. Log of phase spectral density of frequency residuals for the USO pass of May 8, 1990.

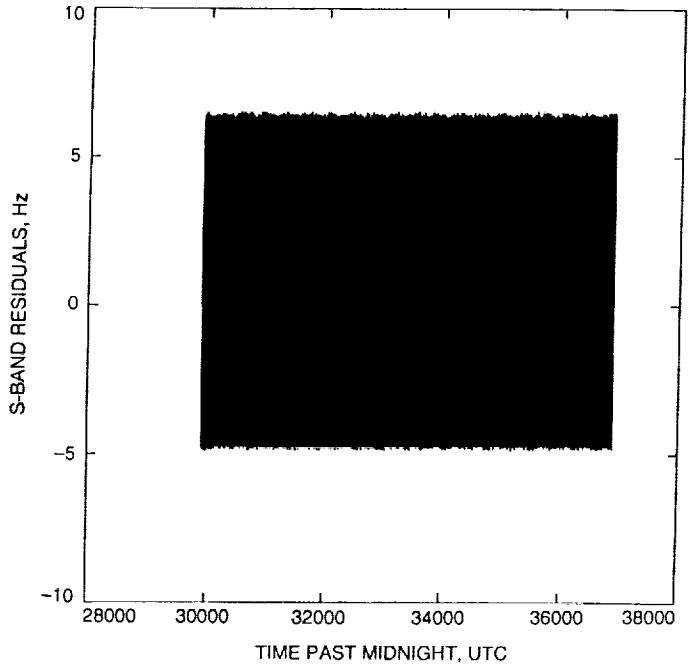


Fig. 8. Frequency residuals of sampled 1/sec Doppler for USO pass of December 16, 1990, where LGA-2 was the spacecraft antenna.

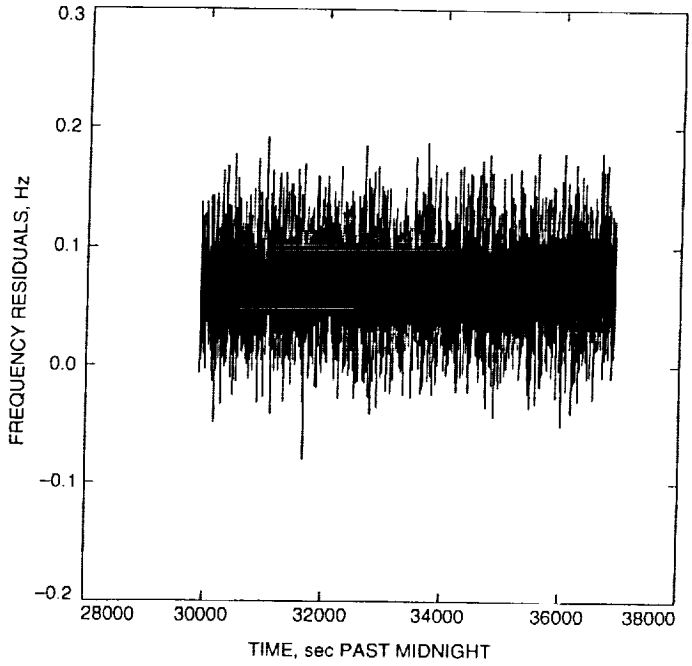


Fig. 10. Frequency residuals after removing the sinusoid fit from residuals displayed in Fig. 8 for USO pass of December 16, 1990, where LGA-2 was the spacecraft antenna.

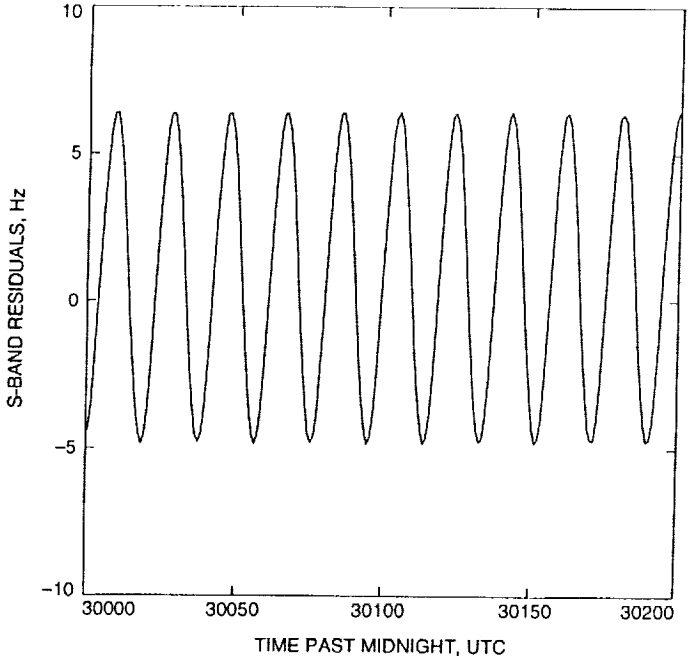


Fig. 9. Selected 200-sec period of frequency residuals of sampled 1/sec Doppler for USO pass of December 16, 1990, where LGA-2 was the spacecraft antenna.

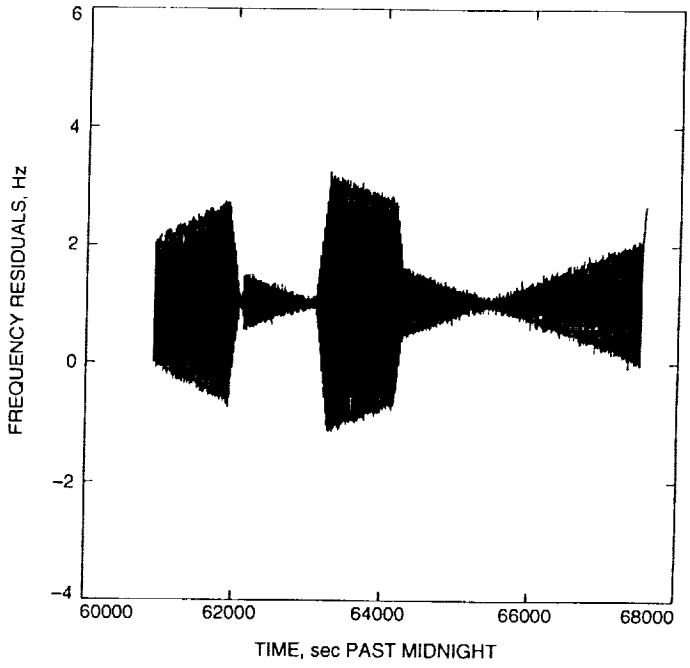


Fig. 11. Frequency residuals of sampled 1/sec Doppler for the USO pass of January 14, 1991, where LGA-2 was the signal source and dynamic motion occurred on board the spacecraft.

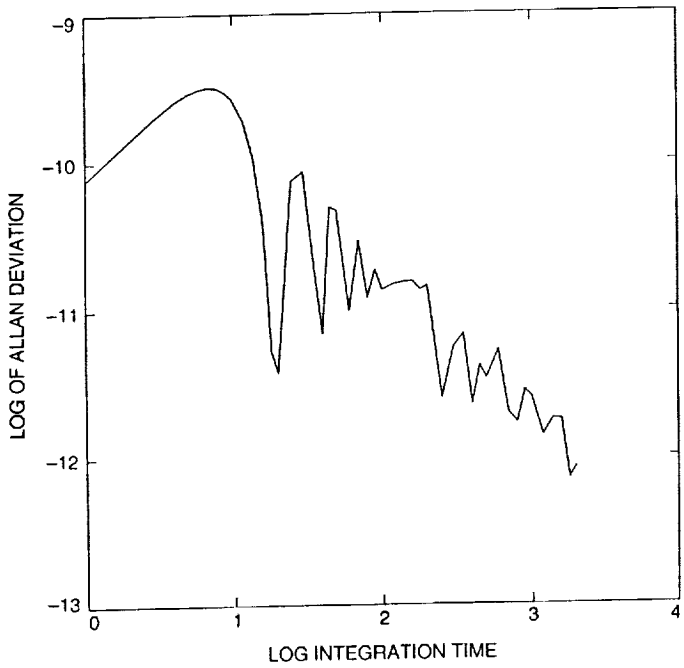


Fig. 12. Log of Allan deviation of frequency residuals of Fig. 11.

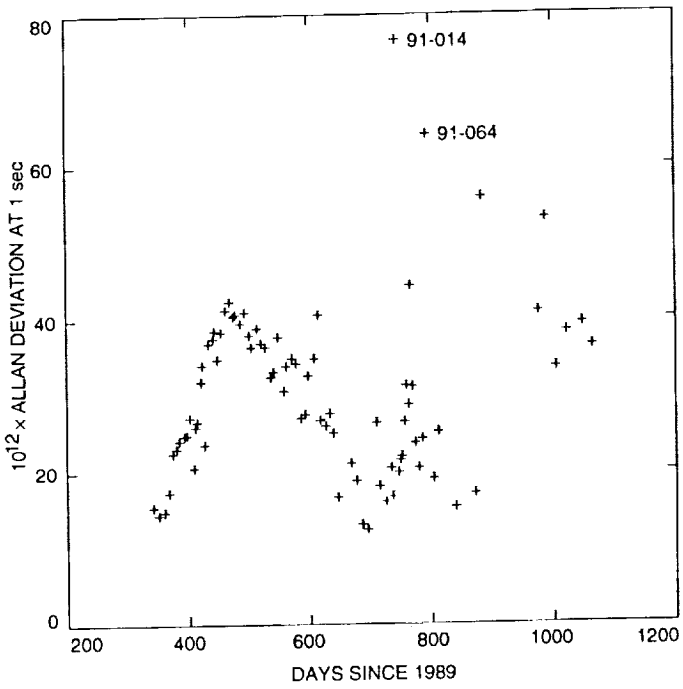


Fig. 14. Allan deviations at 1 sec for the 82 USO passes.

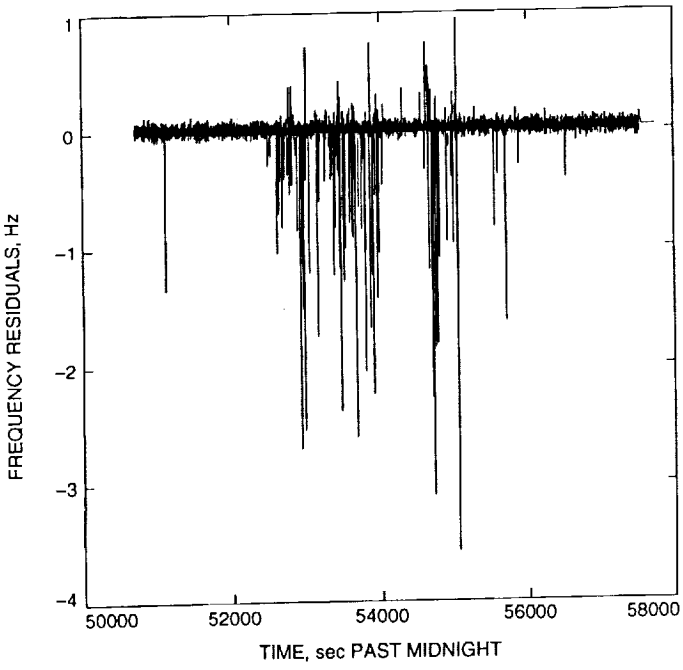


Fig. 13. Frequency residuals of the USO pass of March 5, 1991, where solar activity was known to have occurred.

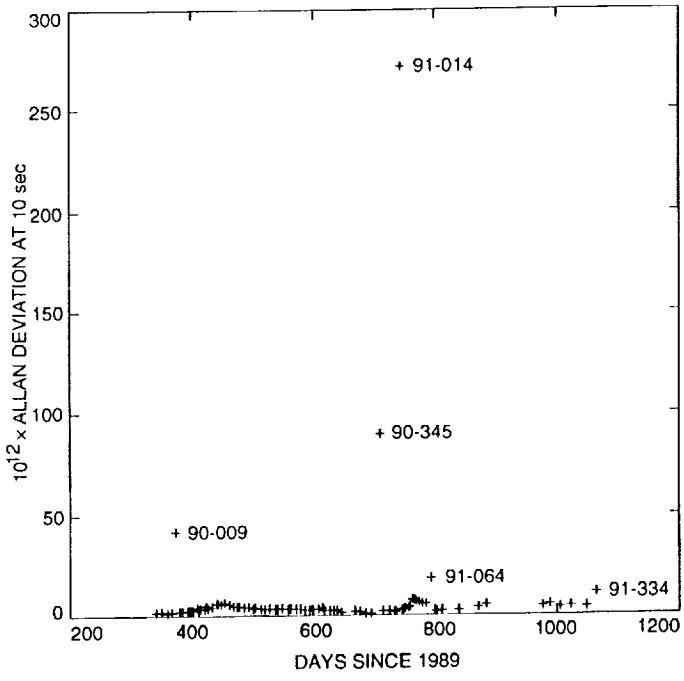


Fig. 15. Allan deviations at 10 sec for the 82 USO passes.

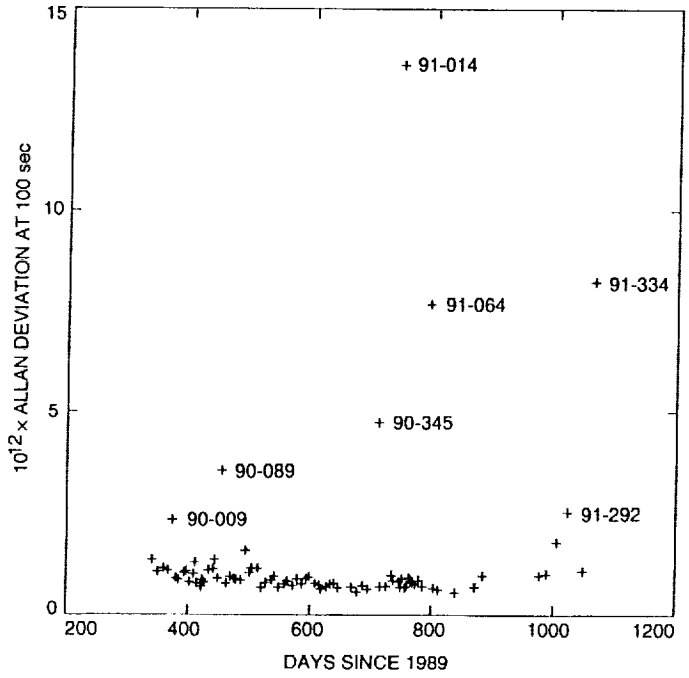


Fig. 16. Allan deviations at 100 sec for the 82 USO passes.

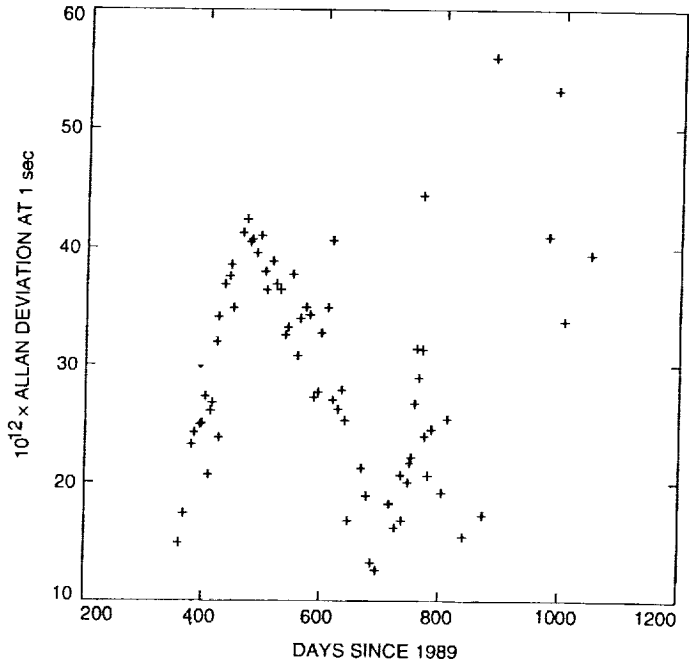


Fig. 18. Allan deviations at 1 sec for 73 USO passes (outliers identified in Figs. 14–17 have been removed).

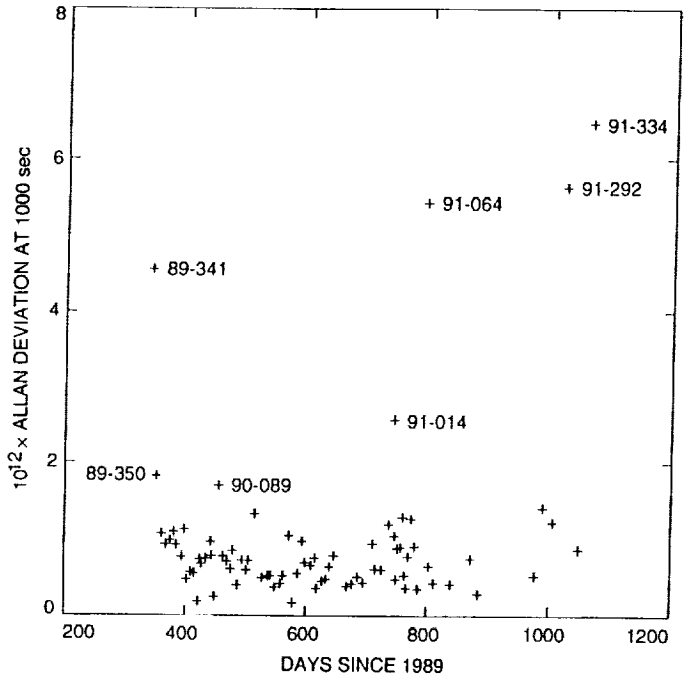


Fig. 17. Allan deviations at 1000 sec for the 79 USO passes.

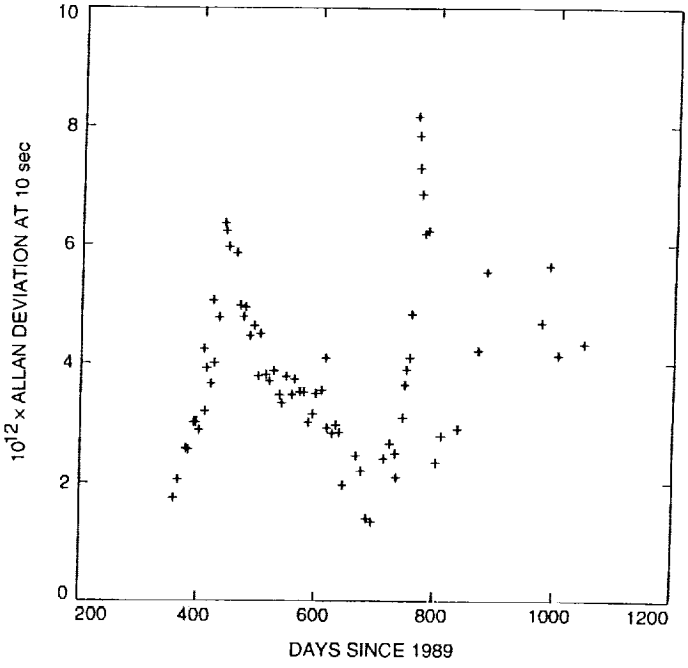


Fig. 19. Allan deviations at 10 sec for 73 USO passes (outliers identified in Figs. 14–17 have been removed).

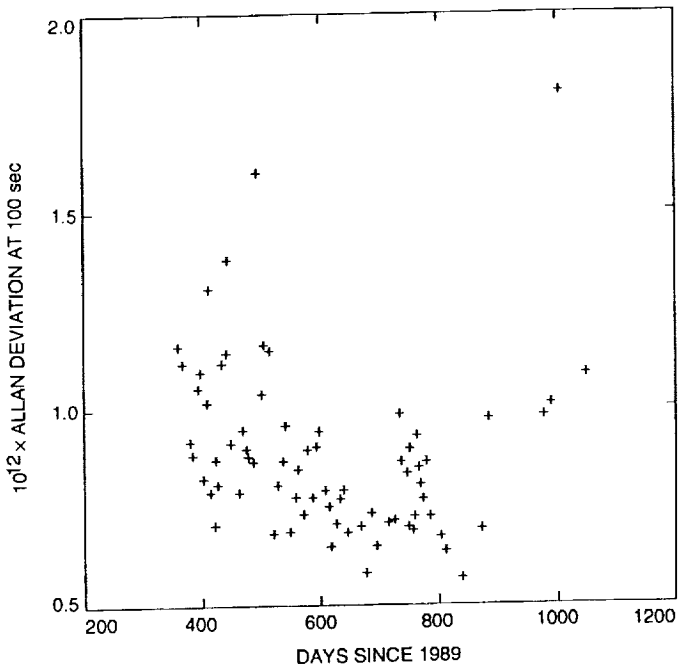


Fig. 20. Allan deviations at 100 sec for 73 USO passes (outliers identified in Figs. 14–17 have been removed).

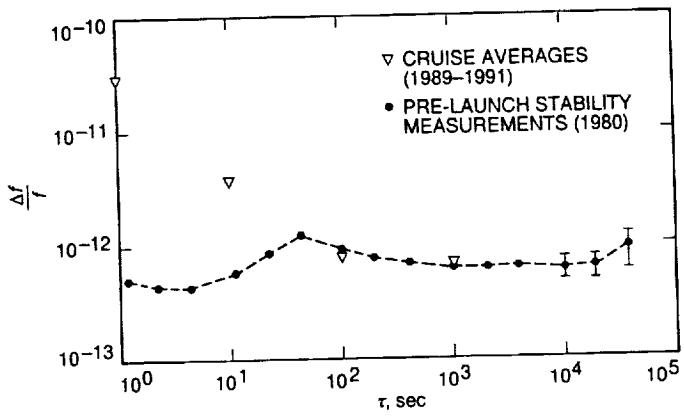


Fig. 22. In-flight USO-pass Allan-deviation measurement averages for 1, 10, 100 and 1000 sec superimposed with pre-flight Allan deviation measurements.

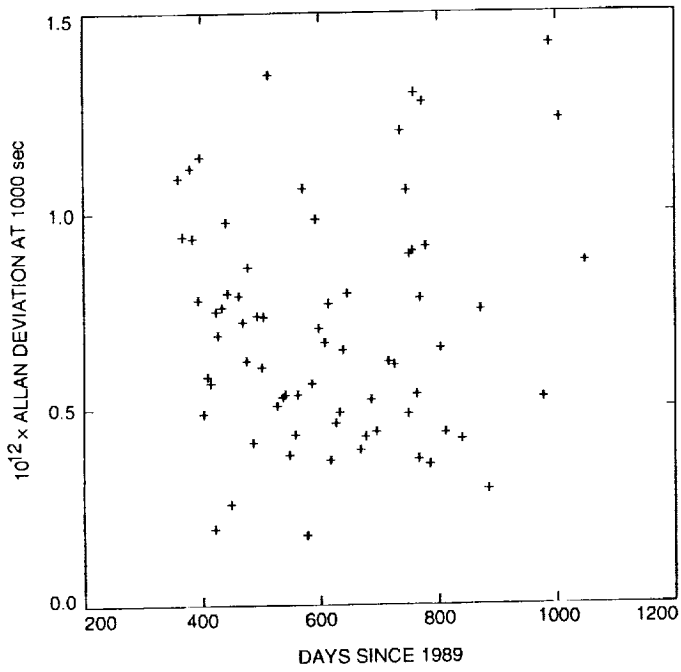


Fig. 21. Allan deviations at 1000 sec for 70 USO passes (outliers identified in Figs. 14–17 have been removed).

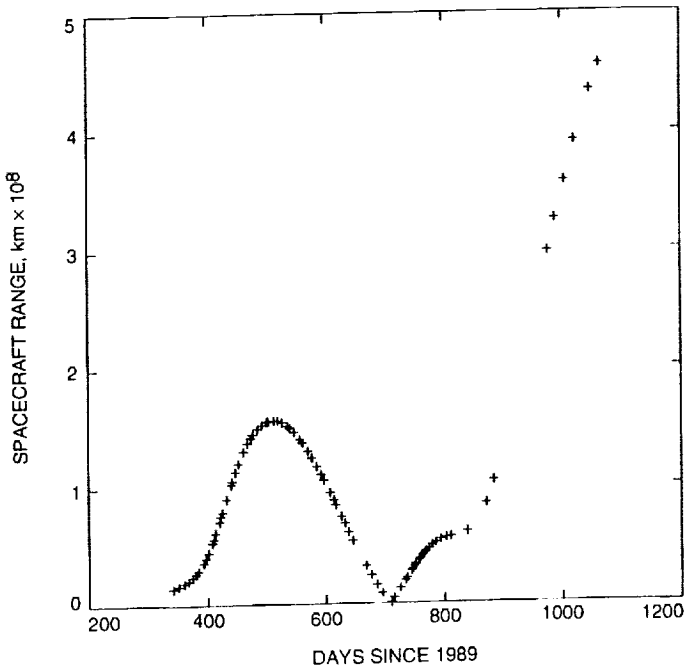


Fig. 23. Spacecraft range for each USO pass. Launch occurred at day 291; the dip at 707 days after 1989.0 was the Earth 1 flyby.

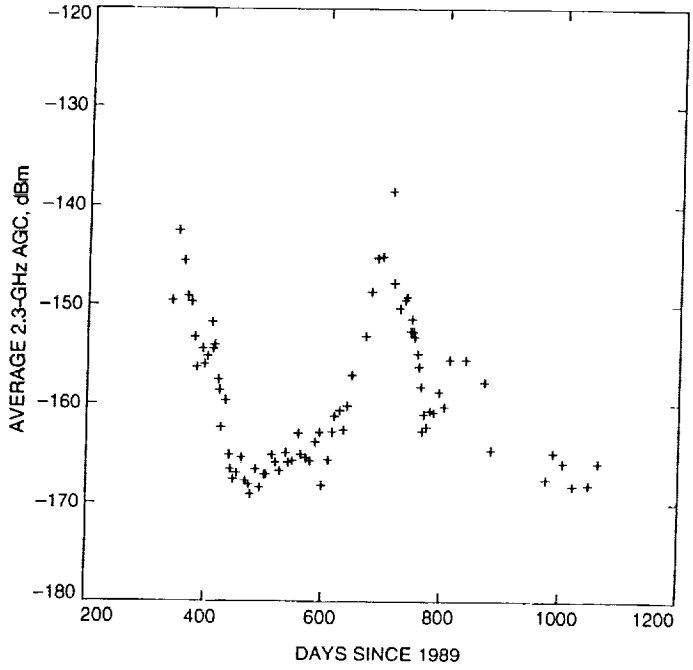


Fig. 24. Signal strength (AGC level) measured for each USO pass.

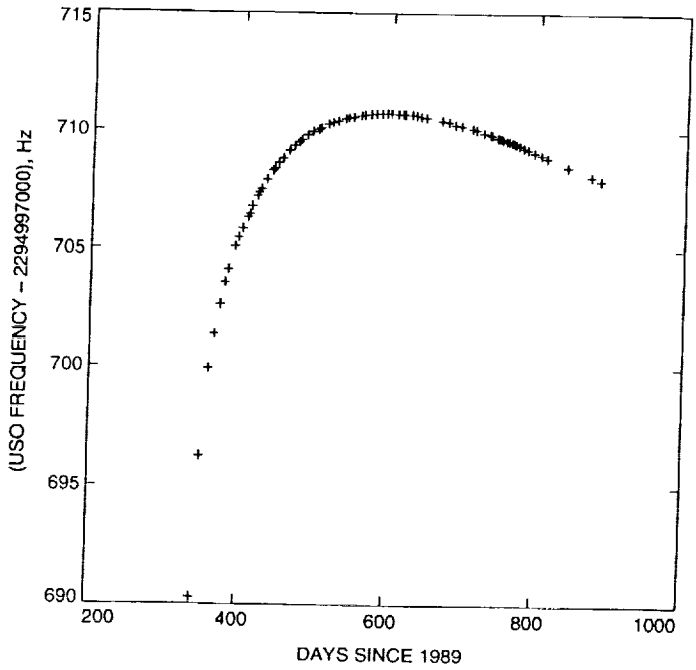


Fig. 26. Estimated spacecraft transmitted frequencies for 76 USO passes during the first USO on–off cycle after launch.

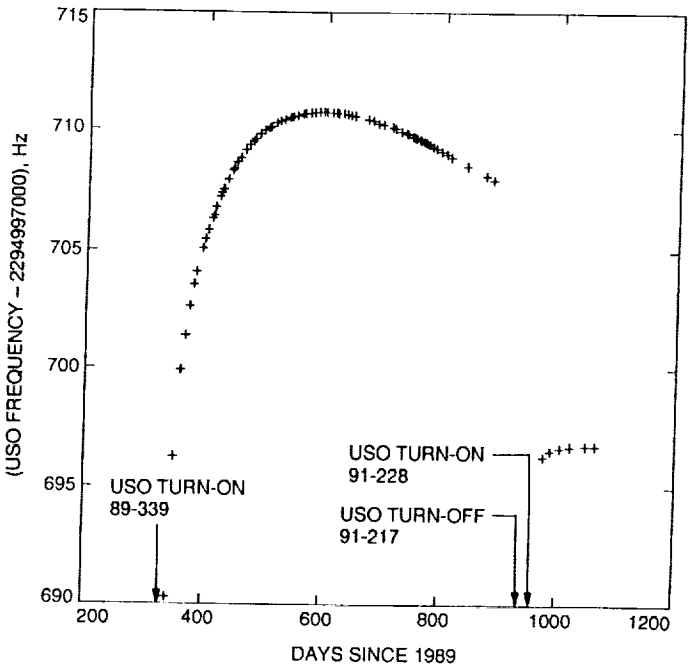


Fig. 25. Estimated spacecraft transmitted frequencies for all 82 USO passes as determined by STBLTY.

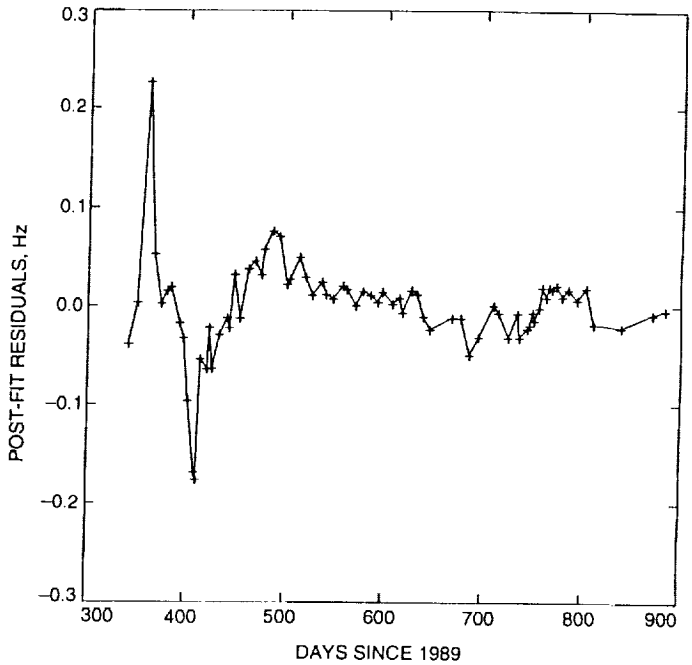


Fig. 27. Post-fit residuals of estimated spacecraft transmitted frequencies of Fig. 26 after fitting and removing an aging model.

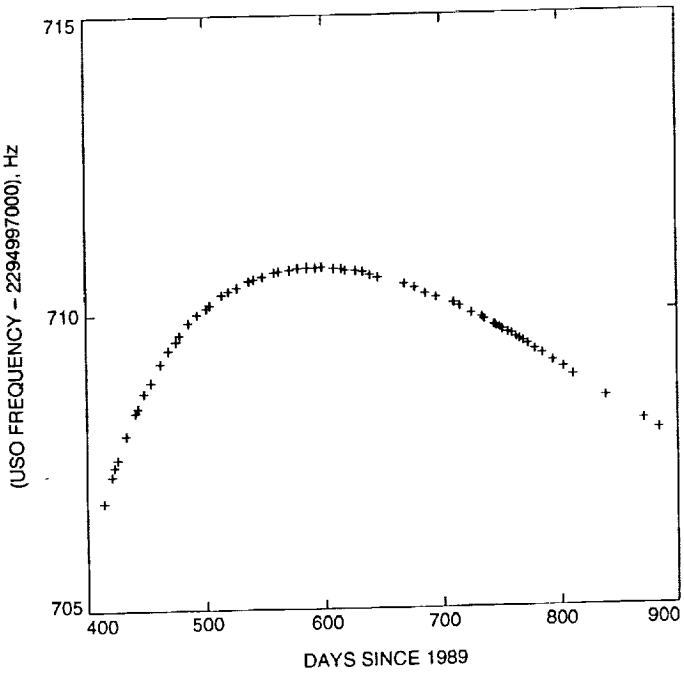


Fig. 28. Estimated spacecraft transmitted frequencies for 64 USO passes during the first USO on–off cycle after launch (the first 12 passes of Fig. 26 have been removed).

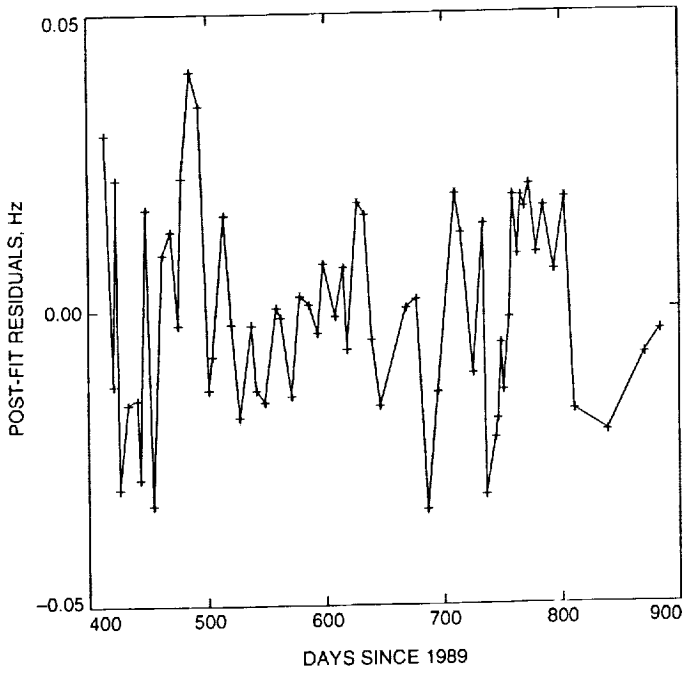


Fig. 29. Post-fit residuals of estimated spacecraft transmitted frequencies of Fig. 28 after fitting and removing an aging model.

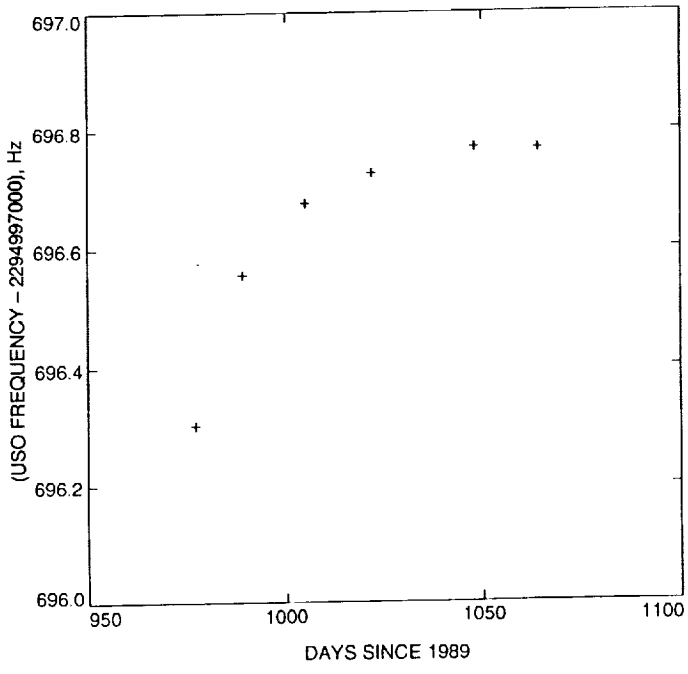


Fig. 30. Estimated spacecraft transmitted frequencies of 6 USO passes conducted after the first USO on–off cycle.

OMIT 1D END

Errata

J. Z. Wilcox has submitted the following errata to [1].

Errors due to averaging tropospheric fluctuations over nonzero water-vapor radiometer beamwidths were presented. An error in the numerical integration code caused reported values to be too large by a factor of 2–10 (the factor depends on the elevation angle and the beamwidth). Corrected values are included in [2].

References

[1] J. Z. Wilcox, "The Effect of Tropospheric Fluctuations on the Accuracy of Water Vapor Radiometry," *The Telecommunications Data Acquisition Progress Report 42-110, vol. April–June 1992*, Jet Propulsion Laboratory, Pasadena, California, pp. 33–51, August 15, 1992.

[2] R. P. Linfield and J. Z. Wilcox, "Radio Metric Errors Due to Mismatch and Offset Between a DSN Antenna Beam and the Beam of a Troposphere Calibration Instrument," *The Telecommunications Data Acquisition Progress Report 42-114, vol. April–June 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–13, August 15, 1993.

Referees

The following people have refereed articles for *The Telecommunications and Data Acquisition Progress Report*. By attesting to the technical and archival value of the articles, they have helped to maintain the excellence of this publication during the past year.

D. A. Bathker
M. Britcliffe
T. Cwik
M. S. Esquivel
F. Hadaegh
R. L. Horttor
W. A. Imbriale
P. W. Kinman
K. M. Liewer
A. Nikora
F. Pollara
P. M. Richter
E. H. Satorius
P. Stanton
J. S. Ulvestad
G. E. Wood
C. S. Yung
M. Belongie
K.-M. Cheung
S. Dolinar
Y. Feria
H. Hemmati
W. Hurd
D. L. Jones
J. H. Kwok
L. Maleki
G. K. Noreen
V. M. Pollmeier
A. A. Riewe
M. Shao
R. L. Sydnor
V. A. Vilnrotter
S. C. Wu
C. Borden
P. W. Chodas
L. Ekroot
W. Gawronski
S. M. Hinedi
W. J. Hyland
J. R. Kahr
J. R. Lesh
R. J. McEliece
J. Peterson
E. C. Posner
E. Rodemich
J. S. Shell
P. J. Tallerico
M. Wert
T.-Y. Yan
D. Boussalis
R. C. Clauss
J. Ellis
J. A. Gleason
D. J. Hoppe
R. Ibanee-Meier
S. Kayalar
B. M. Levine
A. Mileant
S. M. Petty
G. Resch
B. P. Saldua
S. D. Slobin
S. W. Thurman
K. E. Wilson
L. E. Young

END DATE NOV 17, 1993.