

N7811132
[Barcode]

The Deep Space Network Progress Report 42-41

July and August 1977

(NASA-CR-155225) THE DEEP SPACE NETWORK
Progress Report, Jul. - Aug. 1977 (Jet
Propulsion Lab.) 237 p HC A11/MF A01

CSSL 22A

G3/12

N78-11132
THRU
N78-11156
Unclas
52880

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91103

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL
INFORMATION SERVICE
SPRINGFIELD, VA 22161



The Deep Space Network Progress Report 42-41

July and August 1977

October 15, 1977

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91103

Preface

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

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

Description of the DSN

Mission Support

- Ongoing Planetary/Interplanetary Flight Projects
- Advanced Flight Projects

Radio Astronomy

Special Projects

Supporting Research and Technology

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

Network and Facility Engineering and Implementation

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

Operations

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

Program Planning

- TDA Planning

In each issue, the part entitled “Description of the DSN” describes the functions and facilities of the DSN and may report the current configuration of one of the five DSN systems (Tracking, Telemetry, Command, Monitor & Control, and Test & Training).

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

Contents

DESCRIPTION OF THE DSN

Network Functions and Facilities	1
N. A. Renzetti	
The DSN Programming System	4
W. D. Hodgson	
NASA Code 311-03-41-15	
DSN Monitor and Control System, Mark III-77.	10
W. G. Stinnett	
NASA Code 311-03-43-10	

MISSION SUPPORT

Ongoing Planetary/Interplanetary Flight Projects

Viking Extended Mission Support	16
D. W. H. Johnston	
NASA Code 311-03-21-70	
Results of Pioneer – Venus 1978 Sequential Decoding Tests Over a Simulated Lognormal Fading Channel	21
B. K. Levitt	
NASA Code 825-10-21-22	
Pioneer Mission Support	33
T. P. Adamski	
NASA Code 311-03-21-90	
Helios Mission Support	39
P. S. Goodwin and G. M. Rockwell	
NASA Code 311-03-21-50	

SUPPORTING RESEARCH AND TECHNOLOGY

Tracking and Ground-Based Navigation

Near Sun Ranging	43
A. I. Zyguelbaum	
NASA Code 310-10-61-16	
Feasibility Study of Far-Field Methods for Calibrating Ground Station Delays: An Interim Report	51
T. Sato	
NASA Code 310-10-61-08	

PRECEDING PAGE BLANK NOT FILMED

A Technique to Determine Uplink Range Calibration Due to Charged Particles	57
S. C. Wu and F. B. Winn NASA Code 310-10-60-50	

Communications – Spacecraft/Ground

Doppler Profiles of Radio Frequency Interference	82
B. K. Levitt NASA Code 310-20-66-07	

Sampled-Data System Analysis of Antenna Conical-Scan Tracking	86
R. A. Winkelstein NASA Code 310-20-67-03	

Station Control and Operations Technology

Antenna Pointing Subsystem Drive Tape Generator – Another New Interface	101
R. F. Emerson NASA Code 310-30-69-10	

Network Control and Data Processing

An MBASIC™ User Profile	103
R. L. Schwartz NASA Code 310-40-72-05	
CRISPFLOW, A Structured Program Design Flowcharter	112
R. C. Tausworthe and K. C. Landon NASA Code 310-40-72-05	

**NETWORK AND FACILITY ENGINEERING
AND IMPLEMENTATION**

Network

The CTA 21 Radio Science Subsystem – Non-Real-Time Bandwidth Reduction of Wideband Radio Science Data	127
A. L. Berman and J. I. Molinder NASA Code 311-03-43-10	
Electron Density in the Extended Corona – Two Views	135
A. L. Berman NASA Code 311-03-43-10	

S-/X-Band Microwave Optics Design and Analysis for DSN 34 Meter Diameter Antenna	146
D. L. Nixon and D. A. Bathker	
NASA Code 311-03-42-59	
Configuration Control and Audit Assembly	166
A. P. Irvine	
NASA Code 311-03-41-17	
Microcomputer Central Processing Unit Module	172
O. B. Parham	
NASA Code 311-03-41-05	
Fast Algorithms for Computing Mersenne- Prime Number-Theoretic Transforms	176
I. S. Reed and T. K. Truong	
NASA Code 311-03-42-95	

Ground Communications

GCF Central Performance Monitor — Functional Description	206
N. C. S. Lau	
NASA Code 311-06-30-00	

Deep Space Stations

A Thermodynamic Analysis of a Solar-Powered Jet Refrigeration System	209
F. L. Lansing and V. W. Chai	
NASA Code 311-03-44-08	
Estimated Displacements for the VLBI Reference Point of the DSS 14 Antenna	218
H. McGinness	
NASA Code 311-03-42-91	

PROGRAM PLANNING

TDA Planning

Dynamic Modeling and Sensitivity Analysis of Solar Thermal Energy Conversion Systems	226
C. L. Hamilton	
NASA Code 311-03-31-30	

D/
N78-11133

Network Functions and Facilities

N. A. Renzetti

Office of Tracking and Data Acquisition

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

The Deep Space Network (DSN), established by the National Aeronautics and Space Administration (NASA) Office of Tracking and Data Acquisition under the system management and technical direction of the Jet Propulsion Laboratory (JPL), is designed for two-way communications with unmanned spacecraft traveling approximately 16,000 km (10,000 miles) from Earth to the farthest planets of our solar system. It has provided tracking and data acquisition support for the following NASA deep space exploration projects: Ranger, Surveyor, Mariner Venus 1962, Mariner Mars 1964, Mariner Venus 1967, Mariner Mars 1969, Mariner Mars 1971, and Mariner Venus Mercury 1973, for which JPL has been responsible for the project management, the development of the spacecraft, and the conduct of mission operations; Lunar Orbiter, for which the Langley Research Center carried out the project management, spacecraft development, and conduct of mission operations; Pioneer, for which Ames Research Center carried out the project management, spacecraft development, and conduct of mission operations; and Apollo, for which the Lyndon B. Johnson Space Center was the project center and

the Deep Space Network supplemented the Manned Space Flight Network (MSFN), which was managed by the Goddard Space Flight Center (GSFC). It is providing tracking and data acquisition support for Helios, a joint U.S./West German project; and Viking, for which Langley Research Center provides the project management, the Lander spacecraft, and conducts mission operations, and for which JPL also provides the Orbiter spacecraft.

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

From its inception, NASA has had the objective of conducting scientific investigations throughout the solar sys-

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

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

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

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

- The DSN Telemetry System receives, decodes, records, and retransmits engineering and scientific data generated in the spacecraft to Mission Control.
- The DSN Command System accepts coded signals from Mission Control via the Ground Communications Facility and transmits them to the spacecraft in order to initiate spacecraft functions in flight.

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

- The DSN Monitor and Control System instruments, transmits, records, and displays those parameters of the DSN necessary to verify configuration and validate the Network. It provides operational direction and configuration control of the Network, and provides primary interface with flight project Mission Control personnel.
- The DSN Test and Training System generates and controls simulated data to support development, test, training and fault isolation within the DSN. It participates in mission simulation with flight projects.

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

- (1) The Deep Space Stations, which are distributed around Earth and which, prior to 1964, formed part of the Deep Space Instrumentation Facility. The technology involved in equipping these stations is strongly related to the state of the art of telecommunications and flight-ground design considerations, and is almost completely multimission in character.
- (2) The Ground Communications Facility provides the capability required for the transmission, reception, and monitoring of Earth-based, point-to-point communications between the stations and the Network Operations Control Center at JPL, Pasadena, and to the mission operations centers, wherever they may be. Four communications disciplines are provided: teletype, voice, high-speed, and wideband. The Ground Communications Facility uses the capabilities provided by common carriers throughout the world, engineered into an integrated system by Goddard Space Flight Center, and controlled from the communications Center located in the Space Flight Operations Facility (Building 230) at JPL.

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

- Control and coordination of Network support to meet commitments to Network users.
- Utilization of the Network data processing computing capability to generate all standards and limits required for Network operations.
- Utilization of Network data processing computing capability to analyze and validate the performance of all Network systems.

The personnel who carry out the above functions are located in the Space Flight Operations Facility, where mission operations functions are carried out by certain flight projects. Network personnel are directed

by an Operations Control Chief. The functions of the Network Data Processing are:

- Processing of data used by Network Operations Control for control and analysis of the Network.
- Display in the Network Operations Control Area of data processed in the Network Data Processing Area.
- Interface with communications circuits for input to and output from the Network Data Processing Area.
- Data logging and production of the intermediate data records.

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

D₂
N78-11134

The DSN Programming System

W. D. Hodgson
TDA Engineering Section

This article describes the DSN Programming System. The Programming System provides the DSN with implementation tools and practices for producing DSN software. This article provides a general description of the System, as well as its key characteristics, status, plans, and expected benefits.

I. Introduction

The DSN Programming System is a unified body of practices and implementation tools for producing DSN software on time, within budget, conforming correctly to functional requirements, and of predictable and low life-cycle cost. The system incorporates software methodology, software standard practices, language and data base standards, including standard languages, implementation aids, and management aids.

The DSN Programming System is needed because:

- (1) DSN service continues to require more software
 - (a) End user requirements continue to increase, both in quantity and quality.
 - (b) Network functions need to be computerized to stabilize maintenance and operating costs.
- (2) DSN software costs are continuing to be an increasing fraction of the DSN budget. Transferred software in the DSN is increasing at 15% per year.
- (3) The DSN customer, the end user project, requires that DSN software be delivered on time and with full nego-

tiated capability. The Programming System provides assurances, as well as the means for the DSN to meet the end user requirements.

II. Key Characteristics

The DSN Programming System includes the production and maintenance of DSN Software Standard Practice documents, the implementation of DSN standard programming languages according to pre-defined functional requirements, and the implementation and evaluation of the initial concept of the distributed DSN data base (in the Configuration Control and Audit Assembly), as well as software implementation aids and management aids.

There are six complete DSN Standard Practice documents¹ for software that cover implementation including documentation; Software Implementation Guidelines and Practices, Preparation of Software Requirements Documents (SRD), Preparation of Software Definition Documents (SDD), Preparation of Software Specification Documents (SSD), Preparation of Software Operators Manuals (SOM), and Preparation of Software Test and Transfer Documents (STT).

¹These documents are currently available on request from the author.

A seventh Standard Practice is currently in development that will cover the DSN data base, including implementation guidelines and practices.

The Software Implementation Guidelines and Practices describe the overall Tracking and Data Acquisition (TDA) software methodology, policy, and software management plan. The other Standard Practices provide supporting detail as summarized below.

The Standard Practices on Software Requirements Documents and Software Definition Documents cover the early activities of requirements identification and software architectural design. These are the conceptual phases where decisions made during these formative stages tend to have profound effects on the overall costs, commitments, and the general approach. The SRD and SDD de-emphasize formality and permanency. The SDD describes the software architectural design, which permits a $\pm 10\%$ accurate estimate to be made of the cost and schedule of the remaining implementation, which is the program construction. The architectural design can, in fact, be the basis for a bid by an outside contractor for a fixed price implementation. Timing is critical and early reviews facilitate redirection, if needed. Also, informal documentation during these phases facilitates changes as may be needed prior to final approval. The documents covering the preparation of SSDs, SOMs, and STTs require that they be produced concurrently with the program construction and testing activities. They document the true "as-built" and "as-tested" computer program and are used to operate and maintain the transferred (delivered) program. These documents tend to be more formal and are updated, as needed throughout the useful life of the software. The standard practice on data bases is to provide guidelines for standardizing and implementing the operational DSN Data Base. The DSN Data Base includes the data throughout the DSN in continual use for ongoing business, operations, and commitments.

The standard programming languages reduce implementation and sustaining costs because they are both standard and high-level languages. They permit implementers to concentrate on TDA design problems, not on machine details. Furthermore, having standard languages preserves implementer and manager skills and furnishes immunity of DSN applications programs to hardware and operating system changes. The machine-independent design principles being applied reduce the costs of providing the standard languages on various computers the DSN uses. Two standard languages have been identified: a non-real-time language for Operations management use, augmented by a File Editor for data base modification and update, and a real-time language.

The non-real-time language machine-independent design is being completed and implemented on the PDP-10, on the

Univac 1108, and on the DSN standard minicomputer (Ref. 1). The File Editor functional requirements are being defined in preparation for the machine-independent design and subsequent implementation in the Configuration Control and Audit Assembly (Ref. 2). In later years, the real-time DSN standard language will be similarly implemented, initially for the DSN standard minicomputer, and then for DSN Control and Computation Modules (CCMs), the DSN standard microprocessor components.

The implementation aids are to be selected and then implemented, based on the methodology research coordinated under the DSN Advanced Systems program and on the experiences of the Mark III DSN Data Subsystem Implementation Project (MDS). These may include automated finished graphics, computer-based implementation document construction, a program design medium, and perhaps others. The Software Standard Practices will be revised to incorporate the proper use of these aids as they are transferred to Operations, acting in unison with the methodology and languages.

The DSN Programming System is organized in the following structure. (Figure 1 also gives a graphic representation of the same information for clarity.)

- (1) Methodology research
- (2) DSN non-real time language
 - (a) MBASIC² machine independent design (MID)
 - (b) MBASIC² MID extensions (MIDX)
 - (c) MBASIC² batch compiler
 - (d) MBASIC² reference manual
 - (e) MBASIC² implementations
 - (f) File editor
- (3) DSN real-time standard language
 - (a) Design
 - (b) Implementation
- (4) Data base implementation and standards
- (5) Implementation aids and management aids
- (6) Software standard practices
 - (a) Guidelines and practices
 - (b) Software Requirements Document preparation
 - (c) Software Definition Document preparation
 - (d) Software Specification Document preparation
 - (e) Operators Manual preparation

²A trademark of the California Institute of Technology.

- (f) Test and Transfer Document preparation
- (g) DSN data base implementation guidelines and practices

III. Schedule Overview

The DSN Programming System schedule of major deliverable milestones is described in the following paragraphs.

The first six software standard practices have been completed since early FY77. The seventh (Data Base Implementation Guidelines and Practices) will be complete in early FY78.

The Machine Independent Design (MID) of the non-real-time standard language is also complete with one implementation (on the DEC System-10 computer) transferred to operations and in use.

The first volume of the software methodology textbook (Ref. 3) has been published and is available from (among other sources) Prentice-Hall. The second volume is drafted and will be published in early FY78.

It is planned that the non-real-time standard language (MBASIC) interpreter and batch compiler is to be complete and operational on three different computers (MODCOMP II, Univac 1108, and the DEC System-10) in late FY80, followed shortly thereafter by a standard file editor implementation on the same three machines:

The standard real-time programming language will be complete in FY80 for the DSN standard minicomputer, followed

one year later by implementation for the DSN standard micro-processor components (CCMs).

An overview of the Programming System schedule is shown in Fig. 2.

IV. Benefits

Many benefits are expected or are currently being realized because of the application of the Programming System. Table 1 shows in what way each element of the Programming System provides benefits. Some of the benefits can only be predicted based on current software literature (since the Programming System tool is not yet available) but many have actually been measured. For example, for the implementation of the MBASIC language (itself a test bed for the Programming System), costs can be directly compared to the costs of a previous implementation which did not use the Programming System tools. The comparison in this case showed that a machine-independent design was obtained at no additional cost over the old way of machine-dependent design.

V. Summary

All the elements of the DSN Programming System have been defined. Those elements that exist now are being used in Implementation and Operations. The effects on software life-cycle costs and productivity are beginning to be understood. The transfer of all the elements of the Programming System to operations is planned to be complete in 1982.

References

1. Tausworthe, R. C., "Software Production Methodology Test Bed Project," *Deep Space Network Progress Report 42-33*, pp. 186-191, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1976.
2. Bryan, A. I., "A Distributed Data Base Management Capability for the Deep Space Network," *Deep Space Progress Report 42-33*, pp. 32-36, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1976.
3. Tausworthe, R. C., "Standardized Development of Computer Software, Vol. I - Methods; Vol. II - Standards," Jet Propulsion Laboratory, Pasadena, Calif., July 1976.

Table 1. Programming System benefits

Benefits	Elements						
	Standard practices and methodology	Standard NRT interpreter and compiler	Standard RT interpreter and compiler	Standard file editor	Implementation aids	Data base standards	Machine-independent design
Software reliability	b	a	b		b	a	b
Accurate scheduling	b	a	b		b		
Accurate costing	b	a	b		b		
Reduced implementation costs (program development)	a	a	b	b	b		a
Reduced operations costs (OPS people time)	a	b	b	b		b	
Reduced sustaining costs (show 1108 vs PDP, keeping within specs)	b		b		b	a	a
ECO engineering (improvements to software)	a	b			a	a	a
Lower computer production costs (CPU)		c	c				
Reduced training/retraining costs	a	a	b	b	b	b	b
No new file editors				b			
Better management and implementor communication	b	b	b	a	a	a	a
Reduced test time	b	a	b	a	a	a	a
Application software insulated from O.S. changes		a	b				
Same engineer for multiple tasks	b	a	b		a	a	a
Portable software	b	a	b	b	a	b	b

^aPredicted information only

^bSome empirical evidence exists, from both DSN and external data sources

^cBenefits only if interpreter used for development activities and compiler for production

**ORIGINAL PAGE IS
OF POOR QUALITY.**

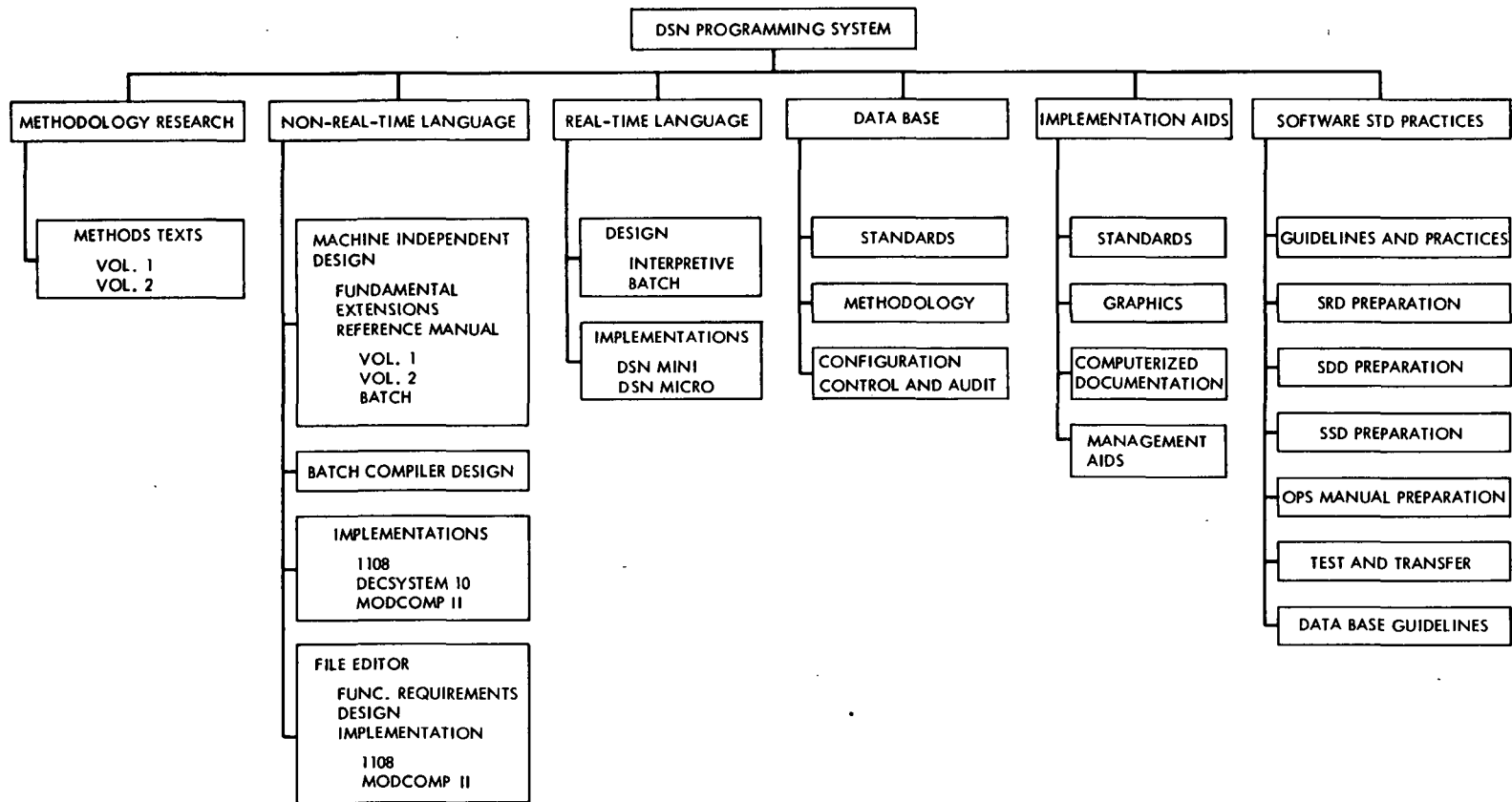


Fig. 1. DSN Programming System

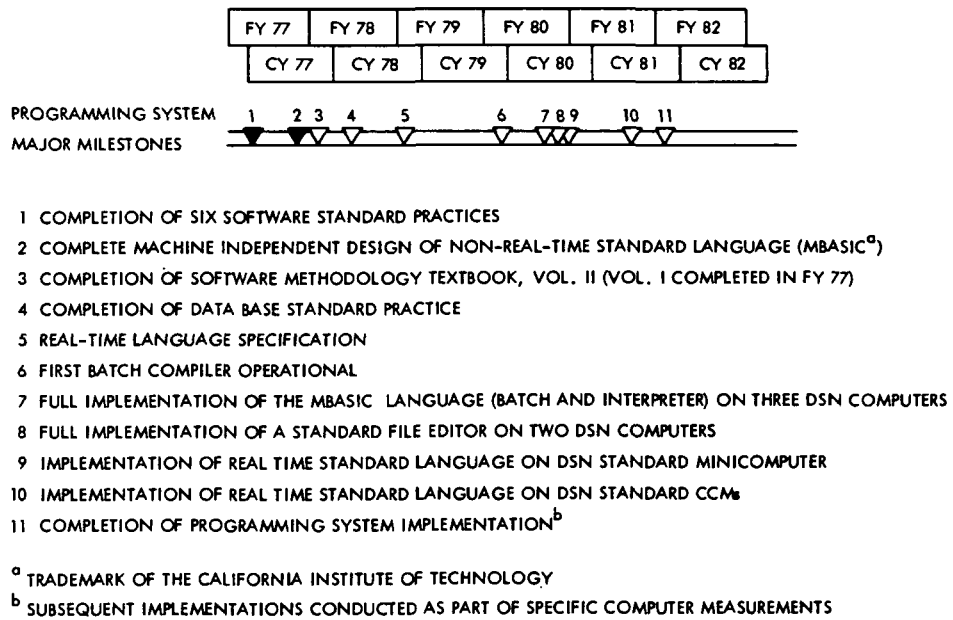


Fig. 2. DSN Programming System overview

**ORIGINAL PAGE IS
OF POOR QUALITY**

N78-11135

DSN Monitor and Control System, Mark III-77

W. G. Stinnett
TDA Engineering Office

A description of the DSN Monitor and Control System, Mark III-77 configuration is discussed. The implementation is noted that explains the evolution from the Mark III-75 to the Mark III-77 configuration.

I. Introduction

The DSN Monitor and Control System, one of the six DSN systems, is currently undergoing implementation that will lead to the Mark III-77 configuration. Implementation is occurring at the Deep Space Stations (DSS), in the Ground Communications Facility (GCF), and at the Network Operations Control Center (NOCC). Implementation is occurring over a 17-month interval at the Deep Space Stations. The first station, DSS-12, completed the Mark III-77 configuration in November 1976. (At the time of submittal of this article, DSS 14, DSS 44, and DSS 62 have completed reconfiguration, with DSS 42 and 43 in process.) The last station to be reconfigured will be DSS 11 in March 1978. Each station is nominally offline for two months for reconfiguration. This two month period is followed by two months (nominal) of test and training before being returned to full flight operations support status. The implementation of the Mark III-77 configuration in the GCF is occurring at two locations – the Deep Space Stations and at JPL. At the Deep Space Stations, the implementation is consistent with the station-by-station Mark III-77 reconfiguration schedule. At JPL, implementation of the Mark III-77 GCF configuration is scheduled for completion in November 1977. At the NOCC the implementation consisted of software modifications to simultaneously process data from the Mark III-77 reconfigured stations and the Mark III-75 configured station. The schedule for this implementation was consistent with the first station (DSS 12) coming on line in the Mark III-77 configuration.

Figure 1 provides a block diagram of the DSN Mark III-77 Monitor and Control System. Major implementation (from the Mark III-75 to Mark III-77 configuration) is noted on the block diagram. This configuration is to support the ongoing Pioneer 6-11, Helios 1 and 2, and Viking missions. Further, this configuration is to support the yet to be launched Voyager 77 and Pioneer Venus 78 missions. (The Voyager spacecraft will be launched by the time this article is published.)

II. Deep Space Station Implementation

Major implementation or modifications are occurring at the Deep Space Stations to develop the Mark III-77 configuration. The major implementation includes:

- (1) The Star Switch Controllers (SSC) to serve as the intercomputer hardware communications device.
- (2) The station centralized input/output device for the computer based subsystems.
- (3) The software in the station monitor computer (XDS 920, Digital Instrumentation Subsystem).
- (4) The Station Monitor and Control (SMC) console.

The Star Switch Controllers provide the heart of the data flow paths in the computer network at the stations. Modcomp

II-25 computers are being implemented, as part of the Mark III Data Subsystem upgrade, for the command, telemetry, tracking, and communication functions. These computers pass system data to JPL, input/output (I/O) data to the centralized I/O device, and subsystem to subsystem data, via the Star Switch Controllers. Figure 2 provides a block diagram of a typical station Star Switch Controller and computer configuration. The centralized I/O device consists of two (prime and backup) CRTs with keyboards and a character printer. This centralized I/O device provides the station with the capability for a single operator to control the command, telemetry, tracking, and communications computers. The CRTs have paging capability with the first page being for operator input and response display, as well as a single line status display for each of the computers being serviced by the centralized I/O device. Page two provides detailed status, upon operator request, for all the computer based subsystems. The character printer provides a centralized hardcopy log of operational alarms and events from the computers.

The software in the station monitor computer underwent major modifications to accommodate the Mark III-77 reconfiguration. Although the function remained similar, significant changes occurred in the software to accommodate data structure changes in the interfaces to the command, telemetry, tracking, and communication computers. Additionally, major changes occurred to accommodate the protocol and timing characteristics of the Star Switch Controllers.

The Station Monitor and Control (SMC) console provides a centralized station operator position for control and monitoring of the non-computer based subsystems. The purpose of implementing this assembly was to minimize station operations costs by minimizing the number of subsystems that required a local operator. The SMC console was installed at all stations during calendar year 1976 and early calendar year 1977.

III. Ground Communication Facility Implementation

The implementation of the Mark III-77 Ground Communication Facility (GCF) high speed data subsystem will be complete in November 1977. However, the nature of the GCF support is such that the Mark III-77 configuration will co-exist with the Mark III-75 configuration through late CY 1978. The existing Mark III-75 configuration at the Jet Propulsion Laboratory will be maintained through the Viking and Helios end of missions (late CY 78). Additionally, the Mark III-75 configuration at the Deep Space Stations will exist until the last station is reconfigured (March 1978, DSS 11). Figure 3

provides a block diagram of this dual Mark III-75 and Mark III-77 GCF configuration.

At the Deep Space Stations, the Communication Monitor and Formatter (CMF) computer is being implemented on a schedule consistent with the station-by-station Mark III-77 reconfiguration. The CMF interfaces with the station computers via the Star Switch Controllers to pass high speed data to and from the station.

At the Jet Propulsion Laboratory, in the GCF Central Communications Terminal (CCT), three types of functional computers are planned to be implemented for the Mark III-77 configuration. An Error Detection Correction (EDC) assembly, a High Speed Switch (HSW) assembly, and a Central Communication Monitor (CCM) assembly are planned for implementation.

The function of the EDC assembly is to establish a protocol with the CMF computer at a Deep Space Station and perform error detection and correction by retransmission for block errors and short line outages. The purpose of this implementation is to provide a more reliable real time data delivery service to the flight projects, as well as cut down on the non-real time playback of data after a spacecraft track. With this error correction by retransmission, it is expected that 100% real time HSD delivery to flight projects will be approached.

The HSW assembly provides for automatic routing of high speed data blocks between the appropriate Deep Space Station, the Mission Control and Computing Center (MCCC), the Network Operations Control Center (NOCC), and Ames Research Center (ARC). The automatic routing function will be accomplished by software by use of destination codes contained within the high speed data messages.

The CCM provides the GCF and Network operations personnel with information regarding the GCF status. Status data is received from the CMF at the Deep Space Stations, and from the equipment in the Central Communication Terminal, and displayed to operations personnel.

IV. Network Operations Control Center Implementation

Major changes to the Network Operations Control Center (NOCC) monitor software occurred for the Mark III-77 reconfiguration. Although similar in function, the high speed data received and processed from the Mark III-77 reconfigured stations changed significantly in structure. Additionally, simultaneous processing of Mark III-75 and Mark III-77 station data was required. Both these items caused significant data processing and software timing changes in the NOCC monitor software.

**ORIGINAL PAGE IS
OF POOR QUALITY**

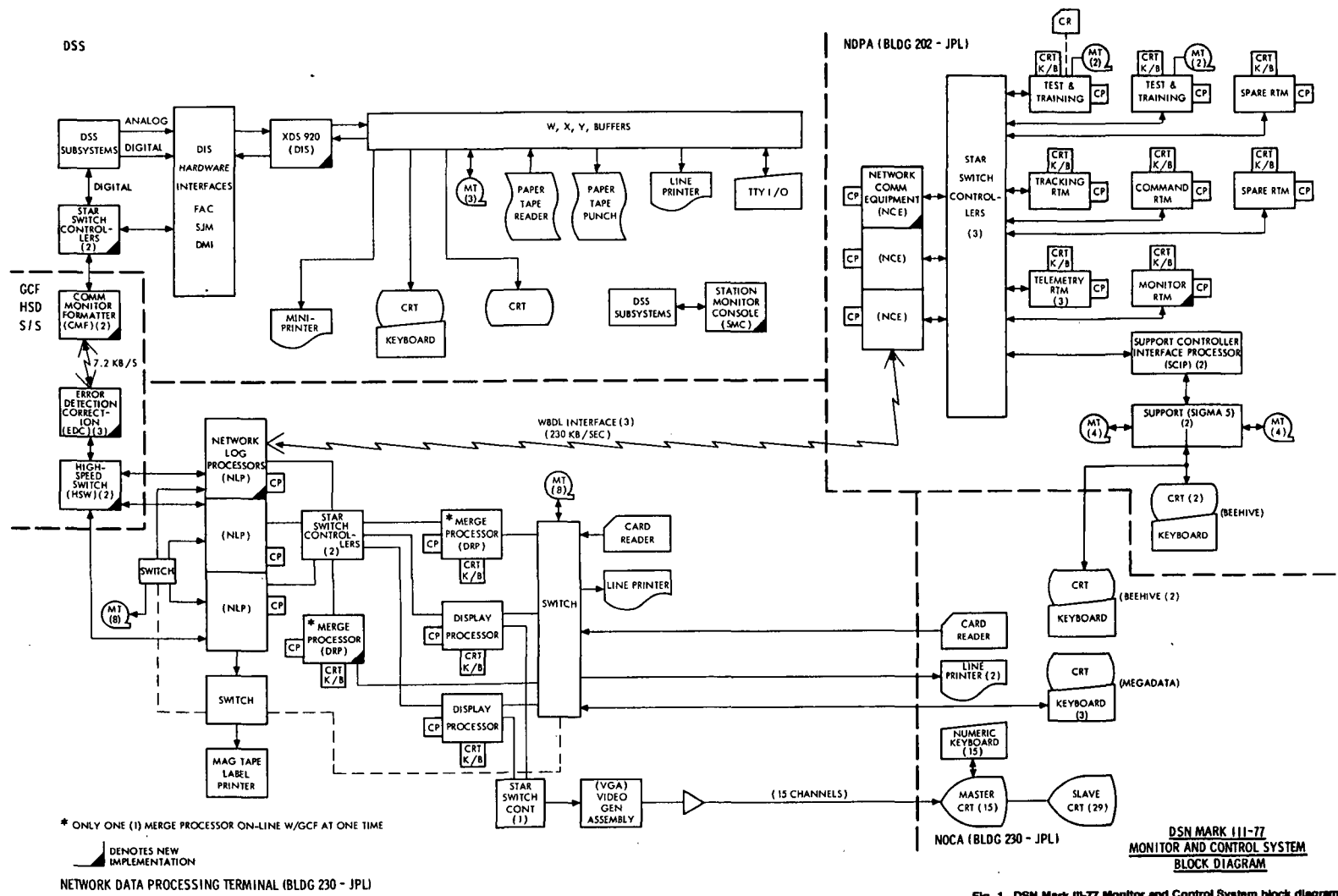


Fig. 1. DSN Mark III-77 Monitor and Control System block diagram

ORIGINAL PAGE IS OF POOR QUALITY

REPRODUCING PAGE BLANK NOT FILLED

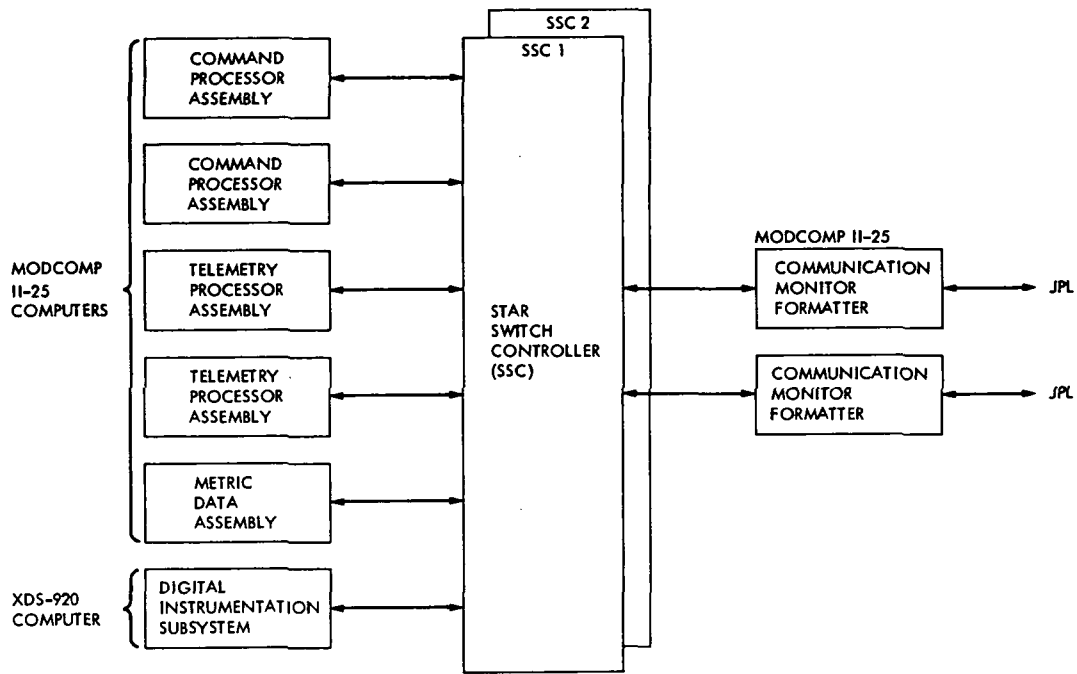


Fig. 2. Deep Space Station Star Switch Controller-computer configuration, Mark III-77

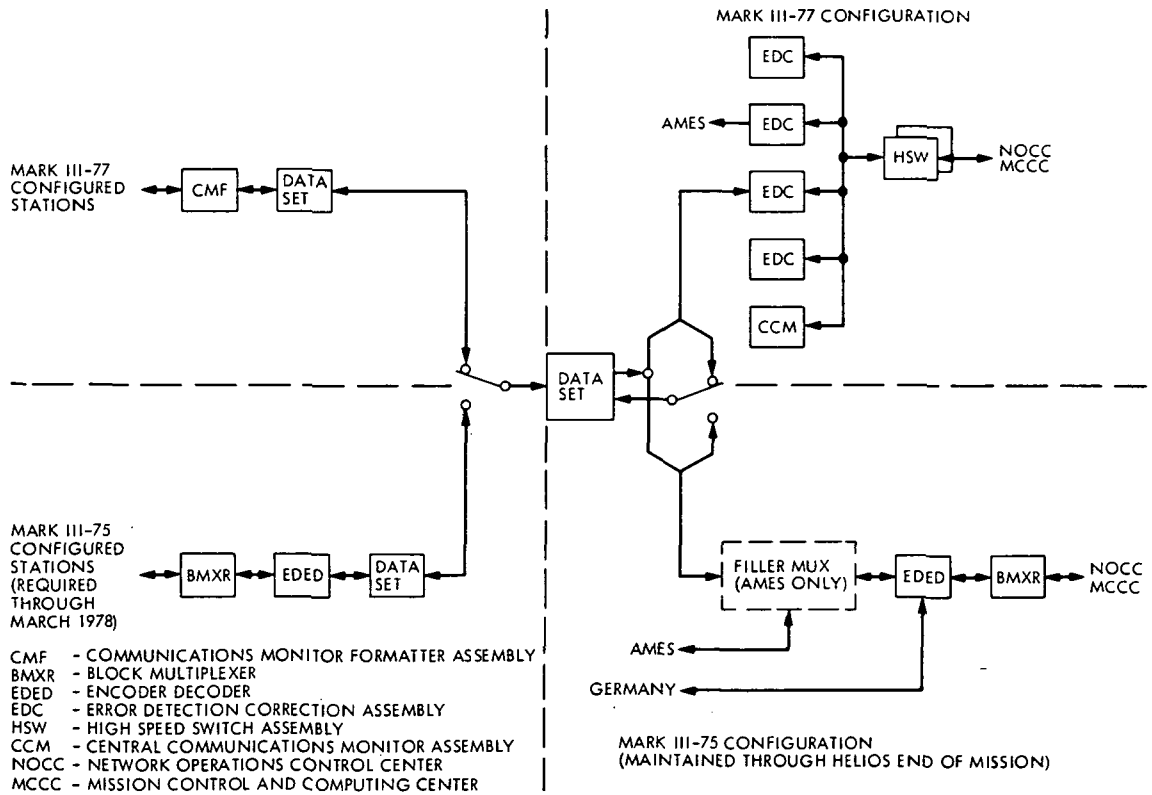


Fig. 3. GCF High-Speed Data Subsystem configuration (1976 through 1978)

PRECEDING PAGE BLANK NOT FILLED
 ORIGINAL PAGE IS OF POOR QUALITY

N78-11136

Viking Extended Mission Support

D.W.H. Johnston
Deep Space Network Operations Section

This report covers the period from May 1 through June 30, 1977 and includes the initial two weeks of post-MDS Viking related testing at DSS 14. Some of the periodic tests routinely carried out with the Viking Orbiter spacecraft by the spacecraft telecomm and DSN teams to ensure optimum performance or detect any degradation in the spacecraft radio or command equipment are also addressed.

I. Viking Operations

A. Status

During May and June 1977 all four Viking Spacecraft continued to perform as planned, with the two landers in their low activity hibernation mode and the two orbiters continuing their photography, weather and water mapping missions.

A major dust storm covering a large part of the entire planet became evident in early June, and excellent pictures of the cloud formations were obtained by the orbiters, although the landers' pictures quickly deteriorated to the point where the sun was practically obscured. At the end of June this dust storm condition remained unchanged.

On May 27, 1977 VO-1 passed within 100 km of Phobos and obtained some spectacular pictures of the Martian moon.

The only orbiter spacecraft performance change noted during May and June was the degradation of 2-5 dB in the performance of the VO-1 Radio Frequency System (RFS) and/or the Command Detector Unit (CDU).

B. VO High-Gain Antenna (HGA) Calibration

Periodically during the Viking mission, calibration tests have been carried out to check the VO HGA pointing angles and confirm that the HGA is optimized to bore sight on the Earth, and check the HGA angles relative to the spacecraft scan platform, etc. A VO-2 high-gain antenna (HGA) calibration test took place on May 23, during the beginning of the DSS 43 view period. The station performed the X-band CONSCAN work per the sequence of events. Each of the 10 calibration points was held for 10 min.

C. Maneuvers

VO-1 MOT-15 took place on May 15, 1977. This was a sunline maneuver to eliminate the possibility of a tangential grazing (or impact) with Phobos during the flyby on May 27, 1977.

VO-1 MOT-16 took place on July 1, 1977. This orbit trim maneuver changed the orbit period to 24 hours and caused the Earth occultations to occur during the DSS 14/DSS 43 overlap period.

II. Radio Science

Viking has been continuing to acquire data for five Radio Science experiments since the end of the superior conjunction period when the intensive coverage for the General Relativity Time Delay and Solar Corona Experiments were completed.

A. General Relativity Dynamics Experiment

The data collected for this experiment is primarily the near simultaneous lander and orbiter ranging data. Orbiter ranging and VLBI data are also useful. Over the past three months the Viking Project has scheduled at least two lander ranging passes per week, and the DSN has acquired good data on most of these passes. The problems during the prime mission with acquiring good near simultaneous lander-orbiter ranging seem to have disappeared due to increased attention that the Viking Project and DSN planning and operations personnel have given to the ranging data acquisition.

The last Orbiter/Quasar VLBI pass for several months occurred on April 14. Since DSS 14 or DSS 43 will be down for Mark III conversion there will not be another opportunity until fall to continue these observations.

B. Mars Physical Ephemerides Experiment

This experiment is concerned with the motion of the Mars pole and spin rate for the purpose of determining the internal mass distribution of the planet. The same near simultaneous lander/orbiter ranging data used for the General Relativity Experiments is also used for this experiment.

C. Earth Occultation Experiment

Orbiter 1 (VO-1) started Earth occultations in mid-March 1977. Occultations will continue to occur through November 1978 (~1½ years). Orbiter 2 (VO-2) started Earth occultation in mid-January and stopped at the end of May. Therefore, there were two occultations every day from mid-March through May 1977. Unfortunately, both orbiters were in non-Earth synchronous orbits with different periods, so the occultation "walked" through the DSN stations at different rates. This fact plus the non-availability of DSS 14 due to station modification and of DSS 63 due to Helios support means that most of the recent Earth occultation support comes from DSS 43. The confusion that has been caused by the changing time of occultations also resulted in the loss of occultation data due to operational procedure errors.

D. Earth Occultation

Two basic station configurations are used to acquire Earth occultation data:

In Configuration One (Fig. 1) the 64-m DSS is locked up two-way to the occulting orbiter at entry, and the DSS acquires S- and X-band Open-Loop Receiver data, and S- and X-band closed-loop high-rate doppler at 10 points/second. At exit, the open- and closed-loop receivers acquire only one-way data, because the spacecraft loses lock when the orbiter disappears behind the planet.

One-way data is inferior to two-way or three-way data because the spacecraft auxiliary oscillator is not as stable as the station frequency reference. Generally one-way data will provide planetary radii and ionospheric profiles but not atmospheric temperature and pressure profiles.

In Configuration Two (Fig. 2) the 64-m DSS is locked up two-way to the non-occulting orbiter throughout the time when the other orbiter is being occulted. The DSS acquires S- and X-band open-loop receiver data from the occulting orbiter, and S- and X-band closed-loop high-rate doppler at 10 points/second from the non-occulting orbiter.

In this configuration the carrier signal from the occulting orbiter passes through the Martian atmosphere and ionosphere plus the interplanetary plasma and the Earth's atmosphere and ionosphere, while the signal from the non-occulting orbiter passes through everything except the Martian atmosphere and ionosphere. These data can then be processed by differencing the signal perturbations, and this difference will be the Martian atmosphere and ionosphere effects, which is the object of the Earth occultation experiment.

As can be seen the Configuration Two observations are more complicated, and therefore less reliable, and more costly to process. When another station is two-way with the occulting orbiter, the open- and closed-loop entry data are three-way, which is as effective as two-way for the Earth occultation experiment (Figs. 3, 4). Because DSS 63 does not have open-loop receivers and recorders, only closed-loop data in Configuration One is obtainable from this station.

E. Gravity Field Experiment

Full orbits of S-band doppler data for both orbit determination and gravity field determination continued to be collected. Very little good S- and X-band doppler short-arc coverage near periapsis for local gravity anomaly analysis has been collected during the entire mission due to conflicts for 64-m coverage, orbiter roll, and orbiter occultations near periapsis.

F. Surface Properties Experiment

Lander-to-orbiter relay link engineering data used for the Surface Properties Experiment continues to be acquired about once per week.

G. Near Simultaneous Lander and Orbiter Ranging

The VL-1 ranging passes during late April and early May occurred during the early part of the DSS 43 pass, followed by VO-1 ranging during the latter part of the passes.

Since VL-2 does not have any direct downlinks, the radio science relativity and planetology experiments depended completely upon the VL-1 direct links. Also because of lack of VL-2 relay links there was competition between the direct link telemetry data and ranging data.

The orbiter near simultaneous ranging is used primarily to calibrate for the ranging group delay due to interplanetary and ionospheric plasma. The orbiter S- and X-band ranging can be used to measure the total electron content in the beam between spacecraft and station and these calibrations are extrapolated back to the lander ranging measurement time to get a more accurate lander ranging measurement. Since the plasma is a time variable, the lander and orbiter ranging measurements must be made as closely together as possible. Normally the time between a lander and orbiter ranging measurement has been about 2 hours.

The near simultaneous lander and orbiter passes have been generally successful during the Viking Extended Mission, and the radio science experimenters will achieve many of their objectives using this data.

III. Network Support

Table 1 gives the DSN tracking support; Table 2 the commands transmitted; and Table 3 the Discrepancy Report status for May and June 1977. The same category statistics for the previous months of 1977 are also shown to give a more complete picture and emphasize any trends in the data.

IV. DSN Mark III Data Subsystem (MDS) Testing

As indicated in the last report of this series the DSN Mark III Data Subsystem has been implemented and tested at

DSS 12, DSS 44, and DSS 62. During this reporting period DSS 14 completed their implementation and system performance testing and in mid-June started Viking operational training and testing. The tests completed up to June 30, 1977 follow.

A. DSS 14 Viking Extended Mission (VEM) OVT I, June 24, 1977, 1900-0330 PDT

The test was unsuccessful. Due to problems with NOCA test and training RTM and the station SCA, telemetry could not be simulated in accordance with the SOE. After TPA 1 failed, the remaining time was turned back to the station for maintenance, as the station had to be ready for the first Viking SIM test on June 25, 1977.

B. DSS 14/NOCA/VMCCCC Viking SIM test, June 25, 1977

This test was very successful. The test was intended to check out the long loop simulation system prior to the initial Viking SIT test on June 28, 1977. This was accomplished in the first three hours of the test, and the remainder of the test time was used to follow the SIT sequence of events, and approximately 70% of the SIT sequence was successfully accomplished.

C. DSS 14 VEM SIT 1, June 28, 1977, 0710-2110 PDT

The test was successful. This test fully demonstrated the capability to support the Viking Mission system requirements with long loop simulated telemetry from VMCCCC to DSS 14 and back to VMCCCC and NOCA.

D. DSS 14 OVT 2, June 29, 1977, 0200-1410 PDT

Test successful and completed SOE items on real-time data; however, lack of time prevented completion of the digital and analog telemetry replay items.

Table 1. VEM tracking support 1977

DSS	Jan.		Feb.		March		April		May		June	
	Tracks	Hours	Tracks	Hours	Tracks	Hours	Tracks	Hours	Tracks	Hours	Tracks	Hours
11	23	135	22	142	10	100:12	17	118:00	38	227:59	40	289:16
12	4	11	1	6			24	175:59	17	119:20	1	04:03
14	52	341	59	392	50	368:35	20	176:21	-	-	-	-
42	21	247	25	226	58	453:24	17	138:36	17	162:29	14	112:42
43	68	721	62	627			63	603:21	60	521:41	57	486:21
44	-0-	-0-	-0-	-0-	7	7:02	1	3:56	-	-	-	-
61	35	261	29	227	12	72:07	40	317:45	54	461:32	51	474:55
62	-0-	-0-	2	7	4	22:25	9	55:10	3	13:54	2	07:32
63	38	327	28	202	66	525:23	15	78:01	23	185:56	15	136:22
Total	241	2043	228	1830	207	1549:38	206	1667:09	212	1692:51	180	1398:39

Table 2. VEM commands transmitted

DSS	Jan.	Feb.	March	April	May	June
11	1521	1394	1027	117	811	-
12	0	0	0	1314	721	-
14	769	1404	1206	274	-	-
42	2072	953	1778	8	1886	1619
43	919	2523	0	2094	1447	972
44	0	0	2	1	-	-
61	605	1116	1328	1925	1922	3838
62	0	0	1	1991	-	496
63	795	472	2039	381	675	383
Total	6681	7862	7381	6180	7465	7308

Table 3. DSN VEM discrepancy reports

DSS	Jan.		Feb.		March		April		May		June	
	Open	Closed	Open	Closed	Open	Closed	Open	Closed	Open	Closed	Open	Closed
11	4		3	4	4	6	1	3	2	3	2	6
12	4						5	2	7	5	0	7
14	14	2	11	19	4	33	3	9	2	2	0	2
42		1	2	3		7		2	0	0	0	0
43	10	13	11	10		12	9	11	8	17	2	14
44						2		1	0	0	0	0
61	1	9	1	6		3		1	1	2	0	6
62				8	1	2	2	1	0	2	0	1
63	1	4	7	3	1	18		6	4	4	3	12
DSN, NDPA, NOCA	4	3	3	9	2	10	4	7	7	12	10	13
Total	38	32	38	62	12	93	24	43	31	47	17	61

ORIGINAL PAGE IS
OF POOR QUALITY

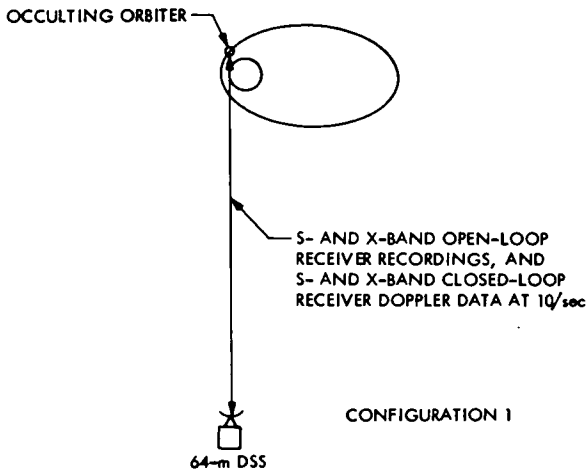


Fig. 1. One 64-m DSS two-way with occulting orbiter at entry and one-way at exit

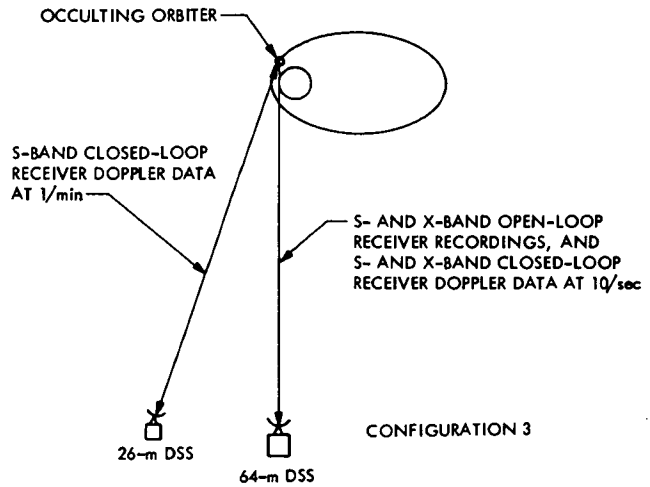


Fig. 3. One 64-m DSS three-way with occulting orbiter at entry and one-way at exit, and one 26-m DSS two-way with occulting orbiter at entry and one-way at exit

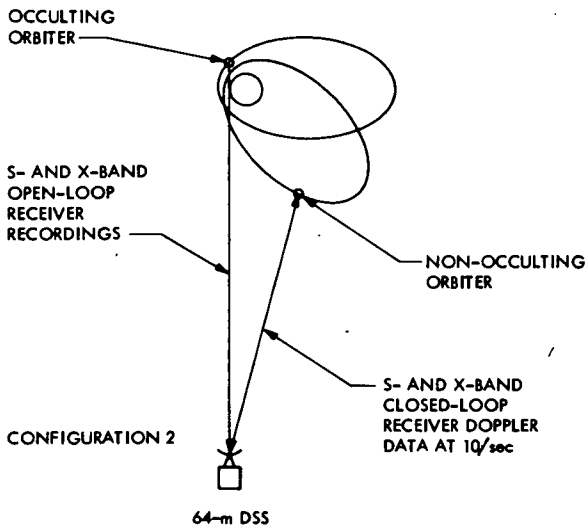


Fig. 2. One 64-m DSS two-way with non-occluding orbiter and one-way with occulting orbiter at entry and exit

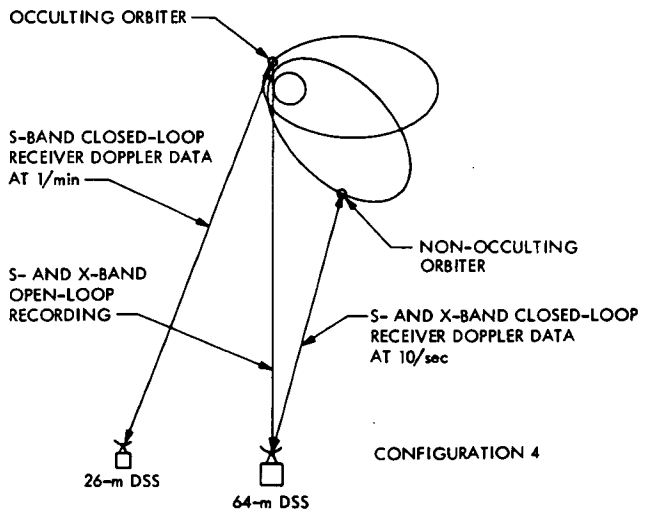


Fig. 4. One 64-m two-way with non-occluding orbiter and three-way with occulting orbiter at entry and one-way at exit

Results of Pioneer-Venus 1978 Sequential Decoding Tests Over a Simulated Lognormal Fading Channel

B. K. Levitt

Communications Systems Research Section

Tests were performed in 1976 at the Jet Propulsion Laboratory's Compatibility Test Area (CTA 21) to investigate the effects of lognormal fading on Pioneer-Venus 1978 (PV78) sequential decoding performance. The results of these simulations were used to refine a previously developed empirical model for predicting PV78 telemetry link performance.

I. Introduction

The PV78 orbiter will dispatch several probes into the planet atmosphere. The telemetry from these probes will be transmitted directly to Earth at data rates between 16 and 256 b/s over a coherent, binary PSK link using a (32, 1/2) convolutional code. These communication links will be degraded by lognormal fading due to turbulence in the atmosphere of Venus (Refs. 1, 2).

In connection with the Helios mission, Layland developed a medium-rate empirical model for analytically determining the effects of a noisy carrier reference on coherent, sequential decoding performance (Ref. 3). For PV78 applications, this model was modified to include the additional degradation of lognormal fading (Ref. 4). However, the accuracy of this extended fading/noisy reference model had not been verified experimentally. Moreover, although many individual decisions over widely varying time intervals are made in sequentially decoding a frame of received data, the models characterized these decoder integration times by a *single* effective interval T_m . In his pioneering work, Layland used a value of T_m which was approximately twice the bit time T_b (Ref. 3, Eq. 4). Still,

it was believed that T_m/T_b was a function of the data rate (Ref. 5, Sect. III), and that a more accurate empirical model could be produced by selecting the value of this parameter that provided the closest agreement with experimental deletion rate behavior (Ref. 4, Sect. IV).

With these considerations, tests were conducted in CTA 21 for several months during 1976 to gather experimental data on the performance of PV78 sequentially decoded telemetry at 256 b/s, with and without simulated lognormal fading. This report describes the results of that effort.

II. Test Procedure

Tests were performed using the standard DSN transmitter/receiver configuration in CTA 21, modified to accommodate the simulated channel fading. Specifically, as shown in Figure 1, the test equipment included the Simulation Conversion Assembly (SCA), a test transmitter, a PIN attenuator modulated by a Gaussian noise generator to simulate lognormal fading (see Appendix), the Y-Factor Assembly (YFA), a DSN Block III receiver, the Subcarrier Demodulator Assembly (SDA), two Symbol Synchronizer Assemblies (SSA's), two

Telemetry/Command Processors (TCP's), the Data Decoder Assembly (DDA), and a magnetic tape unit which provided the test output Original Data Record (ODR).

The SCA supplied a pseudorandom data sequence to the test transmitter. The (32, 1/2) convolutionally coded data was biphasic – modulated onto a 32.768 KHz squarewave subcarrier, which in turn phase-modulated the S-band carrier at the test transmitter. The carrier modulation index θ was set at the test transmitter before each test. The data rate for all tests was 256 b/s.

Based on previous experience, the PIN attenuator had a tendency to drift until it reached its ambient operating temperature: consequently, it was always turned on at least 24 hours prior to a test. With the noise generator off, and the attenuator at its d.c. bias point, the “received” signal-to-noise ratio was set to its approximate test level using the Y-factor measurement technique (Ref. 6). If the test specified log-normal fading with logamplitude variance σ_x^2 , the Gaussian noise generator was then set to a corresponding signal level which depended on the PIN attenuator input/output characteristic slope (see Appendix). A 1 Hz lowpass filter was inserted at the output of the noise generator to simulate the fading bandwidth of the Venus atmosphere (Ref. 1, p. 38, and Ref. 2, p. 45). The received symbol error rate (SER) was then observed on the SSA-TCP monitoring string, and fine adjustments were made in the received signal strength until the predetermined test SER was achieved. The tape unit was then engaged, and test data was collected, usually for several hours to ensure statistical validity.

The SSA-DDA-TCP string operated in real time to provide the ODR with a listing of the number of computations required to sequentially decode each frame of received data, or an indication that a particular frame had been deleted because the system capacity of 25000 computations per second (or about 100 computations per bit at 256 b/s) had been exceeded. Later, off-line computer processing of the ODR determined the test frame deletion rate characteristic as a function of computational speed, up to the real-time sequential decoder limit. The other SSA-TCP string was used to monitor the telemetry channel by identifying drifts in SER and received symbol signal-to-noise ratio due to fluctuations in station parameters.

III. Test Results

The results of these tests are summarized in Table 1. Of the 16 satisfactorily completed runs, 6 involved simulated log-normal fading with $\sigma_x^2 = .049$, which is the level of fading anticipated in the limiting case of a PV78 probe transmitting from the surface of Venus to Earth in a direction 60° off

vertical (Ref. 5, p. 53). Two runs were made at each modulation index θ ; in both the fading and non-fading cases, the central values of θ were selected because they were near optimum in the sense that the PV78 link design deletion rate of 10^{-2} could be achieved with minimum signal power. The received signal-to-noise level, as evidenced by the measured SER, was selected for each test to produce a frame deletion rate in the vicinity of 10^{-2} , based on empirical predictions. Ten independent measurements of SER were made before and after each run, and the average of these is recorded as AVG SER in Table 1.

Prior to the sequential decoding efforts, models had been developed for determining the bit error rate of uncoded, coherent binary PSK telemetry with a noisy carrier reference, with and without fading (Refs. 2, 7). These uncoded models do not have the problem of a variable integration time, characterized by T_m in the sequential decoding case. Consequently, these models are inherently more accurate than their successors, particularly at medium data rates where the bit detection time T_b is on the order of the coherence time of the noisy carrier reference and the channel fading. Based on these uncoded models, curves of SER versus θ were plotted in Figs. 2 and 3 for various signal-to-noise ratios (defined as received signal power P/noise spectral power density N_0). These plots were used to convert the AVG SER's into experimental (EXP) SNR's in Table 1. It is believed that these SNR's are more representative of the actual test values than those derived by the Y-factor method. A basic premise inherent in Figs. 2 and 3, and the work that follows, is that the carrier and data channel losses (I_C and L_D) in the test equipment due to predetection recording, waveform distortion, SSA, and SDA are given by the PV78 design values of 1.00 and 2.08 dB respectively (Ref. 5, p. 52).

Off-line processing of the ODR's generated the plots of experimental frame deletion rate versus sequential decoder computational rate N , in computations/bit, indicated by the X 's in Fig. 4. A smooth curve (not shown in Fig. 4) was drawn through the experimental data for each run, and values of the deletion rate at $N = 5$ and $N = 100$ computations/bit were read off these curves and recorded in Table 1. The sequential decoding performance models of Refs. 3 and 4 compute T_m according to the formula (Ref. 4, Eq. 5)

$$T_m = KT_b \left[1 - \frac{1}{N} \log_2 \left(1 + \frac{N}{2} \right) \right] ; K = 2 \quad (1)$$

One of the objectives of these tests was to determine whether some other value of K ($\cong T_m/T_b$ for large N) would provide better agreement between the empirical models and experimental data. To this end, the distance parameter

$$D(K) = \sum_{\text{run}=1}^{16} \sum_{N=5,100} \left[\log_{10} \frac{\text{EMP. DELETION RATE}(K)}{\text{EXP. DELETION RATE}} \right]^2 \quad (2)$$

was computed over a range of K , where the numerator of the argument above was based on a generalized version of the models in Refs. 3 and 4, with arbitrary K , with P/N_0 given by the EXP SNR values in Table 1, and with telemetry carrier and data channel losses $L_C = 1.00$ dB and $L_D = 2.08$ dB. As shown in Fig. 5, $D(K)$ was minimized near $K = 7$.

The modified sequential decoding models with $K = 7$ were used to generate the deletion rate versus modulation index and received SNR curves of Figs. 6 and 7, for the PV78 computa-

tional rate $N = 100$ computations/bit. The experimental deletion rates in Table 1 at this value of N were superimposed on these curves, yielding the empirical (EMP) SNR's recorded in Table 1.

The difference $\Delta\text{SNR} \equiv \text{EMP SNR} - \text{EXP SNR}$ is a measure of the accuracy of these modified performance models. Based on the 16 test runs, ΔSNR has a mean of $-.11$ dB and a standard deviation of $.33$ dB. The negative mean indicates that the sequential decoding models are optimistic, tending to underestimate the SNR required to achieve a given deletion rate. The models have an apparent accuracy of about $.4 - .5$ dB ($|\text{mean}| + \text{standard deviation}$), and part of this error may be due to incorrect estimates of the telemetry channel losses.

Acknowledgement

This technique for simulating lognormal channel fading was developed by W. J. Hurd, and hardware implementation was provided by W. P. Hubbard, both of Section 331 at Jet Propulsion Laboratory.

References

1. Woo, R., et al., *Effects of Turbulence in the Atmosphere of Venus on Pioneer Venus Radio - Phase I*, Technical Memorandum 33-644, Jet Propulsion Laboratory, Pasadena, Calif., June 30, 1973.
2. Levitt, B. K., and Rhee, M. Y., "Effects of Lognormal Amplitude Fading on Bit Error Probability for Uncoded Binary PSK Signalling," in *The Deep Space Network Progress Report 42-21*, pp. 45-54, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1974.
3. Layland, J. W., "A Sequential Decoding Medium-Rate Performance Model," in *The Deep Space Network, Technical Report 32-1526*, Vol. XVIII, pp. 29-40, Jet Propulsion Laboratory, Pasadena, Calif., December 15, 1973.
4. Levitt, B. K., "Performance Degradation of Uncoded and Sequentially Decoded PSK Systems Due to Lognormal Fading," in *The Deep Space Network Progress Report 42-32*, pp. 58-67, Jet Propulsion Laboratory, Pasadena, Calif., October 15, 1974.
5. Levitt, B. K., "Pioneer Venus 1978: Telemetry Performance Predicts," in *The Deep Space Network Progress Report 42-24*, pp. 51-59, Jet Propulsion Laboratory, Pasadena, Calif., December 15, 1974.
6. DSIF Program Library, Documentation for Y-Factor Computer Program, DOI-5343-SP-B, Jet Propulsion Laboratory, Pasadena, Calif., September 29, 1972 (JPL internal document).
7. Layland, J. W., "A Note on Noisy Reference Detection," in *The Deep Space Network, Technical Report 32-1526*, Vol. XVII, pp. 83-88, Jet Propulsion Laboratory, Pasadena, Calif., October 15, 1973.

ORIGINAL PAGE IS
OF POOR QUALITY

Table 1. Sequential Decoding Test Results

σ_x^2	θ°	Run	Avg. SER $\times 10^{-2}$	Exp. SNR dB	Experimental Frame Deletion Rate		Emp. SNR dB	Δ SNR dB
					@ N = 5 Comp/Bit	@ N = 100 Comp/Bit		
0	48	1	9.22	31.51	2.89×10^{-1}	1.25×10^{-2}	31.72	.21
		2	9.73	31.32	6.2×10^{-1}	5.9×10^{-2}	31.30	-.02
	53	3	7.29	31.72	1.08×10^{-1}	3.7×10^{-3}	31.59	-.13
		4	7.61	31.58	4.2×10^{-2}	8.0×10^{-4}	31.96	.38
	58	5	6.73	31.52	7.5×10^{-2}	3.5×10^{-3}	31.43	-.09
		6	6.48	31.63	3.9×10^{-2}	1.6×10^{-3}	31.66	.03
	63	7	5.88	31.57	5.7×10^{-2}	3.3×10^{-3}	31.54	-.03
		8	5.81	31.60	1.85×10^{-2}	1.0×10^{-3}	31.96	.36
	68	9	5.40	31.62	1.55×10^{-1}	1.6×10^{-2}	31.17	-.45
		10	5.49	31.58	9.0×10^{-2}	6.7×10^{-3}	31.61	.03
.049	59	11	2.70	34.53	6.3×10^{-2}	7.2×10^{-3}	34.66	.13
		12	2.50	34.70	8.8×10^{-2}	1.5×10^{-2}	34.16	-.54
	64	13	2.33	34.54	8.9×10^{-2}	1.8×10^{-2}	33.80	-.74
		14	2.02	34.83	5.9×10^{-2}	5.1×10^{-3}	34.74	-.09
	69	15	2.27	34.44	8.8×10^{-2}	1.3×10^{-2}	34.06	-.38
		16	2.63	34.13	6.0×10^{-2}	2.0×10^{-2}	33.71	-.42

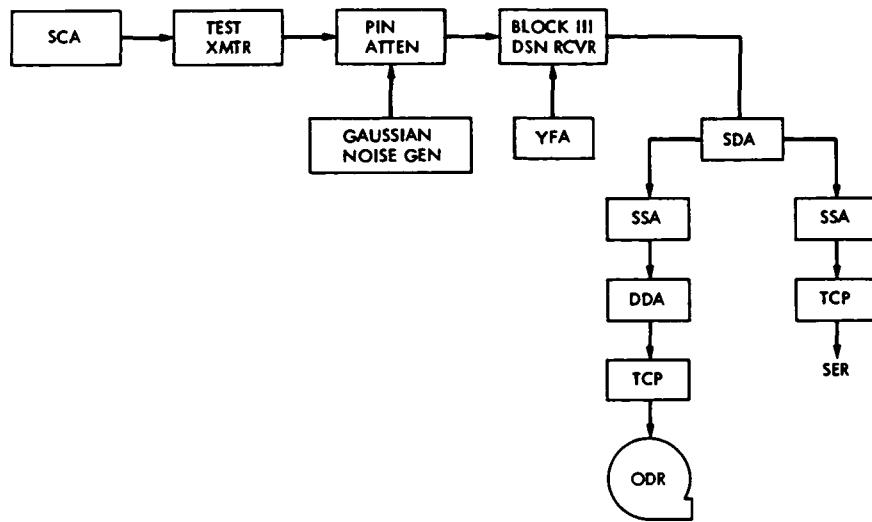


Fig. 1. Test configuration for PV78 sequential decoding tests in CTA 21

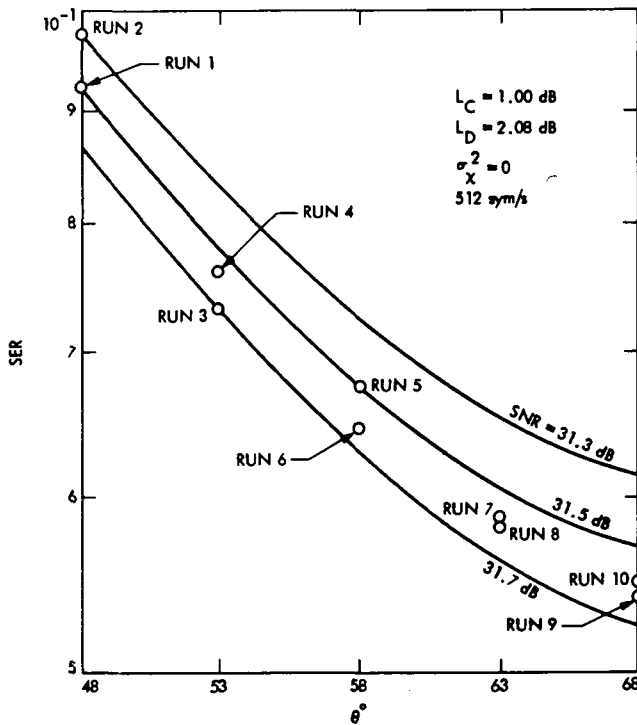


Fig. 2. Theoretical relationship between symbol error rate (SER) and modulation index θ for given received signal-to-noise ratio (SNR). Curves are for uncoded binary data at 512 symbols/s with no fading ($\sigma_X^2 = 0$).

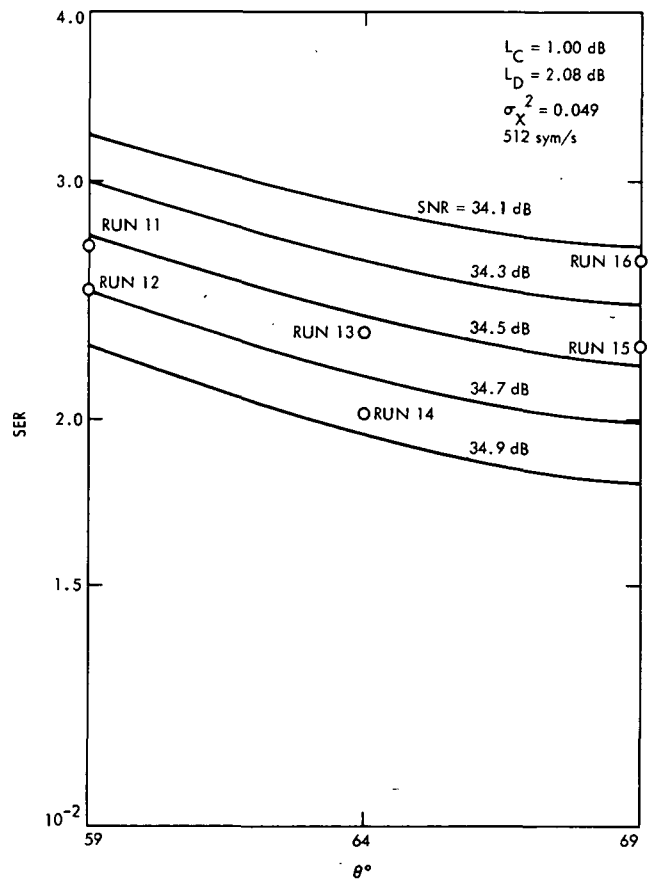


Fig. 3. Same as Fig. 2, except for lognormal fading with $\sigma_X^2 = .049$

ORIGINAL PAGE IS
OF POOR QUALITY

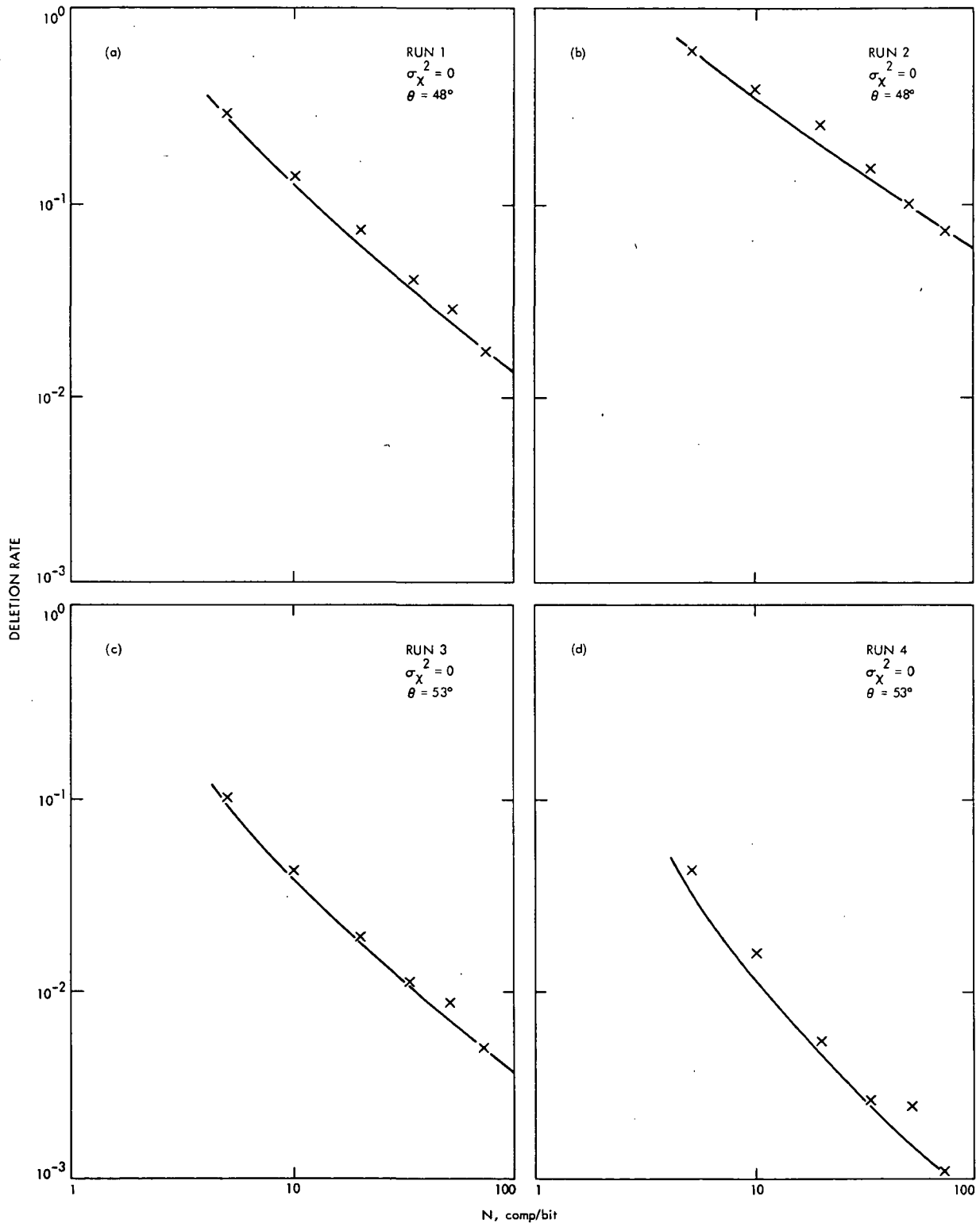


Fig. 4. Frame deletion rate test results

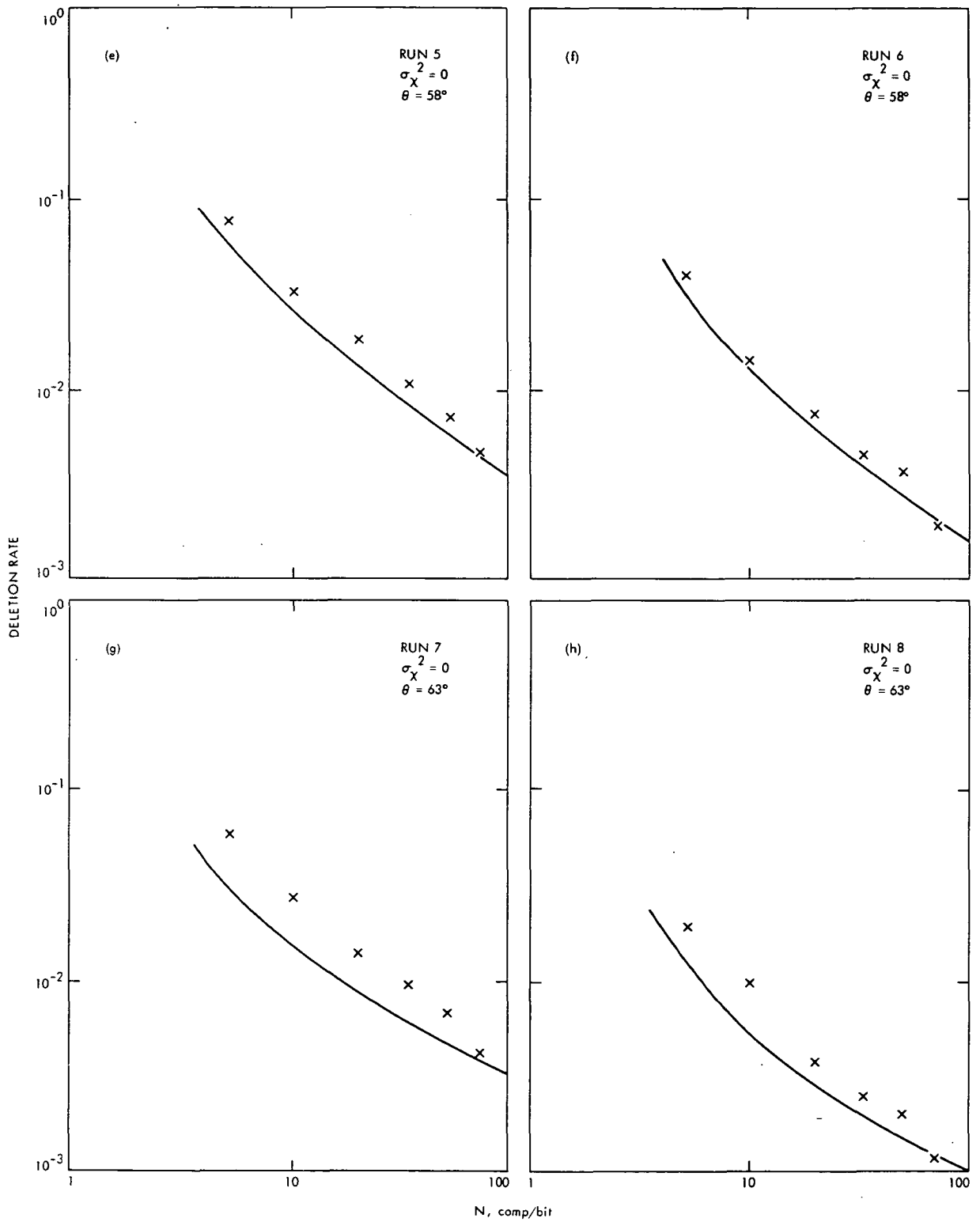


Fig. 4 (contd)

ORIGINAL PAGE IS
OF POOR QUALITY

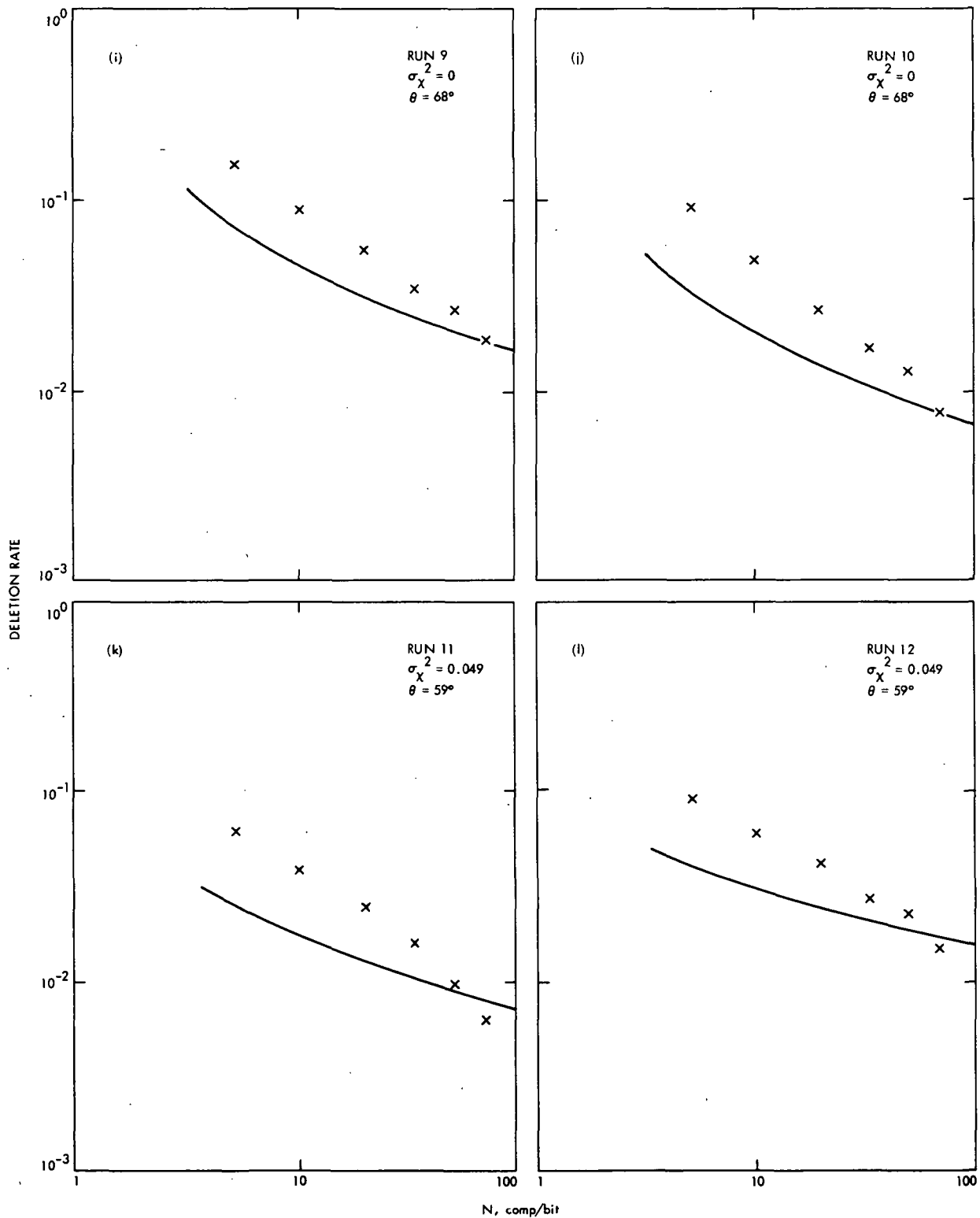


Fig. 4 (contd)

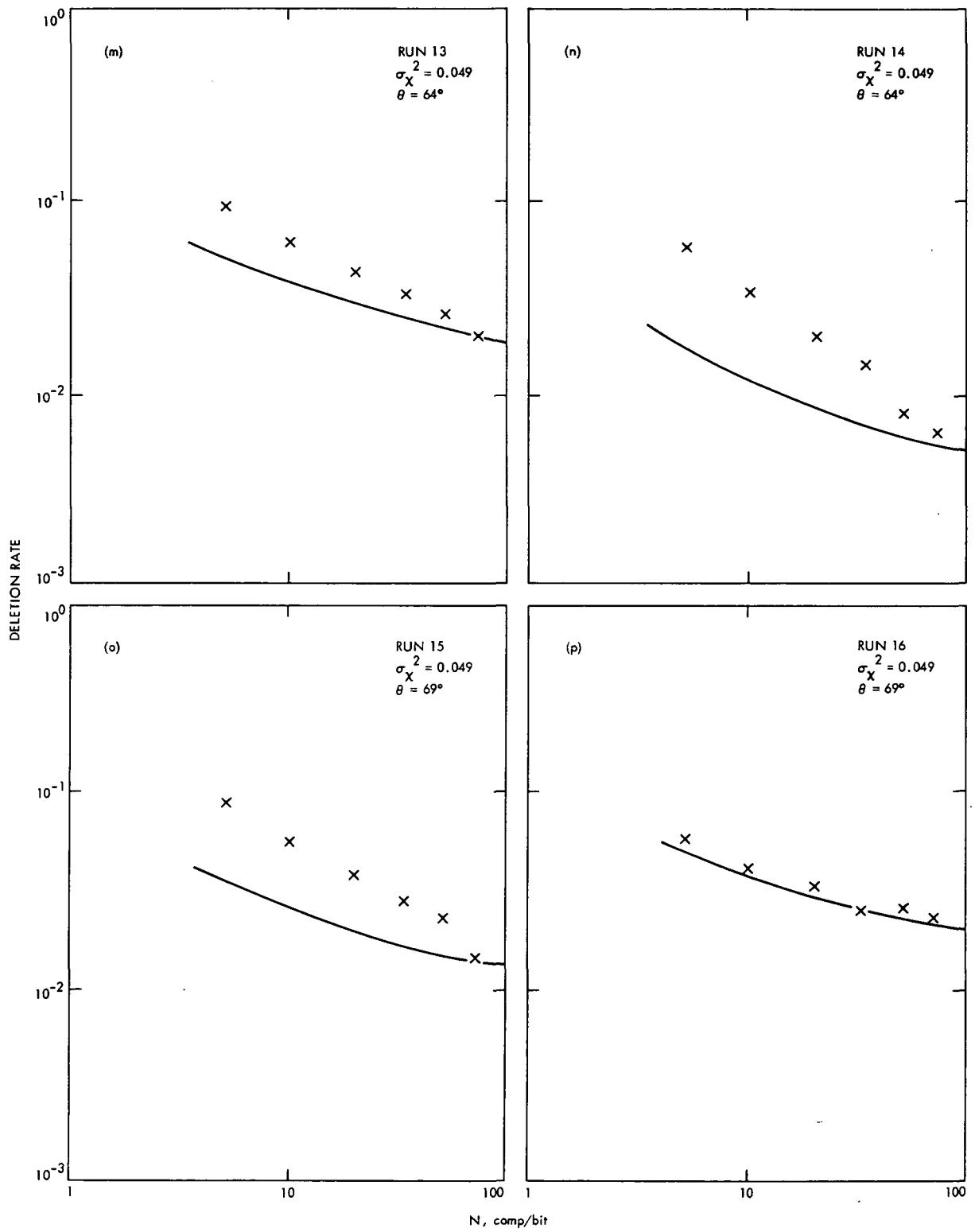


Fig. 4 (contd)

ORIGINAL PAGE IS
 OF POOR QUALITY

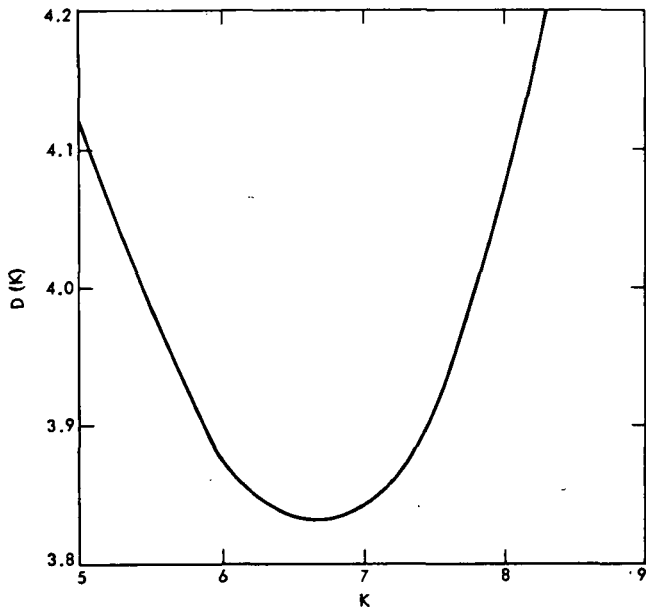


Fig. 5. Determination of $K = \lim_{N \rightarrow \infty} T_m / T_b$ for best agreement between empirical sequential decoding performance models and experimental data, based on distance parameter D

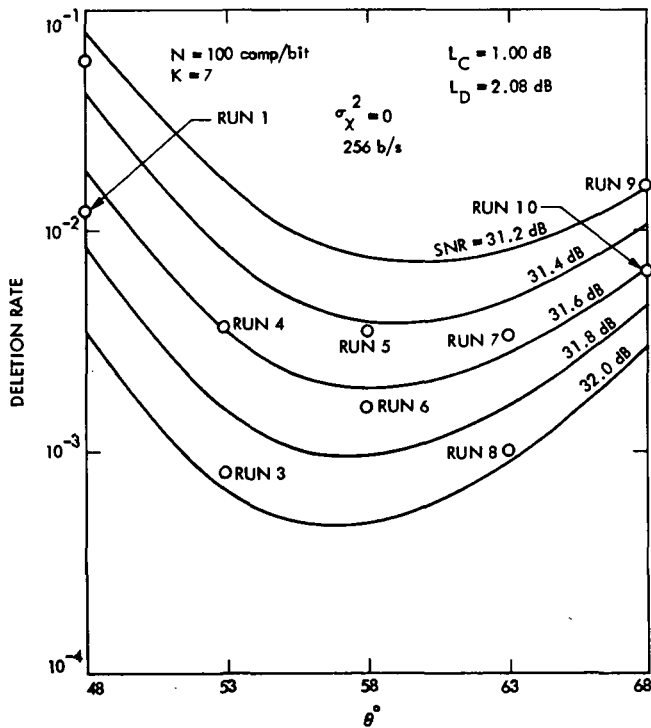


Fig. 6. Sequential decoding performance at 256 b/s with no fading ($\sigma_X^2 = 0$) based on empirical model with parameter $K = 7$.

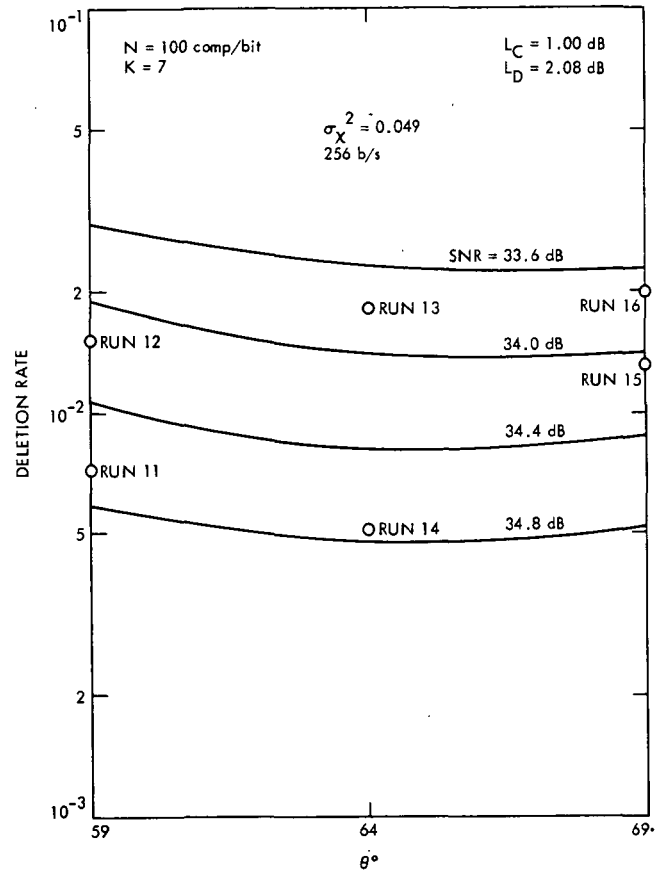


Fig. 7. Same as Fig. 6, except for lognormal fading with $\sigma_X^2 = .049$.

Appendix

Lognormal Fading Simulation

Due to atmospheric fading, PV78 received telemetry will exhibit a time-varying lognormal envelope of the form $Ae^{\chi(t)}$ (Ref. 2, Eq. 2), where $\chi(t)$ is a Gaussian random process, with $\bar{\chi} = -\sigma_{\chi}^2$ based on conservation of energy arguments (i.e., $e^{2\chi} = 1$). In db, the received amplitude has the form

$$a(t) = 20 \log_{10} A + (20 \log_{10} e) \chi(t) \text{ dB} \quad (\text{A-1})$$

In the absence of fading, $\chi(t)$ would be zero, and $a(t)$ would simply be $20 \log_{10} A$. With lognormal fading,

$$\begin{aligned} \bar{a} &= 20 \log_{10} A - (20 \log_{10} e) \sigma_{\chi}^2 \\ \sigma_a &= (20 \log_{10} e) \sigma_{\chi} \end{aligned} \quad (\text{A-2})$$

The PIN attenuator inserted in the transmission path in the test configuration of Fig. 1 has an input/output characteristic of the type illustrated in Fig. A-1. The attenuator receives a d.c. input bias voltage b_0 , selected in a relatively linear region of the characteristic. The Gaussian noise generator injects a zero-mean signal $n(t)$ about the d.c. bias point, at an RMS level

σ_n . Denoting the slope of the attenuator characteristic at the d.c. operating point by m , its gain is given by

$$a(t) \cong a_0 + m n(t) \quad (\text{A-3})$$

provided σ_n is small enough to keep the attenuator in the linear region about a_0 .

To simulate the non-fading channel, the noise generator is turned off, and $a(t) = a_0 = 20 \log_{10} A$. With the injection of $n(t)$,

$$\bar{a} = a_0; \sigma_a \cong m \sigma_n \quad (\text{A-4})$$

To simulate the fading channel, a comparison of Eqs. A-2 and A-4 shows that the RMS noise level should be set at

$$\sigma_n = \left(\frac{20 \log_{10} e}{m} \right) \sigma_{\chi} \quad (\text{A-5})$$

and the received signal strength should be reduced by $(20 \log_{10} e) \sigma_{\chi}^2$ dB.

ORIGINAL PAGE IS
OF POOR QUALITY

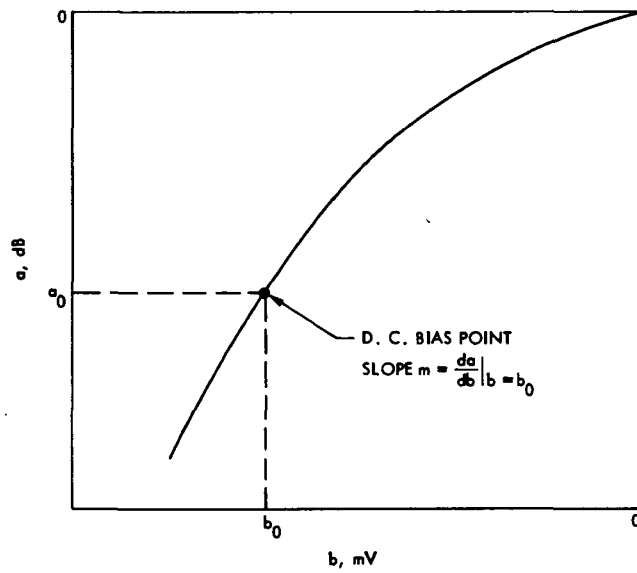


Fig. A-1. Form of PIN attenuator input/output characteristic

Pioneer Mission Support

T. P. Adamski
Deep Space Network Operations

This article reports on activities within the Deep Space Network in support of the Pioneer Project's in-flight spacecraft during the period April through July 1977. The amount of tracking coverage provided by the Network and a summary of operational testing of the Mark III Data Subsystems at DSS 14 are presented.

I. Pioneers 6, 7, 8, and 9

Coverage of these spacecraft has increased dramatically during this reporting period. The tracking coverage times shown in Table 1 represent more than a ten-fold increase over the last reporting period (Ref. 1). However, the majority of the tracks supported were for the acquisition of radio metric data only, in order to update the spacecraft trajectory data. Adequate telemetry data and command capability were provided to the Project during this period to allow spacecraft performance assessment and scientific data monitoring.

II. Pioneers 10 and 11

A. Mission Operations and Status

Pioneer 10 continues to operate satisfactorily. The spacecraft is now more than 13 astronomical units (AU) from the Earth and the round trip light time is over 3 hours 40 minutes. Tracking coverage has increased somewhat over the last reporting period, due primarily to greater availability of the 64-meter stations. Tracking coverage for the last four months is tabulated in Table 1 and coverage during the last twelve months is shown in Figure 1.

On 7 May, the spacecraft passed through superior conjunction – its fifth since launch – at a minimum Earth-Sun-probe angle of 2.83 degrees (approximately $10.4 R_{\odot}$). Telemetry data were degraded during the week surrounding conjunction, necessitating a decrease in bit rate from the normal rate of 128 bits per second to 32 bits per second. Radio metric data were also affected, as indicated by an increase in Doppler noise – although the maximum noise observed was less than expected (Fig. 2).

Pioneer 11 is also operating normally. Earth-spacecraft range is now more than 6 AU and the round trip light time is over 1 hour 50 minutes. No further problems with the spacecraft thrusters have been encountered, but the health of the entire attitude control and propulsion system continues to be closely monitored. Tracking coverage for Pioneer 11 is shown in Table 1 and Figure 3.

B. Mark III Data Subsystems Support of Pioneer

Since the previous report (Ref. 1), Mark III Data System (MDS) verification testing for Pioneers 10 and 11 has been completed at DSS 14. Configuration control for Pioneer was established on 2 August. Table 2 is a listing of representative demonstration passes performed by the station and a summary of the problems encountered. All major anomalies have been

cleared by the issuance and successful demonstration of either updated software or hardware changes. The majority of the minor anomalies encountered were related to operator training and familiarization. These problems decreased as operations personnel became more proficient and more knowledgeable of the new systems.

DSSs 42 and 43 are currently undergoing upgrading to MDS configurations. Verification testing of these stations for Pioneer support will begin in October of this year and will follow the same pattern of demonstration passes as used for DSSs 12, 14, 44, and 62. A future article will report on this testing.

Reference

1. Adamski, T. P., "Pioneer Mission Support," in *The Deep Space Network Progress Report 42-39*, pp. 9-16, Jet Propulsion Laboratory, Pasadena, California, June 15, 1977.

Table 1. Pioneer tracking coverage

Month	Spacecraft	Station Type	Tracks	Tracking time hr:min
April	Pioneer 6	26 m	3	27:52
	Pioneer 7	26 m	4	33:26
	Pioneer 9	26 m	1	12:47
	Pioneer 10	26 m	4	35:32
		64 m	31	192:59
	Pioneer 11	26 m	80	611:10
		64 m	7	38:13
May	Pioneer 6	26 m	18	112:43
	Pioneer 7	26 m	11	67:46
		64 m	6	32:10
	Pioneer 9	26 m	15	91:48
	Pioneer 10	26 m	2	7:05
		64 m	30	201:48
	Pioneer 11	26 m	84	651:11
		64 m	11	43:21
June	Pioneer 6	26 m	8	45:01
	Pioneer 7	26 m	22	154:54
	Pioneer 8	64 m	3	14:41
	Pioneer 9	26 m	7	47:20
	Pioneer 10	26 m	4	21:40
		64 m	32	331:07
	Pioneer 11	26 m	87	660:05
		64 m	6	27:28
July	Pioneer 6	26 m	11	57:42
	Pioneer 7	26 m	22	138:56
	Pioneer 8	64 m	1	1:30
	Pioneer 9	26 m	10	55:46
	Pioneer 10	26 m	4	20:34
		64 m	32	244:48
	Pioneer 11	26 m	85	654:58
		64 m	3	18:08

**ORIGINAL PAGE IS
OF POOR QUALITY**

Table 2. Pioneer MDS demonstration pass summary

Date	Duration (hr:min)	Space-craft	Anomalies	Date	Duration (hr:min)	Space-craft	Anomalies
6/24/77	5:33	10	<ul style="list-style-type: none"> (1) Unable to access CPA for data transfer test. CMF had been initialized for 22-bit GCF error polynomial instead of 33-bit code. Reinitialized CMF. (2) Continuous time errors observed by Project in telemetry data at acquisition. TPA had been initialized for default bit rate (2048 bits/second) instead of actual rate (16 bits/second). TPA reinitialized. (3) Antenna went to brake due to erroneous indication of low film height at pad 3 (4) Lost MDA/DIS interface at the DIS. MDA/DST interface remained active. Reestablished interface by DIS type-in. (5) Unsuccessful timed frame length change due to erroneous procedures. 				<ul style="list-style-type: none"> sisted throughout the pass. Data output restarted at the DIS each time. (3) Received series of command alarms (bit verify, symbol period abort limit, missing symbol interrupt, symbol period warnings). (4) Excessive angle residuals displayed by the tracking RTM. Residuals of over four degrees in each angle were displayed.
				6/26/77	7:07	11	<ul style="list-style-type: none"> (1) Unable to enter commands into the manual buffer. Illegal type-in format used by station personnel. (2) SSA dropped lock immediately upon entry of a timed frame size change scheduled for the following radio day. (Program characteristic, corrected by software update.) (3) Observed three bit verify errors from prime CPA during the pass. (4) TPA changed telemetry frame size consequent with a bit rate change. (5) Unable to perform telemetry replays post-pass. Suspected procedural errors.
6/24/77	4:00	11	<ul style="list-style-type: none"> (1) Lost high speed data line from station due to failure of carrier selector amplifier at GCF 10. (2) TPA changed telemetry frame length consequent with a bit rate change. 				
6/26/77	9:09	10	<ul style="list-style-type: none"> (1) Unable to access either CPA for command data transfer. Program reloaded. (2) Unable to maintain DIS high speed output. Program reload decreased frequency of problem, but the condition per- 	7/15/77	5:42	10	<ul style="list-style-type: none"> (1) CMF ODR stopped logging. Reloaded program. (2) SPD maser failed. Reconfigured to Mod III maser.

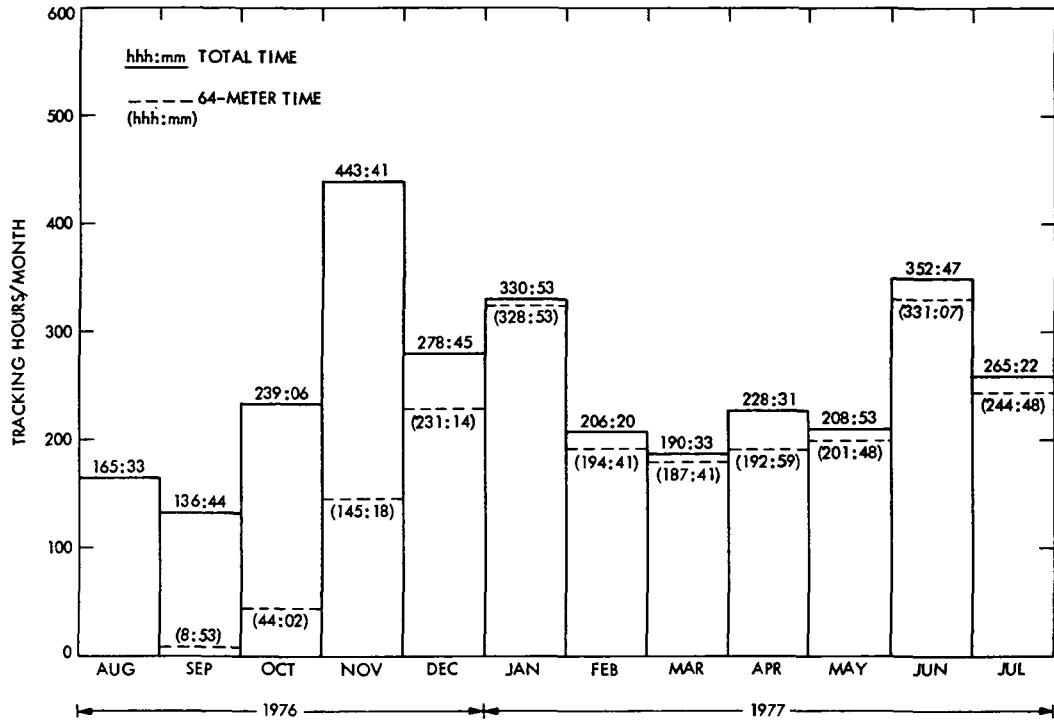


Fig. 1. Pioneer 10 tracking times

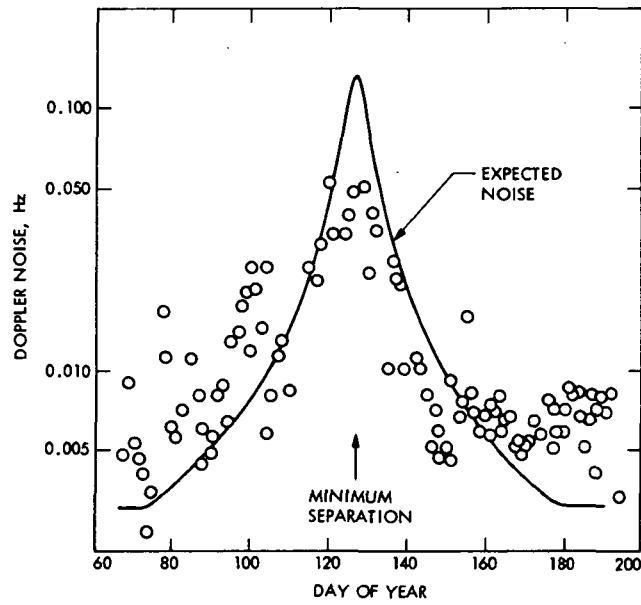


Fig. 2. Doppler noise during Pioneer 10 superior conjunction

ORIGINAL PAGE IS
 OF POOR QUALITY

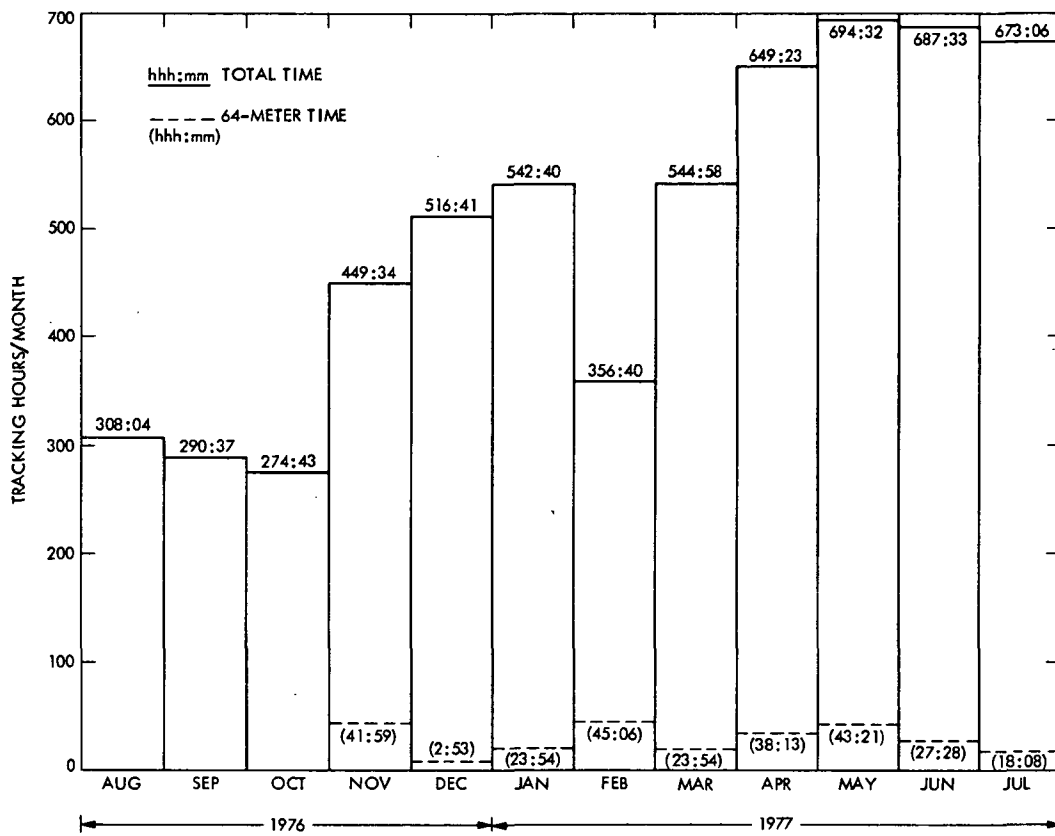


Fig. 3. Pioneer 11 tracking times

Helios Mission Support

P. S. Goodwin
TDA Mission Support

G. M. Rockwell
Deep Space Network Operations

This article reports on activities of the Network Operations organization in support of the Helios Project from 15 June through 15 August 1977.

I. Introduction

This article is the seventeenth in a continuing series of reports that discuss Deep Space Network support of Helios Mission Operations. Included in this article is information on Helios-2 superior conjunction period, a further report on Faraday Rotation experimentation (Ref. 1), the DSN Mark III Data Subsystems (MDS) update at Deep Space Station 14 (Goldstone, California), and other mission-oriented activities.

II. Mission Operations and Status

A spacecraft emergency was declared for Helios-1 on July 12, 1977, when DSS 42 (Australia) lost the downlink, and subsequent contingency sweeps to reacquire the downlink were unsuccessful. The spacecraft was then commanded, through the Australian 64-meter station, DSS 43, to go to 8 bits per second. Following this data rate change, lock of the downlink signal was achieved. The first analysis of this data showed that the transmitter in the spacecraft had gone to the medium-gain antenna from the high gain, the power regulator had gone from 1 to 2, and all experiments had turned off. The regulator switch, for reasons unknown, had caused a power drop causing the spacecraft's high-gain antenna to move off point. This resulted in the loss of the downlink. The spacecraft

was then commanded back to the high-gain antenna, and the normal configuration restored. All experiments were turned on with the exception of Experiment 10, which would be restored when sufficient power was available. On July 21, 1977, Experiment 10 was restored.

On July 18, 1977, the Helios-2 spacecraft passed through its fifth aphelion at 0025 GMT. The spacecraft was at a round-trip light time of 31 minutes, 31.7 seconds. Ground station coverage was over DSS 44 (Australia) at a bit rate of 32 bits per second coded, having a signal-to-noise ratio of 5.7 dB and a downlink AGC of -157.6 dB.

As reported in the last article (Ref. 1), Helios-2 entered a blackout period on June 2, 1977. The exit from blackout occurred on June 28, 1977, pass 531 over DSS 63 (Spain). On this first day, the downlink signal would not hold solid receiver lock and no telemetry was processed until the next day. On June 29, the first data (8 bits per second) were processed over DSS 63 with the spacecraft at a Sun-Earth-Probe (SEP) angle of 0.868 degrees. The first uplink was performed on July 2, 1977, over DSS 63. Also, the Faraday Rotation Experiment and the Solar Wind Experiment (Ref. 1) were resumed following Helios-2 exit from blackout. The progress of these experiments will be discussed later in this article.

Overall tracking time for both Helios spacecraft for this period is shown in Table 1.

III. Special Activities

A. DSN Mark III Data Subsystems (MDS) Support of Helios

Since the last article, which reported that DSS 14 was undergoing MDS implementation, DSS 14 has entered the MDS testing and training phase (as of June 24, 1977). The planned combined data flow test (Ref. 1) was canceled and each project proceeded with its own training program. The first Helios demonstration pass was conducted on June 25, 1977. Since that time, a total of 24 demonstration passes have been conducted over DSS 14. Overall, these tests have been a success with the majority of problems being operational rather than equipment related. These operational difficulties have at present been minimized through continued training. As a result, the station personnel have become highly knowledgeable of the MDS system and its operation. As of August 1, 1977, DSS 14 has been placed under configuration control for Helios operational support, and the test and training period for Helios was concluded. The standard Helios configuration for DSS 14 is shown in Figure 1.

On July 15, the Australian conjoint complex (DSS 42/43) was taken down for MDS implementation. They are scheduled to begin their MDS test and training phase on September 30, 1977. The progress of DSS 42/43 will be covered in the next article of this series.

B. Radio Science Activity

As mentioned in the last article, the scientific interest in this phase of the Helios-2 mission is very high (May through October 1977). The two experiments being conducted are Faraday Rotation and Solar Wind. These were discussed in detail in the last article. This article will give the status of each. Also, some preliminary results of the data collection accomplished to date will be presented.

The Faraday Rotation Experiment was performed by making solar occultation measurements. These measurements were made in two phases, entry and exit. The entry phase occurred

from May 15, 1977 through June 2, 1977. Tracking coverage for the entry phase was provided by DSS 43 and DSS 63. Two types of data, polarization angle and ellipticity, were collected during these passes. Of the 17 passes over DSS 43, 10 resulted in good polarization data, but none for ellipticity. Over DSS 63, from a total of 12 passes, 8 provided good polarization data while 5 provided good ellipticity data. Overall coverage of this entry phase went well, and the data are in the final stages of processing at this time.

The exit phase of the experiment occurred from June 26, 1977 through the present. Through this phase, coverage has been provided with 18 passes from DSS 14, 17 passes from DSS 43, and 19 passes from DSS 63. No processing of data has occurred so far, but preliminary observations of the data packages show some poor stripchart labeling and tape handling from DSS 14, which may hinder analysis. The other stations' packages appear in order and should not present any difficulties. Procedures will be refined so that the above problems will not occur in the future.

The other major experiment reported earlier (Ref. 1) was the Solar Wind Experiment (SWE). This experiment has been progressing rather smoothly. The experimenters have had many tracks and the coverage has been excellent, especially near the Sun. Over 100 hours of simultaneous two-station, long baseline observations of Helios-1 and Helios-2 high-rate doppler residuals have been collected. The experimenters have found that the one-way data contain the effects of instability of the on-board oscillator as well as the solar wind. They have come up with a scheme to mathematically remove the effects of this oscillator instability while providing a measurement of solar wind velocity at the same time. The experimenters have also begun a series of spectral broadening and one- and two-station intensity scintillation observations at DSS 13 and 14. They have demonstrated that the solar wind velocity can be measured using DSS 13 and 14 intensity scintillations, and plan on doing so for the heliocentric distance range of 5.5 - 25 R_{\odot} . These measurements will complement the SWE and allow comparison of solar wind velocities determined by two different methods. To date, the experimenters have collected DSS 13 and 14 data on 11 days and a more extensive set of single-station spectral broadening data.

Both of the above experiments will be reported on in the next article, when further results will become available.

Reference

1. Goodwin, P. S., Burke, E. S., and Rockwell, G. M., "Helios Mission Support" in the Deep Space Network Progress Report 42-40, Jet Propulsion Laboratory, Pasadena, California, June 15, 1977.

Table 1. Helios Tracking Coverage

Month	Spacecraft	Station Type	Number of Tracks	Tracking time (Hours, min)
June	Helios 1	26 m	57	393:56
		64 m	2	8:22
	Helios 2	26 m	1	5:55
		64 m	16	29:43
July	Helios 1	26 m	47	302:24
		64 m	2	8:18
	Helios 2	26 m	5	4:44
		64 m	58	221:22

**ORIGINAL PAGE IS
OF POOR QUALITY**

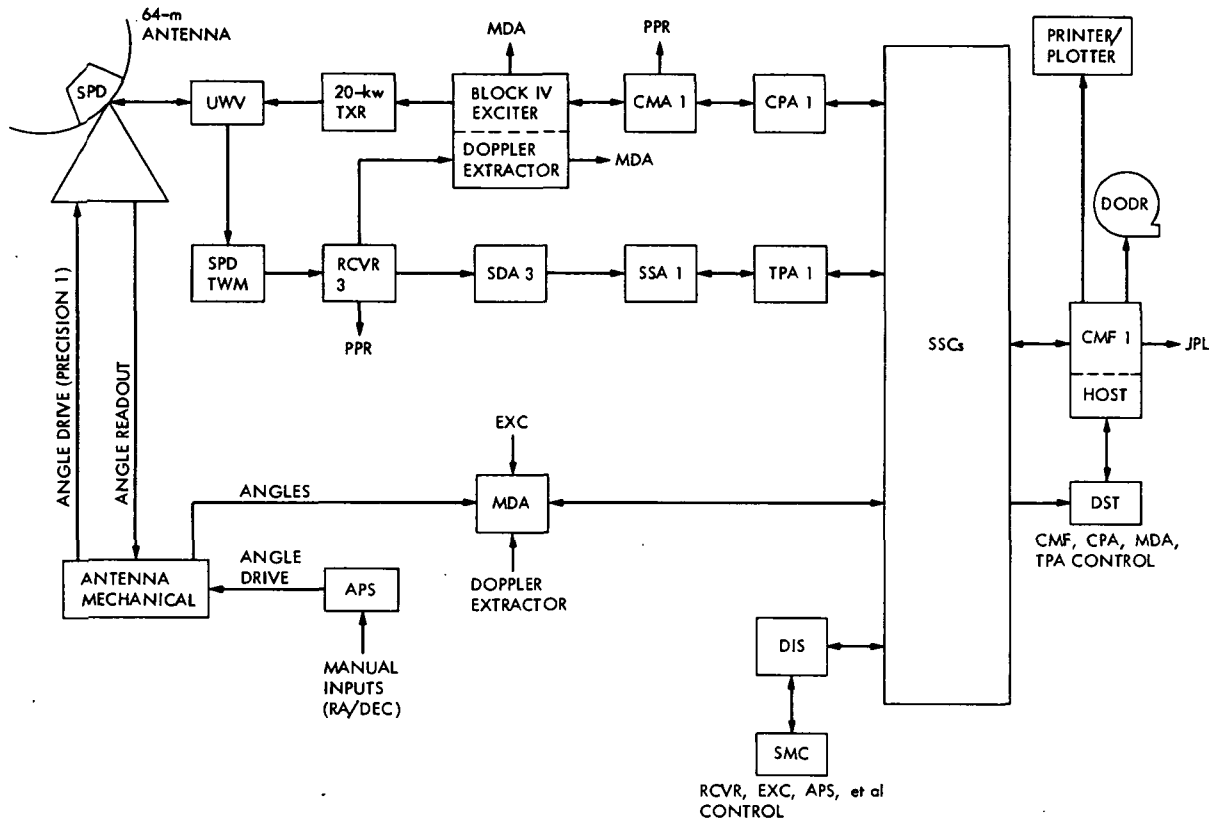


Fig. 1. Standard Helios configuration for DSS 14

Near Sun Ranging

A. I. Zygielbaum

Communications Systems Research Section

Near superior solar conjunction, radiowaves traveling between Earth and a spacecraft graze the sun. Ranging, or determining the round-trip radio signal propagation time, provides measurements of signal delay induced by the solar corona and gravity field. This article describes techniques which enhance range data quality during the harsh signal conditions existing at solar conjunction.

I. Introduction

During superior solar conjunction, the line of sight between the Earth and a spacecraft passes near the sun. Radio signals traveling between the Earth and spacecraft pass through and are perturbed by the solar corona and gravity field. Measurements of the round-trip propagation time of these signals can be used to determine the density of the solar coronal plasma and to test theories of relativity. Unfortunately the harsh signal conditions at solar conjunction hinder the ranging systems making the measurements.

This article discusses the special techniques required for successfully obtaining range data when the signal raypath intersects the solar corona. As an introduction to this subject, Section II is a brief description of DSN ranging equipment while Section III describes the sun's effect on the signals to be processed. The reader familiar with ranging may skip either of these sections without loss of continuity. Section IV describes techniques to overcome the signal degradation at solar conjunction. The last section, V, presents some of the Viking results achieved through these techniques.

II. The Equipment — An Overview

DSN Ranging currently employs the binary sequential technique first described by Goldstein (Ref. 1). A detailed description of this technique may be found in Reference 2. The brief description that follows is presented to facilitate reading of this article.

Range from a DSN tracking station to a spacecraft is measured by means of a ranging signal or code consisting of a series of coherently related squarewaves phase modulated onto a S-band carrier. The resulting signal is transmitted from the tracking station to the spacecraft where the ranging signal is coherently detected and remodulated onto a different but frequency coherent S-Band carrier and also, in the case of spacecraft such as Viking, a frequency coherent X-Band carrier for transmission back to earth. Due to the round-trip propagation time a phase offset will exist at the tracking station between the transmitted and received range codes. This offset translates into a measure of range. The periods of the different squarewaves or codes in the series are binary multiples of each other. Range measurement precision is established by the shortest period code used while the longest period limits the ambiguity with which the range is measured.

In the absence of relative motion between the spacecraft and tracking station, the phase offset between the transmitted and received range codes could be measured by direct cross-correlation. Motion of the Earth and spacecraft causes a frequency change, or Doppler shift, in the received signal precluding such direct correlation. A local model of the received signal is therefore constructed by generating a range code which is phase and frequency locked to the transmitted code until some time, designated $T\phi$. At $T\phi$, the frequency of the local model is modified by an amount equivalent to the Doppler shift. The range measurement is now obtained by determining the phase difference between the local model and received code. This phase offset represents the backward-looking time-of-flight existing at time $T\phi$.

III. The Sun's Influence on Ranging

The solar corona and gravitational field perturb the measurement of spacecraft range. Solar influence on ranging is characterized by three effects: interaction of radio waves and the solar plasma, direct reception of the sun's microwave emissions, and relativistic propagation delay induced by gravity.

The solar corona is a plasma whose density is inversely proportional to radial distance from the sun's center. According to Tyler, et al. (Ref. 4), the coronal plasma density may be modeled by:

$$N_e = \left[\frac{2.99}{\rho^{16}} + \frac{1.55}{\rho^6} \times 10^{14} + \frac{3.44}{\rho^2} \times 10^{11} \right] \times \left[\cos^2\theta + \frac{1}{64} \sin^2\theta \right]^{\frac{1}{2}} \quad (1)$$

where

N_e is in electrons/m³
 ρ is radial distance in solar radii
 θ is the solar latitude

Given a $1/\rho^6$ dependence far from the sun and $1/\rho^{16}$ near the sun, plasma effects grow enormously as the sun is approached.

Plasma in the solar corona distorts the velocities of radio-waves. The predominant effect is an increase in the signal's phase velocity and a corresponding decrease in the signal's group velocity. In a tenuous plasma the phase and group velocities, respectively V_p and V_g , are given approximately by (Ref. 3):

$$V_p = c \left(1 + \frac{1}{2} \frac{\omega_p^2}{\omega^2} \right) \quad (2a)$$

$$V_g = c \left(1 - \frac{1}{2} \frac{\omega_p^2}{\omega^2} \right) \quad (2b)$$

where

$$\omega \gg \omega_p$$

ω_p is the plasma frequency, a term proportional to plasma density

ω is the carrier frequency

c is the speed of light in a vacuum

Because the range code propagates at the group velocity, the increase in range delay due to the corona is inversely proportional to the square of the carrier frequency. This dependence of range delay on frequency provides an important mechanism for measuring the electron density in the ray-path. For example, the Viking spacecraft radio system generates S- and X-band downlink (spacecraft to earth) carriers by multiplying the received uplink (earth to spacecraft) carrier frequency by 240/221 and $(11/3) \times (240/221)$, respectively. The difference between the S- and X-band range delays can be converted to measured downlink integrated columnar electron density by (derived from Ref. 3):

$$I_D = \frac{\Delta s x f^2 G^2 c}{2A} \frac{1}{1 - (3/11)^2} \text{ electrons/m}^2 \quad (3)$$

where

$c = 2.998 \times 10^8$ m/sec
 $f \approx 2.112 \times 10^9$ Hz uplink carrier frequency
 $G = 240/221$ Transponder S-band ratio
 $A = 20.15$ Conversion Constant
 $\Delta s x =$ S-X Range in seconds

The determination of I_D provides a means for separating plasma induced delay from range delay measurements. I_D is also converted, through additional analysis, into N_e , the measure of electron density, for comparison against solar models such as (1) above.

Plasma dynamics or turbulence causes scintillation of traversing radiowaves, which is manifested as broadening of the signal carrier spectrum. The tracking station and spacecraft

ORIGINAL PAGE IS
OF POOR QUALITY

receivers must therefore contend with both degraded carrier signal-to-noise ratio and randomly gyrating carrier frequency. Since the downlink S- and X-band carrier frequencies are derived from the uplink received carrier, the downlink jitter has two components, the scintillation encountered during downlink propagation and the uplink jitter multiplied by 240/221 and S-band and $(11/3) \times (240/221)$ on X-band. The magnitude of the scintillation effect appears to be proportional to $1/\rho^{1.5}$ far from the sun and $1/\rho^{3.5}$ near the sun (Ref. 5). From practical experience, scintillation sets the minimum sun-earth-probe-angle limit at which ranging can be accomplished.

The sun is also a powerful microwave emitter and can increase a tracking station's noise temperature. Using an empirical technique, S. T. Rockwell derived a formula for the degradation in signal to noise ratio due to front end noise temperature (Ref. 6):

$$\Delta \text{SNR}(\text{db}) = 10 \log \frac{[(T_Z + T_{EL})^2 + T_{SUN}^2]^{1/2}}{T_Z + T_{EL}}$$

where

$$\begin{aligned} T_Z &= \text{noise temperature with antenna at zenith} \\ &\quad \text{(measured)} \\ T_{EL} &= 25.9e^{-.066(EL)} \\ &\quad \text{for 64m station where } EL \text{ is elevation in degrees} \\ T_{SUN} &= 6.1 \exp [5.2/(SEP + .34)] \end{aligned}$$

where SEP is sun-earth-probe angle in degrees.

This model is apparently a good approximation at SEP's greater than about 1.5° . Within 1.5° , disturbances in the antenna radiation pattern — known as quadrupod effects — can cause significant error in the term T_{SUN} . Whenever the sun appears in one of the sidelobes caused by an antenna's quadrupod structure, the system noise temperature will become much larger than T_{SUN} predicts.

IV. Operations

In the previous section, the modelable plasma/solar effects were discussed. The variations in these effects are difficult or impossible to model. It has been stated that the mean and σ of any measurement of coronal density are the same (Ref. 5). Under such dynamic conditions, only real-time monitoring can assure optimal range data return. Therefore, this section is devoted to the operations required to optimize near-sun ranging.

A. Station Optimization

The least sophisticated method to overcome low signal-to-noise ratios is to use as much power as possible in the signal. During the Viking superior solar conjunction 100 Kw was the typical DSN transmitter power. Use of high power improves overall system performance and assures that the spacecraft receiver stays in lock.

Another "obvious" technique is to avoid ranging during periods of maximum quadrupod effect. Programs exist (Ref. 7) which when given spacecraft, sun, and tracking station geometries compute front-end noise temperature as a function of time. During Viking for instance, plots of noise temperature were used to plan Lander ranging times. Because the Lander transponders were limited to operational periods as short as 15 minutes, such planning was vital to the success of the conjunction experiments.

Apparent carrier frequency spread due to scintillation, as well as degraded SNR, combine to cause the DSN receiver to "drop lock" and slip cycles. A large static phase error (SPE) in the receiver's phase-locked-loop will cause these slips to be in only one direction. Because the ranging system is "Doppler rate-aided", the result, at best, is an erroneous drift in range derived measurements of plasma variation and, at worst, is a loss of ranging due to smearing of the code. If the SPE is kept near zero, the slips occur with equal probability in either direction and become equivalent to random noise. The SPE is maintained near zero by using a Programmed Local Oscillator (PLO), driven by predicts, to track the spacecraft signal within the receiver pass-band. An alternative is to approximate the predicts with ramps generated by the Block IV receiver's Programmed Oscillator Control Assembly (POCA).

Carrier scintillation must also be compensated by adjusting the receiver carrier loop bandwidth. The 10 Hz bandwidth normally used for tracking is inappropriate when the carrier is spread over 100 Hz. A wider bandwidth can, however, degrade an already poor SNR. The optimum setting can be based on spectral broadening measurements.

B. Ranging System Optimization

Ranging yield can be improved through the use of multiple range acquisitions. When operating near the sun, the success probability of any particular acquisition is low. If the acquisition time is short, random noise alone limits the success probability. If the acquisition time is increased to combat the additive noise, receiver cycle slips distort the local reference code through the doppler rate-aiding mechanism and severely corrupt the range measurement. As noted above, the interference from the sun is highly dynamic, and may be quiet one moment and severe the next. In practice the acquisition time is

set to a compromise value and pipelining used to increase the number of acquisition attempts and hence the chance of at least one successful acquisition during the pass. Multiple measurements at a time of large uncertainties also confirm system operation and increase confidence in the data obtained.

During solar conjunction, the apparent modulation phase jitter can exceed the period of the usual 2 μ sec high frequency range code. Hence a code of larger period must be used to prevent loss of data. The Mu-II Ranging System can use any code from 8 MHz to 2 Hz as a first code. Initial codes as long as 32 μ sec were used for the Viking conjunction. A longer initial code minimizes the number of components, thereby minimizing the chance of a component error within an acquisition.

Short term variations in the coronal plasma can be observed by monitoring the resulting phase changes of the returned range code. Because the highest frequency code allows the most accurate measurements of these changes, the highest frequency code is sometimes mixed with the lower frequency components of the range acquisition sequence. Its phase offset is thereby continuously measured.

Near the sun this practice becomes a liability. Combining two range codes splits the available power between them and complicates cross-correlation of the received range code against the local model. The effect is to decrease the probability of correctly measuring the phase of the lower frequency code thus jeopardizing acquisition success. The Mu-II is normally operated during solar conjunction with code mixing disabled during the acquisition.

C. Post-processing

Non-real time processing can be used to both improve and verify the quality of range data. The more general of these techniques is "pseudo-DRVID" (Ref. 8). Applicable to either the Mu-II or Planetary Ranging Assembly (PRA), pseudo-DRVID compares the relative change in spacecraft range indicated by successive range points to the change measured by the integrated Doppler cycle count. The degree of agreement is good evidence of the reliability of the range data.

Beyond verification, range data can actually be improved through the use of a maximum likelihood technique. Reference 2 explains that the ranging algorithm used by both the Mu-II and PRA measures the phase of only the highest frequency component of the range acquisition. The contributions of all subsequent components are determined merely from the sign of their inphase correlation voltages. Information contained in the quadrature channel is ignored.

A better algorithm for range determination was described by Erickson and Layland in Reference 9. Their maximum likelihood method treats the phase measurement of each component (within the limits of its ambiguity and resolution) as an independent determination of range. The algorithm uses the Mu-II correlation voltages to reconstruct the correlation function between the local reference code and all possible delays of the received code. Assuming that additive white gaussian noise disturbs the received code, the correlation function can be treated as a likelihood function, e.g., the a posteriori probability of all range values. After accounting for the delays introduced by the Mu-II algorithm's manipulation of the reference code, the likelihood functions of all components are directly added to produce a likelihood function for the range determined by the entire acquisition. The most probable range point is selected at the peak of this combined function. According to Timor in Reference 10, an improvement of up to 1.5 dB in ranging signal-to-noise ratio can be obtained with this technique.

In addition to providing a range number, the maximum likelihood program plots the acquisition likelihood function. With some knowledge of the process, the form and symmetry of this plot can be used as an indicator of range quality. Analysis by this technique offers an additional evaluation tool to the range data user or experimenter.

V. The Results — 1976 Viking Superior Solar Conjunction

The latest use of the techniques espoused here was in support of the 1976 Viking Superior Conjunction Experiments. Both the solar corona and the relativity experiment involved ranging under very adverse conditions. The quality of the range data is reflected in Figure 1 showing electron density, determined by S- and X-band range, plotted versus the ray-path's normal distance from the sun (Ref. 4). Individual points depict 312 independent acquisitions. The solid line is the density predicted by equation (1). Note that the highest data point was derived from a successful Mu-II range acquisition at $\sim 57^\circ$ SEP. The S-X differential was just greater than 61 μ sec. This point represents the closest measurement (in terms of SEP) of coronal total columnar electron density ever made.

Viking also afforded the opportunity for the best test to date of Einstein's Theory of General Relativity. The interested reader is directed to Reference 11 for more information. The precision of the Mu-II ranging system was spectacularly demonstrated. Figure 2 shows a linear fit to residuals obtained by differencing range data (corrected for the coronal plasma effect) from predicts generated by Dr. Irwin Shapiro and his associates at MIT (a D.C. bias has been removed). The points

are scattered about the line by about 95 nanoseconds. This 15 meter scatter will, hopefully, be improved with more processing. The data does, however, yield a measurement of the relativistic induced delay good to .5% – a two-fold improvement over any previous experiment and a four-fold improvement over any previous test involving a spacecraft.

Figure 3 shows the precision during a single tracking pass. Unfettered with day to day plasma variations and ephemeris errors, the residuals show a scatter of less than 10 nano-

seconds. Ten nanoseconds out of a round-trip-light-time of nearly 2500 seconds is equivalent to five parts in 10^{12} , which, according to Dr. Shapiro of MIT, is the most precise measurement of distance yet made (Ref. 12).

The Viking experience does not signal the end of solar conjunction experiments. It is hoped that this paper will serve as a starting point and aid to the ranging experimenters on Voyager and later missions.

Acknowledgements

The author respectfully and gratefully acknowledges the pioneering work of W. L. Martin in both the field of ranging as well as superior conjunction strategy. At the same time the outstanding support received from the Mu-II operators at DSS 14 as well as the PRA operators and other support personnel throughout the DSN cannot be overlooked. The success of any spacecraft experiment is dependent upon the cooperation of many diverse people. The success of the Viking relativity and corona experiments is a tribute to all involved.

**ORIGINAL PAGE IS
OF POOR QUALITY**

References

1. Goldstein, R. M., "Ranging With Sequential Components," Space Programs Summary 37-52, Vol. II, pp. 46-49, Jet Propulsion Laboratory, Pasadena, Calif., July 31, 1968.
2. Martin, W. L., A. I. Zygielbaum, "Mu-II Ranging," Technical Memorandum 33-768, Jet Propulsion Laboratory, Pasadena, Calif., May 15, 1977.
3. MacDoran, P. F., "A First Principles Derivation of the Differenced Range Versus Integrated Doppler (DRVID) Charged Particle Calibration Method," Space Programs Summary 37-62, Vol. II, pp. 28-34, Jet Propulsion Laboratory, Pasadena, Calif., March 31, 1970.
4. Tyler, G. L., J. P. Brenkle, T. A. Komarek, A. I. Zygielbaum, "The Viking Solar Corona Experiment," Journal of Geophysical Research, to be published this year.
5. Private Conversation with Dr. G. L. Tyler of Stanford, June 7, 1977.
6. Rockwell, S. T., "A Model of SNR Degradation During Solar Conjunction," The Deep Space Network Progress Report 42-38, January and February 1977, Jet Propulsion Laboratory, Pasadena, Calif., April 15, 1977.
7. Huneycutt, B. L. "QUADEF - Program to Predict DSS Antenna Noise Temperatures During Near-Sun Tracking," JPL Interoffice Memo 3395-77-029, Jet Propulsion Laboratory, Pasadena, Calif., February 9, 1977 (JPL internal document).
8. Spradlin, G., R. Schlafer, "Training Module 1: The DSN Ranging Subsystem," JPL Interoffice Memo 421G-75-066, Jet Propulsion Laboratory, Pasadena, Calif., April 2, 1975 (JPL internal document).
9. Erickson, D. E., J. W. Layland, "An Experiment in Remote Monitoring of Mu-Ranging Operation at Mariner Mars 1971 Superior Conjunction," The Deep Space Network Progress Report for March and April 1973, JPL TR-32-1526 Vol. XV, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1973.
10. Timor, U., "Sequential Ranging With the Viterbi Algorithm," The Deep Space Network Progress Report for January and February 1971, JPL TR-32-1526, Vol. II, Jet Propulsion Laboratory, Pasadena, Calif., April 15, 1971.
11. Shapiro, I. I., R. D. Reasenberg, P. E. MacNeil, R. B. Goldstein, J. Brenkle, D. Cain, T. Komarek, A. Zygielbaum, W. F. Cuddihy and W. H. Michael, Jr., "The Viking Relativity Experiment," Journal of Geophysical Research, to be published this year.
12. Statement released by Dr. I. I. Shapiro of the Massachusetts Institute of Technology at the Viking Press Conference, January 6, 1977, Jet Propulsion Laboratory, Pasadena, Calif.

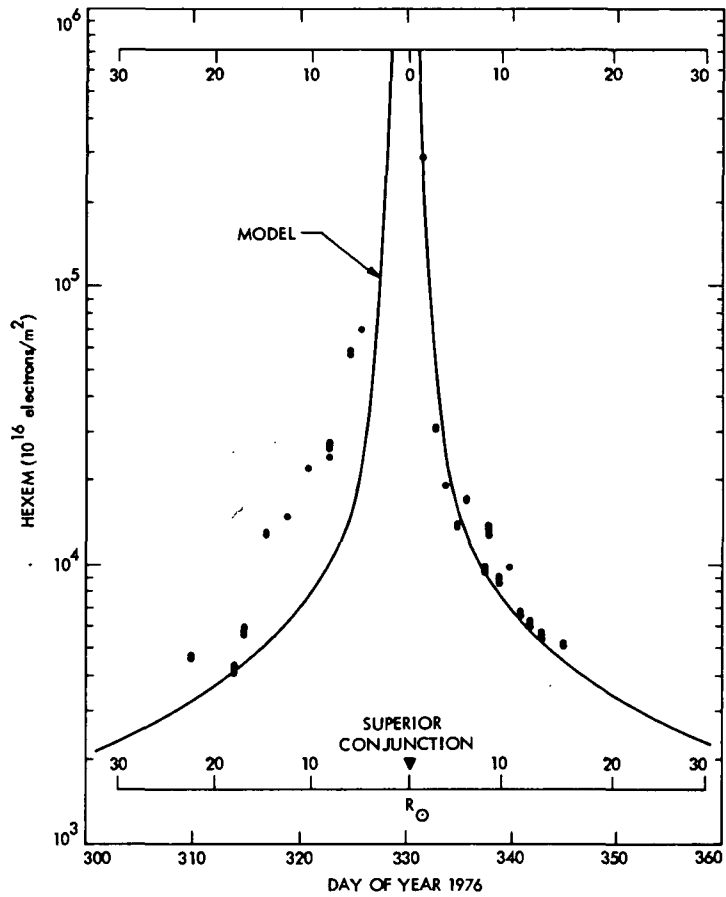


Fig. 1. Integrated electron density measured by Viking S-X range (R_{\odot} = solar radius)

ORIGINAL PAGE IS
OF POOR QUALITY

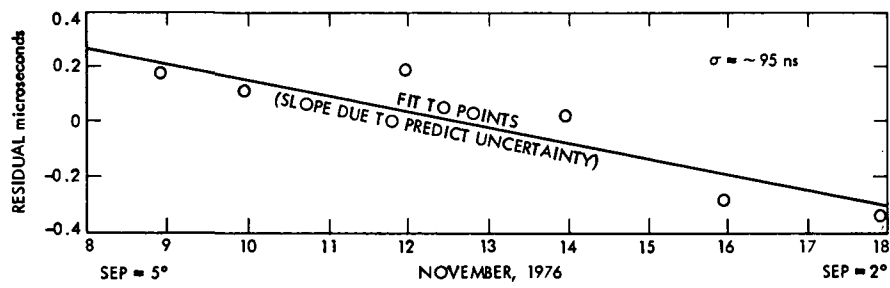


Fig. 2. Inter-day range residuals, Viking Lander 2

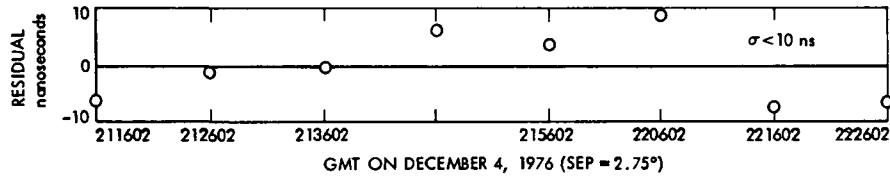


Fig. 3. Intra-day range residuals, Viking Lander 1

Feasibility Study of Far-Field Methods for Calibrating Ground Station Delays: An Interim Report

T. Sato

Radio Frequency and Microwave Subsystems Section

This interim report presents the results of a study to survey and arrive at cost/performance estimates of various methods of implementing far-field calibrations of ground station delays. Both direct and indirect methods for far-field calibrations are discussed.

I. Introduction

This is an interim report of a study to survey and arrive at cost/performance estimates of the various methods of implementing far-field calibration of ground station delays.¹ Determination of ground station delays is of vital importance in the overall ranging accuracy problem. Ideally, the calibration scheme should involve a radiator of electromagnetic waves in the far field of the antenna thus minimizing the effects of multipaths, particularly those involving antenna structural elements. Experimental verification of this approach was performed by Otoshi and Brunn (Ref. 2) utilizing the Viking Spacecraft as the signal sources. While the results were promising, the essentially one-shot nature precludes using such spacecraft over extended periods of time for a deep experimental investigation of the significant factors causing the station delay uncertainties.

This study will address itself to both the direct and indirect approach. The direct schemes will involve placing a transponder at or beyond the classical far-field distance (Ref. 3)

¹The initial report on this study has been presented in Ref. 1.

given by $2D^2/\lambda$ where the antenna diameter D and wavelength used λ are in the same units. Just precisely where the transponder is located will depend on other factors as well as depending on what platform is used to transport the transponder. The indirect approach will be dependent on finding a valid means of transforming measurements made utilizing a transponder located in the near-field region to those results obtainable if it were possible to locate the transponder in the far field.

II. Antenna Parameters

The criterion of whether a source is located in the far field of an antenna is determined if the antenna-source distance R exceeds

$$R = 2D^2/\lambda$$

where D is the diameter and λ is the wavelength received by the antenna measured in the same units. This relationship defines the condition where a wave incident on the antenna does not deviate more than $\lambda/16$ across the aperture. The

source is considered to be a point source thus radiating a spherical wave, and it is this curvature that is mismatched to the plane of the antenna aperture.

The antenna diameters of greatest interest are 26 and 64 meters. Frequencies of interest are 2295 and 8415 MHz. Table 1 shows the far-field distance for the two antenna diameters at the two frequencies.

It is immediately evident that a collimation tower can be used in the far field effectively at 2295 MHz with a 26-m antenna, but it cannot be used for a 64-m antenna. At 8415 MHz, neither the 26-m nor the 64-m antenna's can utilize collimation towers located in the far field because of the tower height required for a reasonable elevation angle of greater than 5 deg.

III. Far-Field Platforms

Discounting the situation of investigating sources of antenna delay variations solely of a 26-m diameter antenna at 2295 MHz, it is clear from Table 1 that the transponder must be attached to some type of free-moving platform because of the great distance and height required. A problem of considerable complexity is presented, that of determining with sufficient accuracy the near-instantaneous location of this platform. Range to a collimation tower is easier and less expensive to determine and can be repeatedly surveyed to reduce range errors. The aircraft, balloon, and satellite approaches discussed in the next three sections all require an independent ranging system.

A ranging scheme has been suggested using a tracking laser system (Ref. 4) similar to those used in connection with extremely precise tracking of geodetic satellites. Results from the NASA LAGEOS Project, which requires the precise ranging and tracking of a satellite in earth orbit, indicate that accuracies of better than 5 cm are possible.

The rationale of using lasers is to operate at a wavelength sufficiently far removed from microwaves to avoid those effects deleterious to microwaves but not to lasers. Lasers in turn have problems, particularly clouds, so any advantages won in one area may be lost in becoming more weather dependent.

These high-performance laser trackers are very expensive because of their ability to track rapidly moving targets with great precision. Depending on the degree of computer control and various safety equipment desired, the cost of a laser tracker can exceed \$250,000.

A. Aircraft

A high-performance jet-powered aircraft can be used as a transponder platform. The outstanding advantage of an aircraft is the positive control of the transponder both as to its position and operation. Precise maneuvers can be executed along a flight path thereby somewhat easing the range determination problem. Further, the danger of losing a transponder or suffering a malfunction is eased since the entire system is retrievable at anytime.

The main disadvantage of the aircraft is the maximum altitude (Ref. 5) achievable being limited to about 13.5 km (8.39 miles). This altitude translates into an elevation angle of about 12 deg with a slant range of 65 km. A 26-m diameter antenna operating at 8415 MHz could satisfactorily use the aircraft-mounted transponder since the elevation angle is about 18 deg for the necessary far field slant range of 40 km at an altitude of 13.5 km. A 64-m diameter antenna requires a minimum far-field slant range of 238 km resulting in a 3-deg elevation angle, which is below the present operating limit set by transmitter and land-mask considerations.

The cost of developing an aircraft-based transponder platform depends on many factors. Methods of mounting as well as the type of aircraft selected all strongly influence the cost. Special techniques must be employed to minimize bothersome reflections that can cause a multipath error. The total cost of an aircraft-based program can be quite low if leased aircraft are used. Some planes suitable for this application are available for about \$1,000/h.

Use of powerful laser ranging systems on manned aircraft is always cause for concern. An unmanned jet-powered aircraft is being developed for military use that can be useful for transporting a transponder package. The USAF Compass Cope (Ref. 6) can carry a payload of 500 kg to a height above 18 km. One of the projected uses of Compass Cope by the USAF is as a communications relay.

B. Balloons

Balloons have in the past been used to calibrate antenna systems and constitute a relatively inexpensive means of lifting lightweight targets or transponders to moderate heights. The lifting of relatively sophisticated instrument packages such as a ranging transponder and associated equipment requires a fairly large balloon system, certainly much larger than those Rawinsonde systems routinely used by the Weather Bureau.

The Stratoscope Project (Ref. 7) utilized a free balloon system capable of lifting 3600 kg of equipment to a height in excess of 25 km. These balloons are expensive, quite difficult to launch, and require a considerable amount of special han-

ding equipment as well as a highly trained operations crew. Further, use of these large balloons is highly weather dependent and recovery of the transponder for subsequent flights is problematical at best. Flight paths of the balloon can often be quite unpredictable as with any free flying craft.

An interesting possibility, however, is the powered balloon. This is a free flying but powered balloon and therefore has a limited maneuvering capability. Such a craft should be able to hover over the same sub-earth point making a suitable transponder platform for a 64-m antenna at 2295 MHz, but would still be of little use at 8415 MHz.

The cost of balloon systems for all but the very small weather balloon would be expected to be rather expensive, about \$25,000 each, and are not reusable. Powered balloons would cost more than free balloons but would be less likely to be lost along with the transponder package than a free balloon.

C. Satellites

A satellite can be used as a transponder platform provided it does not move at high angular velocities. Geostationary satellites are in orbit such that they appear to the antenna as essentially fixed in the sky. Because of this fact, such a transponder platform will be available anytime of the day and could be placed so that favorable elevation angles can be achieved. The orbital height is determined by the physical constants of the earth and is about 36,000 km in altitude above the earth, nearly a tenth of the distance to the moon. Accurate ranging of the transponder platform at these distances becomes considerably more difficult than at aircraft or balloon ranges if the goal of determining this distance with an accuracy of better than 50 cm is to be achieved. Accuracies in the region of 40 cm have been achieved at lunar distances, but only after the accumulation of considerable data.

Costs of satellite systems have been determined and a Cost Estimate Relation (CER) derived by Hadfield (Ref. 8). Applying these CERs to a geostationary satellite in the 140-kg weight class, the cost estimate is about \$5,000,000. These CERs are relatively insensitive to the type of electronics flown in the satellite and one may assume that a ranging transponder is not extraordinarily complicated nor sophisticated enough to alter these CERs.

An interesting variant to using satellites in this range delay problem is to use the ALSEP (lunar package transmitter) as a far-field source and use a delay comparison method. The delay of a horn antenna is measured and is compared with a medium size (9-m diameter) antenna such as the ARIES antenna while both are aimed at a collimation tower located in the far field of the medium sized antenna. The derived delay for the medium sized antenna is now used to view ALSEP simultane-

ously with a 26-m or 64-m antenna while located adjacent to the large antenna. The outputs of both antenna are then suitably processed, thus yielding a difference in delays through the respective antenna microwave optics. The delays through the major electronic components in the implementation of this method will be determined using existing techniques (Ref. 9).

IV. Near-Field Methods

The previous sections describing the various far-field methods have shown that except for the ALSEP bootstrap, all of these approaches require an independent means of accurately determining distances preferably without the use of microwave systems. The laser ranger needed is quite complex and therefore expensive since it must be agile enough to track distant and often rapidly moving targets.

A collimation tower by its very nature provides a very stable platform and whose distance can be determined by modern conventional surveying to a high order of accuracy. Here again lasers can be used but they will not be tracking systems and therefore will be much less expensive.

Collimation towers are relatively inexpensive and a number of them can be erected at various distances. For a 26-m diameter antenna, a far-field and near-field tower can both be erected (Ref. 10) and studies performed to arrive at suitable models that allow near-field measurements to be transformed to those simulating far-field measurements. The cost of a 30-m collimation tower of the type used as broadcast antennas with an equipment elevator can be erected for about \$20,000 (Ref. 11). The exact total cost will depend on its location with such factors as earthwork requirements, type of surface material, safety equipment, and other logistical factors.

At Goldstone, the various elevation angles are listed for both the existing 2-m collimation tower (Ref. 12), and the proposed 30-m collimation tower in Table 2.

The main disadvantage of using collimation towers is that measurements must be made at low elevation angles. Antenna structural sag as well as ground reflections may pose problems. A reflection abatement program forms an integral part of this near-field method. As mentioned earlier, the near-field method depends upon finding a valid means of transforming measurements made with the transponder located in the near field to those results obtainable if it were possible to locate the transponder in the far field. In concept it would appear that the same transformation technique used in near-field gain measurements can be applied. This approach needs to be investigated thoroughly by an antenna expert and may prove to be a new separate theoretical area of study that needs to be supported and funded.

V. Conclusions and Recommendations

Table 3 summarizes the various methods for far-field ground station delay calibration methods described in this report. It provides a quick reference for purposes of cross comparisons of the salient characteristics, advantages and disadvantages of the various methods.

The study of range-delay calibrations and their uncertainties is an important topic and should be approached by various

routes. The moderate sized (26-m) antenna appears to be a very important element in this study since it is quite feasible to construct a transponder platform in the near field and one in the far field. Comparison of measurements taken with this system will verify the validity of antenna range-delay models with minimum expense. If extensive testing on this range yields an acceptable model, one may extend it to the larger (64-m) antenna, erect a collimation tower that necessarily must be in its near field, and commence study of important factors contributing to its delay uncertainties.

**ORIGINAL PAGE IS
OF POOR QUALITY**

Acknowledgment

The author wishes to thank D. Curkendall, G. Resch, and J. Williams for valuable and helpful discussions. T. Otoshi provided technical direction for this study.

References

1. Sato, T. "Feasibility Study of Far-Field Methods for Calibrating Ground Station Delays", in *The Deep Space Network Progress Report 42-37*, pp. 55-56, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1977.
2. Otoshi, T. Y. and D. L. Brunn, "Multipath Tests on 64-m Antennas Using the Viking Orbiter -1 and -2 Spacecraft as Far-Field Illuminators", in *The Deep Space Network Progress Report 42-31*, pp. 41-49, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1976.
3. Hollis, J. S., "Antenna Measurements", *Microwave Journal*, pp. 39-46, Feb. 1960.
4. Couvaut, C., "Help Aboard Key to Laser Mission Goals", *Aviation Week and Space Technology*, pp. 36-43, May 12, 1975.
5. Jane's *All the World's Aircraft*, McGraw-Hill, 1976.
6. O'Lone, R. G., "Compass Cope Seen in Anti-Tank Role", *Aviation Week and Space Technology*, pp. 60-62, June 13, 1977.
7. Ashbrook J., "The Latest Flight of Stratoscope II", *Sky and Telescope*, Vol. 39, No. 6, pp. 365-367, June 1970.
8. Hadfield, B. M., "Satellite-System Cost Estimation", *IEEE Transactions on Communications*, Vol. COM-22, No. 10, pp. 1540-1547, Oct. 1974.
9. Komarek, T. and T. Otoshi, "Terminology of Ranging Measurements and DSS Calibrations", in *The Deep Space Network Progress Report 42-36*, pp. 35-40 Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1976.
10. Rusch, W. V. T. and P. D. Potter, *Analysis of Reflector Antennas*, Academic Press, 1970.
11. Walker, A. P. Ed., *NAB Engineering Handbook*, pp. 2-41-2-43, McGraw-Hill, 1960.
12. Jackson, E. G., private communication.

Table 1. Far-field distance as a function of frequency and antenna diameter

Antenna diameter, m	$2D^2/\lambda$, km	
	2295 MHz	8415 MHz
26	11	40
64	65	238

Table 2. Elevation angles to collimation towers

Collimation tower location	DSS 13, 26 m		DSS 14, 64 m	
	Target location	Target elevation angle, deg	Target location	Target elevation angle, deg
Near field				
Existing tower $0.14 \left(\frac{2D^2}{\lambda} \right)$	2 m from ground on 176-m high hill	6.6	--	--
Proposed tower $0.10 \left(\frac{2D^2}{\lambda} \right)$	30 m from ground that slopes -2.2 deg; near same azimuth as Tiefort Mtn.	0.92	30 m from ground on 70-m hill at azimuth of 50 deg	0.88
Far field				
Existing tower Tiefort Mtn. $1.8 \left(\frac{2D^2}{\lambda} \right)$	2 m from ground on Tiefort Mtn.	0.5	--	--

Table 3. Advantages and disadvantages of methods

Method	Advantages	Disadvantages
Aircraft	Positive control of operation and of flight path; relatively inexpensive if leased aircraft used	Good only for S-band on 64-m antenna; needs laser ranger
Balloon, free	Can reach higher altitude than aircraft	Good only for S-band on 64-m antenna; needs laser ranger; requires good weather for launch
Balloon, powered	Same performance height as free balloon; limited maneuver capability, can hover over a sub-earth point	Good only for S-band on 64-m antenna; needs laser ranger
Satellite, geosynchronous	Continuously in position; in far field for both S- and X-band of 64-m antenna	Very expensive, satellite and launch cost greater than \$5M; needs laser ranger
Near-field collimation tower	Continuously in position; least expensive, \$20k per tower; easier to determine distances; does not need laser ranger	Works at low elevation angles; needs near- to far-field transform; will require ground reflection abatement

D
N78-11141

A Technique to Determine Uplink Range Calibration Due to Charged Particles

S. C. Wu and F. B. Winn

Tracking Systems and Applications Section

A technique is presented to determine uplink range change due to line-of-sight charged particles using downlink S/X range and doppler and 2-way DRVID data. The line-of-sight relative charged-particle distribution is first estimated from these data by a point-matching method. The uplink range change is then synthesized by the weighted sum of the delayed downlink range changes. Simulation analysis shows 2 to 6 cm agreement between the "true" and the synthesized uplink range changes. Demonstration using Viking 1975 radio metric data shows 4 to 22 cm consistency between the uplink range calibration solutions synthesized from S-DRVID and X-DRVID. The consistency is an order of magnitude superior to that equating uplink and downlink effects.

I. Introduction

During the previous interplanetary missions (including Viking '75) there have been no requirements for the calibration of the range change due to charged-particle effects. These effects are calibrated only for the doppler. For interplanetary tracking over ~ 10 A.U. (e.g., Voyager '77) the range change due to charged particles also needs to be calibrated. This will require both uplink and downlink range calibrations.

The radio metric data currently available are S- and X-band DRVID (Ref. 1), S/X dual-frequency doppler and range (Ref. 2). The types of these data are summarized in Table 1. S-DRVID, despite being 2-way data, provides only relative range-change information. X-DRVID, containing S-band uplink and X-band downlink, also is a relative measurement. S/X doppler is a "cleaner" downlink measurement. Although it is a relative measurement as well, it can be fitted to S/X range data (downlink) to infer absolute range calibration due to charged-particle effects. The uplink range calibration, on

the other hand, is still a problem to be solved. The determination of this quantity from the above radio metric data requires the knowledge of the line-of-sight charged-particle distribution.

Winn, et al. (Ref. 3) proposed a scheme to estimate the mean location of the charged-particle concentration along the line-of-sight under the assumptions that the charged-particle cloud is (1) singular, (2) discrete, and (3) fixed in space. The uplink calibration can then be inferred from the delayed downlink calibration. This calibration scheme is indeed an improvement over that equating uplink to downlink. However, the above three constraints limit its application.

In this article a technique is presented to determine uplink range calibration with the above first two constraints removed. This technique requires only that the line-of-sight *relative* distribution of the charged particles be fixed over a period of time. The line-of-sight is first divided into a finite number of segments. A set of constrained linear algebraic equations of the

relative charged-particle distribution is set up by a point-matching method. (Ref. 4). The solution of this relative charged-particle distribution provides the weights with which the delayed downlink calibrations are summed to synthesize the desired uplink calibrations.

Simulation analysis shows that the synthesized uplink calibration has an accuracy of better than 6 cm for a 20-minute one-way light time. A 20-cm (1σ) noise is assumed for one of the two observables to simulate the DRVID noise (Table 1). Application to the Viking 1975 radio metric data is demonstrated. The uplink calibration solutions synthesized from S-DRVID and X-DRVID show a 4 to 22 cm agreement.

II. Formulation

In the following analysis it is assumed that 2-way S-band DRVID and downlink S/X doppler and range data are available. Figure 1 depicts the geometry of the spacecraft, the deep-space station, and the charged-particle cloud. Let the charged-particle *relative* distribution be $w(t')$ at a distance t' light-minutes away from the spacecraft along the line-of-sight. This distribution $w(t')$ will be assumed fixed over the whole pass. The *absolute* distribution of the charged particles at any time t is the product of $w(t')$ and a temporal function $p(t)$. The a priori knowledge of $p(t)$ is not needed as will be seen. The uplink and downlink range changes due to the charged particles can be written as

$$R_u(t) = \int_0^T w(t') p(t - T - t') dt' \quad (1)$$

and

$$R_d(t) = \int_0^T w(t') p(t - T + t') dt' \quad (2)$$

where T is the one-way light time, and t is measured at the DSS. It is understood that $w(t')$ is non-negative and normalized:

$$\int_0^T w(t') dt' = \int_0^T |w(t')| dt' = 1 \quad (3)$$

The equation relating the uplink and downlink range changes from Eqs. (1) and (2) would be too complicated to be practical. An approximation accurate enough for practical uses can be made on the following basis. Comparing Eqs. (1) and (2) we see that for each infinitesimal component of downlink

$R_d(t)$ there corresponds an equal uplink component which will appear only at a time $2t'$ later as received by the DSS. Hence an approximate expression for $R_u(t)$ would be

$$R_u(t) = \int_0^T w(t') R_d(t - 2t') dt' \quad (4)$$

This expression contains no temporal function $p(t)$. The uplink range change $R_u(t)$ can be synthesized once the relative distribution $w(t')$ is determined.

Note that both $R_u(t)$ and $R_d(t)$ in Eq. (4) are *absolute* range changes. Since DRVID (2-way) and S/X doppler (downlink) provide only relative range changes it is desired to reduce Eq. (4) in terms of relative quantities. Let the absolute uplink and downlink range changes at t_0 be $R_u(t_0)$ and $R_d(t_0)$, respectively. Then Eq. (4) can be rewritten as

$$\begin{aligned} R_u(t_0) + f_u(t) &= \int_0^T w(t') [R_d(t_0) + f_d(t - 2t')] dt' \\ &= R_d(t_0) + \int_0^T w(t') f_d(t - 2t') dt' \end{aligned}$$

or

$$f_u(t) = \int_0^T w(t') f_d(t - 2t') dt' + R_d(t_0) - R_u(t_0) \quad (5)$$

where f_u and f_d are relative range changes with zero references at $t = t_0$. The existence of the quantity $R_d(t_0) - R_u(t_0)$ in this equation presents no difficulties in the solution of $w(t')$. Although this quantity can be eliminated by the substitution of Eq. (4) with $t = t_0$, it is desirable to retain it as an additional solve-for unknown to eliminate the effect of DRVID noise at t_0 . Elimination of this term will result in a bias in $f_u(t)$ when DRVID data at t_0 is inaccurate; this, in turn, will deteriorate the solution of $w(t')$. It should be noted that $f_d(t)$ has to be available at least $2T$ prior to $f_u(t)$ for Eq. (5) to be valid.

A simple approach to solve Eq. (5) is that of point matching (Ref. 4). The line-of-sight is divided into N segments of equal lengths. Over each segment the distribution is considered to be uniform. Hence Eq. (5) is reduced into the following summation:

$$f_u(t) = \sum_{n=1}^N W_n \int_{(n-1)T/N}^{nT/N} f_d(t - 2t') dt' + R_{d-u}(t_0) \quad (6)$$

Here, $f_d(t)$ represents the downlink S/X doppler. The (relative) uplink $f_u(t)$ is obtained from the difference¹ of S-DRVID and S/X doppler.

The unknowns W_n , $n = 1, 2, \dots, N$ and $R_{d-u}(t_0)$ can be solved by $M (> N + 1)$ equations resulting from the substitution of M different values of t in Eq. (6) subjected to the constraints reduced from Eq. (3):

$$\frac{T}{N} \sum_{n=1}^N W_n = 1 \quad (7)$$

and

$$W_n \geq 0, n = 1, 2, \dots, N. \quad (8)$$

A linear least-squares process with constraints (Ref. 5) can be applied to solve for the $N + 1$ unknowns.

Once the W_n are determined, the absolute uplink range change $R_u(t)$ can be synthesized in two steps: (1) the S/X doppler is fitted to the S/X range points to yield the absolute downlink $R_d(t)$, and (2) the uplink $R_u(t)$ is calculated by the reduced form of Eq. (4):

$$R_u(t) = \sum_{n=1}^N W_n \int_{(n-1)T/N}^{nT/N} R_d(t - 2t') dt' \quad (9)$$

III. Simulation Analysis

To illustrate the uplink range calibration technique a simulation analysis is given in this section. A temporal function $p(t)$ describing the line-of-sight charged-particle content is assumed to be that shown in Fig. 2. Also assumed are the three models of line-of-sight relative charged-particle distribution as shown in Fig. 3. Here a 20-min one-way light time between spacecraft and DSS is assumed. The first model contains a single plasma uniformly distributed over a distance of one light minute; the second contains a nonuniform plasma over 2 light minutes; and the third contains two discrete plasmas of different spreads and densities.

The simulated S/X doppler is calculated by Eq. (2) with the zero-reference set at $t_0 = 2T$. The simulated S-DRVID is calculated by summing Eqs. (1) and (2), also with $t_0 = 2T$ as zero

¹Or, alternatively, from the difference of X-DRVID and $(3/11)^2$ times S/X doppler, where 3/11 is the ratio between downlink X- and S-band carriers.

reference. The DRVID noise is simulated by adding a gaussian noise of 0.2-m standard deviation. These simulated S/X doppler and DRVID are plotted in Fig. 4.

To reduce the effect of DRVID noise, the number of equations M is made larger than the number of unknowns $N + 1 = 21$. The solutions of W_n using $M = 60$ are shown in Fig. 5 for the three models. The assumed true distributions of Fig. 3 are redrawn for comparison. The computed distributions indeed display the plasma concentrations. The larger smeared effect in Model 3 is believed due to the error in Eq. (4) with widely spread distribution.

The synthesized uplink range changes from Eq. (9) and the combined 2-way range changes are plotted in Fig. 6. The true uplink, downlink, and 2-way range changes are also shown for comparison. The synthesized quantities are essentially indistinguishable from the corresponding true quantities. The 2-way range changes obtained by simply doubling the downlinks are inserted to show how much they deviate from the true and the synthesized solutions. Table 2 compares the RMS deviation between these solutions. Synthesized results for a reduced number of equations ($M = 40$) are also included to demonstrate the convergence of solutions with increasing M .

The computed W_n with $M = 40$ are plotted in Fig. 7. The synthesized range changes are essentially identical to those with $M = 60$ (Fig. 6). A comparison of Figs. 5 and 7 indicates that the range change can be accurately determined from a not-so-accurate spatial distribution. This is due to the fact that the former is a linear combination of the latter; a first-order approximation of the distribution provides a second-order approximation of the range change. Hence, detailed resolution of the line-of-sight spatial distribution is not needed in the synthesis of uplink range change. Resolution of the order of light minutes is considered appropriate.

IV. Viking 1975 Radio Metric Data Demonstration

This section provides a demonstration for the calibration technique through applying to the Viking '75 radio metric data. Since actual uplink range changes are not known, the accuracy is estimated by comparing the results from different sets of observables undergoing the same media effect. That is, the result determined from the correlation between S/X doppler and S-band DRVID can be compared with that from S/X doppler and X-band DRVID. Hence, only passes with both S- and X-DRVID available are to be selected for the demonstration. In the following eight passes from Viking B Orbiter are selected in which the first six are for SEP (Sun-Earth-probe angle) ~ 30 deg and the remaining two are for SEP ~ 8 deg (Fig. 8).

Figure 9 displays the S/X doppler as fitted to the S/X range points. The square-wave distortion effect has been removed from each range point by adding the difference of the nearby S-DRVID mean and X-DRVID mean.² Spacecraft delay, station delay and Z-correction are not removed from these range points. Figure 10 displays the corresponding S- and X-DRVID data during these passes. S-DRVID appears to have less noise than X-DRVID.

Since the downlink S-band carrier frequency differs from that of uplink by a factor $240/221$, f_d in Eq. (6) is now given by $(240/221)^2$ times S/X doppler; f_u , as before, is given by S-DRVID subtracted by S/X doppler or X-DRVID subtracted by $(3/11)^2$ of S/X doppler. Also, since the S/X doppler data in most passes are taken once per minute the line-of-sight is conveniently divided into cells one light minute long.

Figure 11 compares the synthesized 2-way (uplink and downlink) range changes using S-DRVID and using X-DRVID. The RMS difference is 4 to 22 cm. The larger discrepancy is due to either (1) charged-particles having weak second-order dynamic that the observables contain little information about its distribution, as in Fig. 11 (d), or (2) X-DRVID being too noisy, as in Fig. 11(b) (cf. Fig. 10(b)). In all cases the consistency is better than that between 2X downlink³ (continuous lines in Fig. 11) and either of these synthesized 2-way range changes. Typically it is an order of magnitude better. Table 3 summarizes the statistic of the comparison.

It should be pointed out that on October 26 each X-DRVID point was taken 10 s after the corresponding S-DRVID point and that the synthesized range changes actually have a 10-s relative shift. The RMS difference in Fig. 11(g) would have been reduced if the same S/X doppler set had been used for both synthesized range changes. For the remaining seven passes the S-DRVID and X-DRVID were taken alternately at 1-minute intervals; once the relative charged-particle distributions were reconstructed, the synthesized range changes were calculated on the same time frame.

As an intermediate result, the line-of-sight relative charged-particle distribution is shown in Fig. 12. The distance between the spacecraft and DSS along the line-of-sight is ~ 20 light minutes for the first six passes (Fig. 12(a) to (f)), and is ~ 21 light minutes for the remaining two passes (Fig. 12(g) to (h)). For comparison, the mean plasma locations determined by the method of Ref. 3 are also included (shown as vertical arrows).

²As proposed by T. Komarek of Telecommunication Systems Section, Jet Propulsion Laboratory.

³The 2X downlink has been shifted upward by a factor $[1 + (240/221)^2]/2$ to account for the lower uplink carrier frequency.

From Fig. 12 it is observed that the computed charged-particle distributions have a concentration near 13 to 14 light minutes from the spacecraft in most cases. This agrees with the nearest-to-Sun point on the line-of-sight which is 12.8 light minutes from the spacecraft, as depicted in Fig. 8, during these passes. The better consistency among the results using S-DRVID once again indicates lower S-DRVID noise.

It should be pointed out that the computed concentration is only for the varying component of the charged-particle cloud. The concentration of the stationary charged-particle cloud is left out of the solution. This however, will not degrade the range calibration accuracy since the location of the stationary charged-particle cloud is immaterial in the synthesis of uplink range from downlink.

V. Discussions

Uplink range calibration due to charged particles can be determined from current radio-metric data. One may simply equate the uplink effect to the downlink effect by using S/X doppler and range (both downlink only). Unless the charged-particle cloud is concentrated near the spacecraft or is stationary, the accuracy is considered to be poor. With the knowledge of the Sun-Earth-probe geometry, one can improve the accuracy by equating the uplink effect to the delayed downlink, the delay t' being twice the light time between the spacecraft and the nearest-to-Sun point on the line-of-sight. The accuracy can be further improved by the mean-plasma-location approach (Ref. 3) which determines the delay t' by correlating the S/X doppler to DRVID. As the charged-particle cloud spreads out along the line-of-sight over several light minutes, the accuracy of these solutions will deteriorate and one will have to rely on the synthesis approach to maintain good accuracy of uplink range calibration.

Two-way doppler, which is the time rate of change of 2-way range, is directly available from the synthesized 2-way range change. Although DRVID does provide this information, the inherent noise prevents its direct application before undergoing proper smoothing process. The smoothing process, however, will filter out any short-term charged-particle variation effects. The reduction of 2-way doppler from synthesized 2-way range requires no smoothing process and hence, will retain these short-term effects.

VI. Conclusions

A technique to determine uplink charged-particle range calibration has been presented. The charged-particle effect is synthesized by the solution of relative charged-particle distribution along the line-of-sight. Resolution of distribution is not stringent; resolution of the order of light minutes is appro-

priate. Simulation analysis has shown an accuracy better than 6 cm for a 20-min one-way light time with a 20-cm (1σ) noise in the simulated DRVID data. Application to the Viking '75 radio-metric data has demonstrated a 4 to 22 cm consistency between the results synthesized from S-band DRVID and from X-band DRVID. The only constraint for this technique is the assumption that the line-of-sight *relative* distribution of the charged-particles be fixed over the pass.⁴ This is not an unreal-

⁴It is sufficient to assume such condition over a period $\sim 3\times$ the one-way light time between spacecraft and DSS.

istic assumption considering the ~ 10 hours required for a solar plasma to move, at ~ 500 km/s speed (Ref. 6), over a distance of one light minute.

The line-of-sight charged-particle distribution is a by-product of the calibration technique. This information is valuable in the investigation of solar plasma activities. For a detailed solar-plasma spatial structure a more accurate expression would be required in place of Eq. (4). This is currently being studied.

References

1. MacDoran, P.F., "A First-Principles Derivation of the Differenced Range Versus Integrated Doppler (DRVID) Charged-Particle Calibration Method," *Space Programs Summary 37-62*, Vol. II, pp. 28-34, March 31, 1970.
2. Madrid, G.A., "The Measurement of Dispersive Effect Using the Mariner 10 S- and X-Band Spacecraft to Station Link," *DSN Progress Report 42-22*, pp. 22-27, August 15, 1974.
3. Winn, F.B., Wu, S.C., Komarek T., Lam, V.W., and Royden, H.N., "A Solar Plasma Stream Measured by DRVID and Dual Frequency Range and Doppler Radiometric Data," *DSN Progress Report 42-37*, pp. 43-54, Feb. 15, 1977.
4. Harrington, R.F., *Field Computation by Moment Methods*, Macmillan, 1968.
5. Lawson, C.L., and Hanson, R.J., *Solving Least Squares Problems*, Prentice Hall, Inc. 1974.
6. Snyder, C.W., Neugebauer, M., and Rao, U.R., "The Solar Wind Velocity and Its Correlation with Cosmic Ray Variations and with Solar and Geomagnetic Activity," *J. Geophys. Res.*, Vol. 68, p. 6361, 1963.

Table 1. Viking 1975 radio metric data

Data type	Uplink/downlink	Absolute/relative	Typical noise level
S-DRVID	2-way	Relative	0.2 m
X-DRVID	Uplink ^a	Relative	0.2 m
S/X doppler	Downlink	Relative	0.01 m
S/X range	Downlink	Absolute	0.2 m

^aX-DRVID actually consists of uplink S-DRVID plus downlink X-DRVID that equals uplink S-DRVID with ~ 6% error

Table 2. RMS Deviation from "true" 2-way range—simulation

	Model 1	Model 2	Model 3
Synthesized 2-way (M = 60)	1.9 cm	3.4 cm	3.7 cm
Synthesized 2-way (M = 40)	3.4 cm	4.1 cm	5.8 cm
2 X downlink	75.7 cm	77.4 cm	65.8 cm

Table 3. Statistic of synthesized 2-way range changes—Viking '75 radio metric data

Pass	Date	DSS	Number of cells	Number of equations	Residual of equations		Difference from 2X downlink		(S) - (X)
					(S)	(X)	(S)	(X)	
(a)	8/16/76	43	20	31	0.49 m	0.55 m	0.39 m	0.34 m	0.08 m
(b)	8/16/76	63	20	26	0.32	0.50	0.41	0.20	0.22
(c)	8/22/76	63	20	38	0.28	0.58	1.17	1.17	0.04
(d)	9/23/76	63	20	46	0.24	0.51	1.14	0.95	0.22
(e)	8/24/76	63	20	22	0.23	0.35	0.88	0.77	0.11
(f)	8/25/76	63	20	36	0.16	0.46	1.12	1.16	0.07
(g)	10/26/76	43	21	50	0.52	0.82	2.61	2.61	0.20
(h)	11/01/76	43	21	30	0.51	0.55	1.48	1.56	0.17

(S) 2-way range change synthesized from S-DRVID

(X) 2-way range change synthesized from X-DRVID

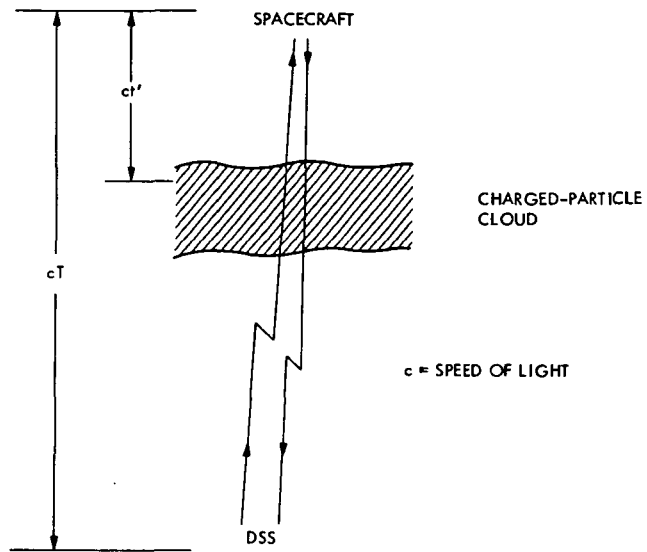


Fig. 1 Geometry of the problem

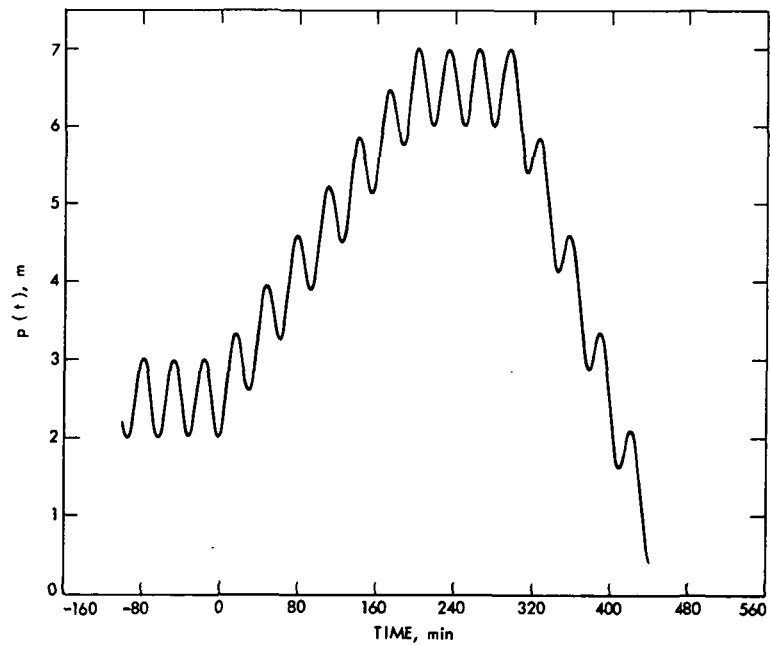


Fig. 2. Total line-of-sight charged-particle content (in terms of range change)

ORIGINAL PAGE IS
OF POOR QUALITY

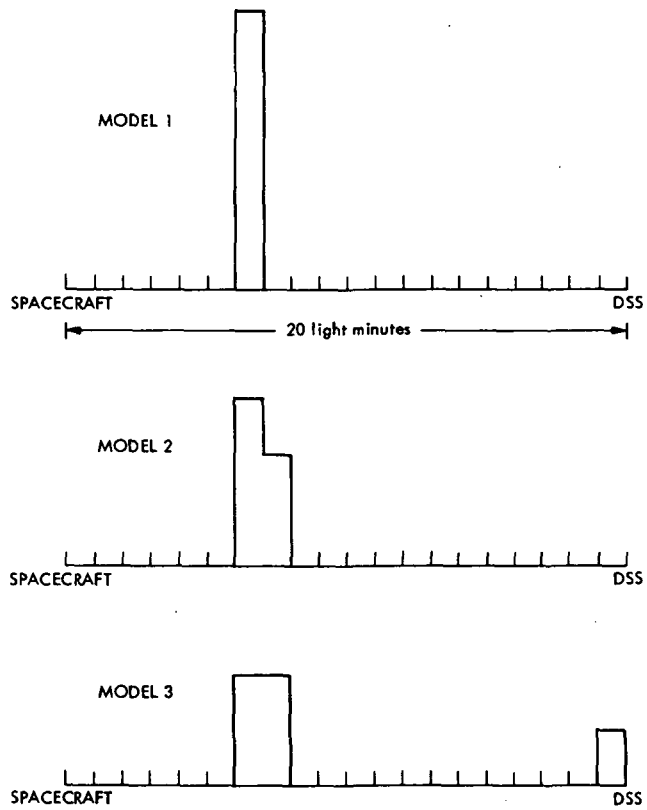


Fig. 3 Line-of-sight charged-particle relative distribution models

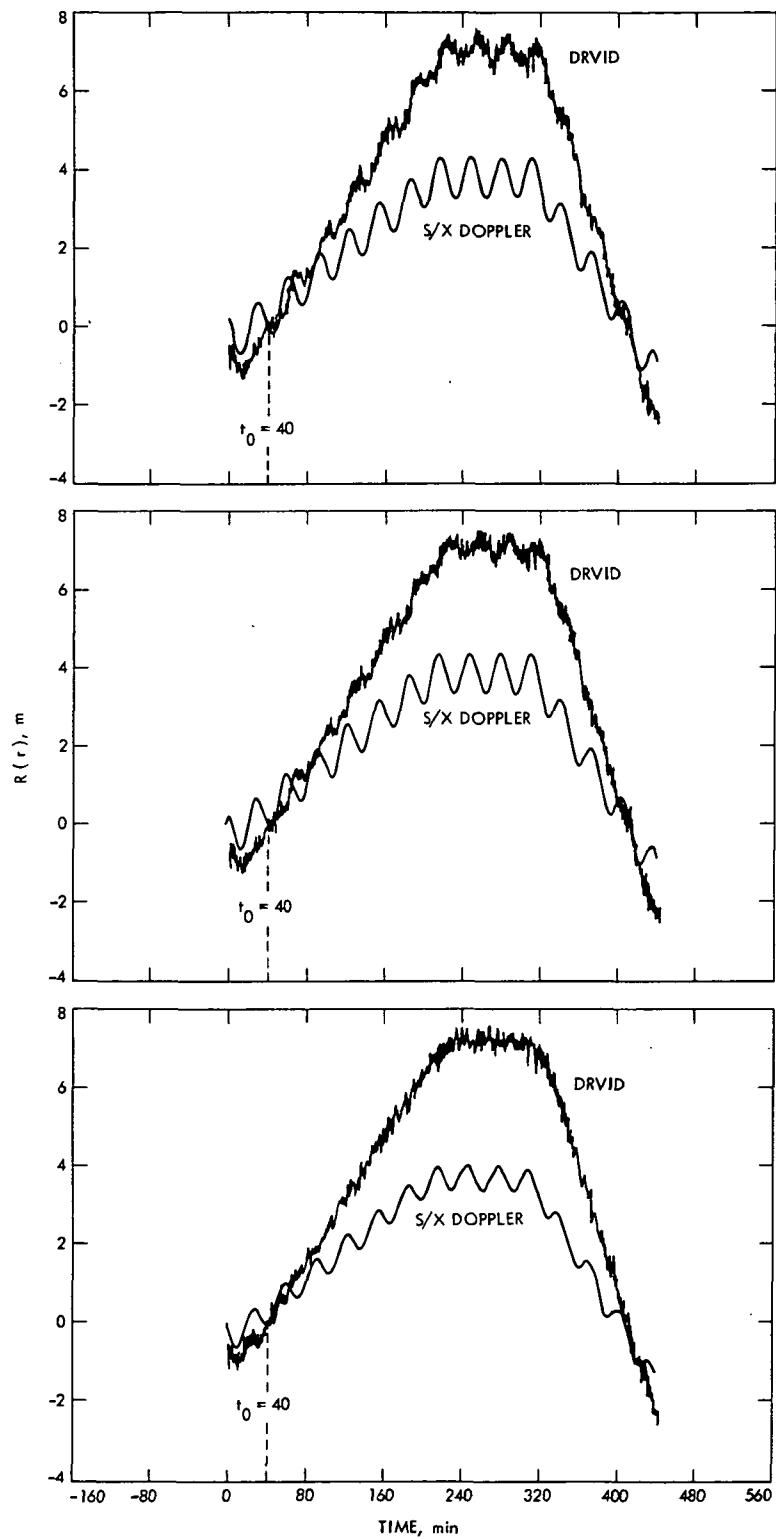


Fig. 4. Simulated DRVID and S/X band doppler (a) Model 1, (b) Model 2, (c) Model 3

ORIGINAL PAGE IS
OF POOR QUALITY.

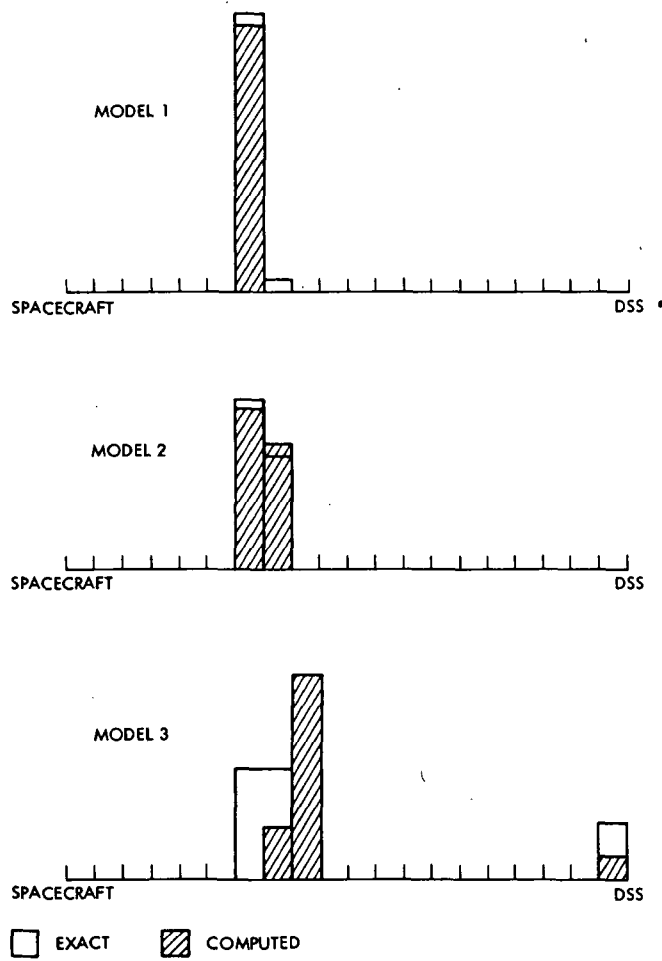


Fig. 5. Solutions of line-of-sight charged-particle relative distribution ($M = 60$)

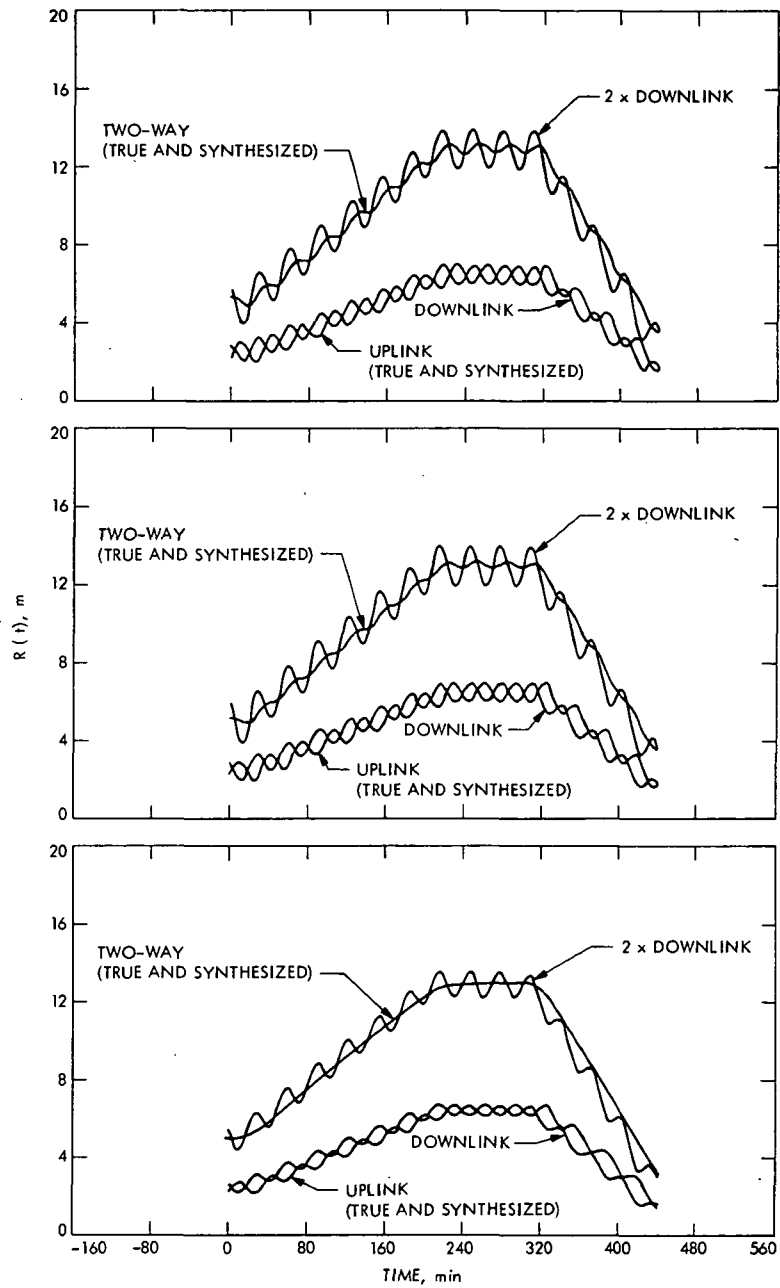


Fig. 6. Range changes due to charged particles (a) Model 1, (b) Model 2, (c) Model 3

ORIGINAL PAGE IS
OF POOR QUALITY

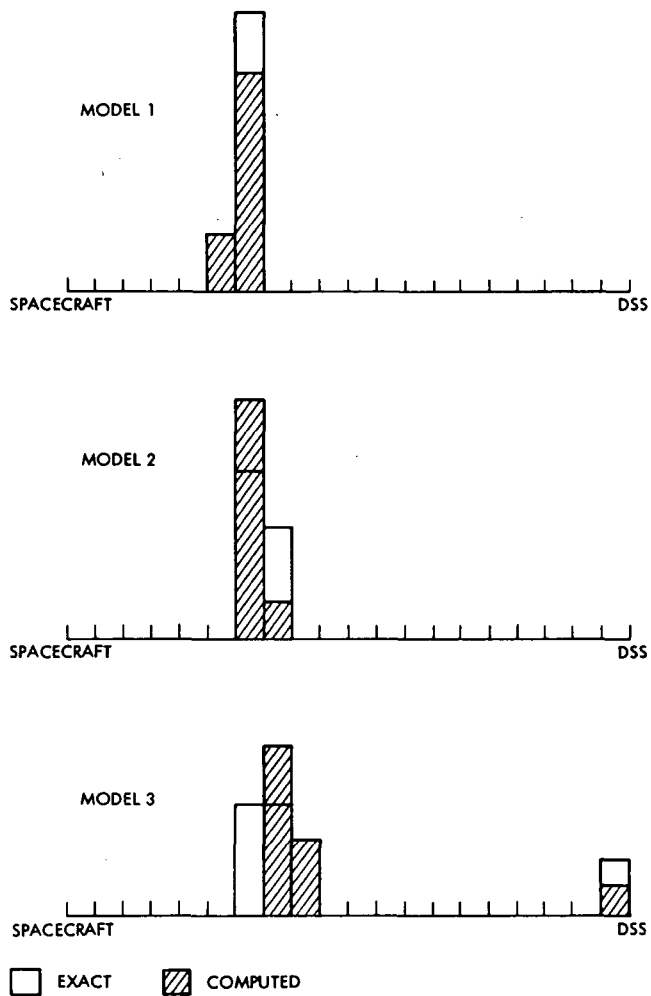


Fig. 7. Solutions of line-of-sight charged-particle relative distribution ($M = 40$)

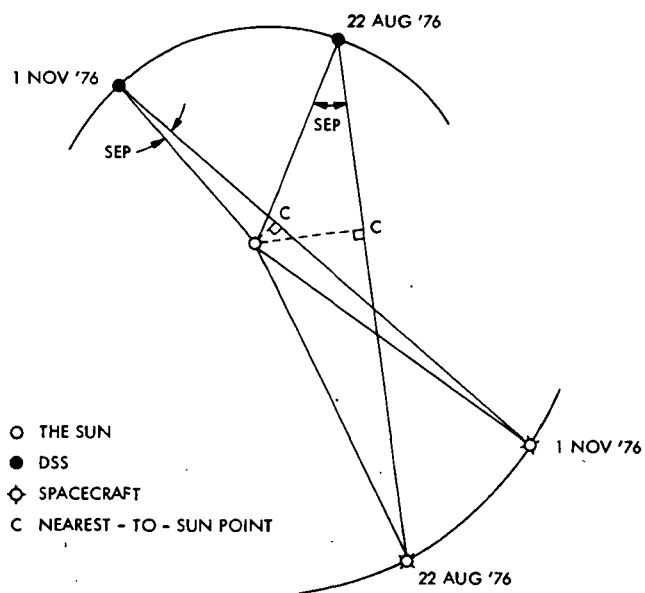


Fig. 8. Helio-centric geometry of deep space station and spacecraft (Viking B)

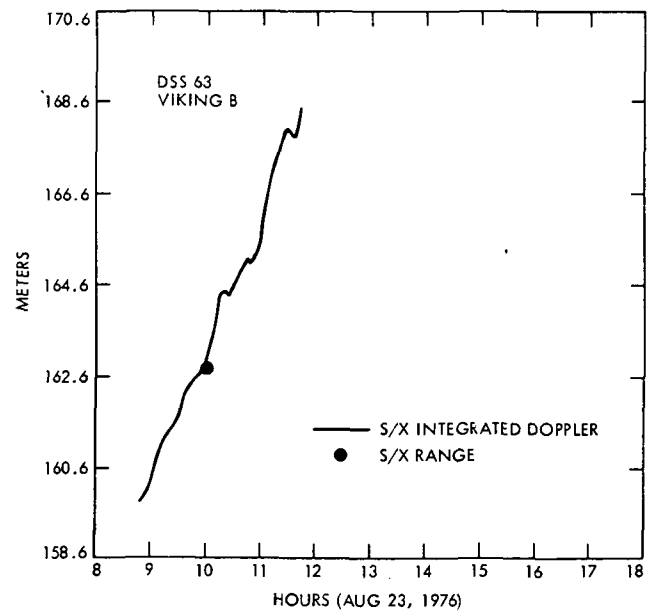
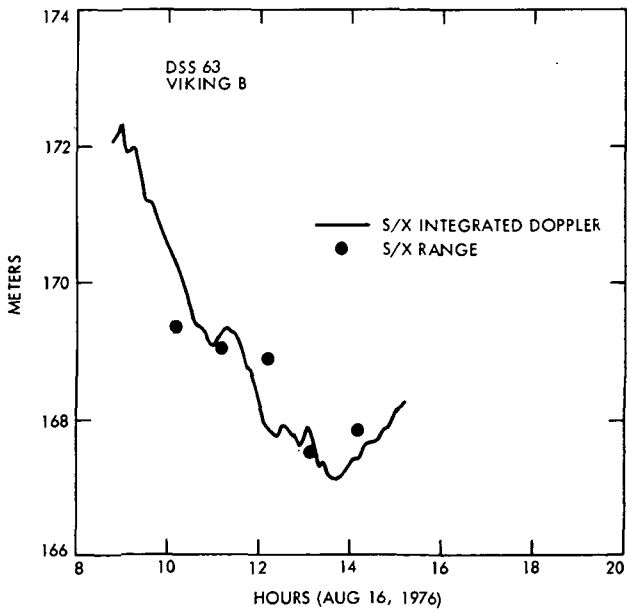
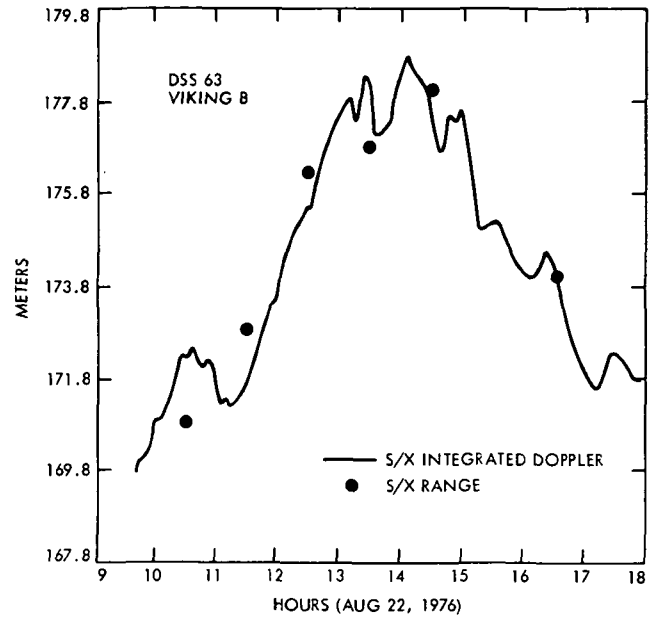
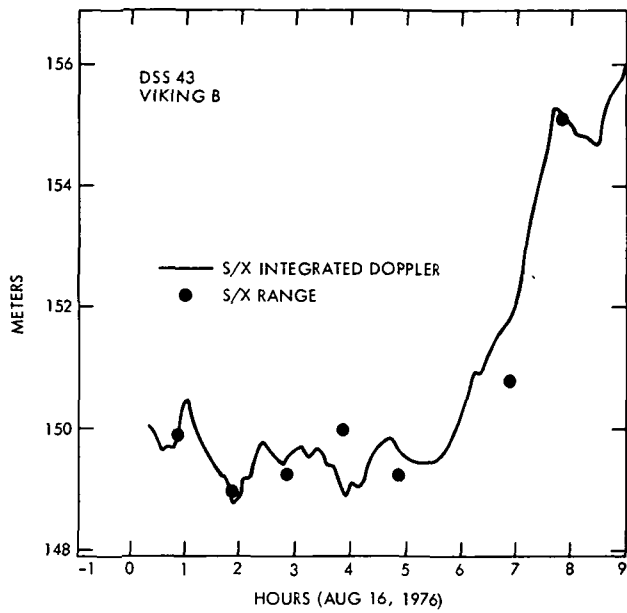


Fig. 9. S/X integrated doppler and S/X range

ORIGINAL PAGE IS
OF POOR QUALITY.

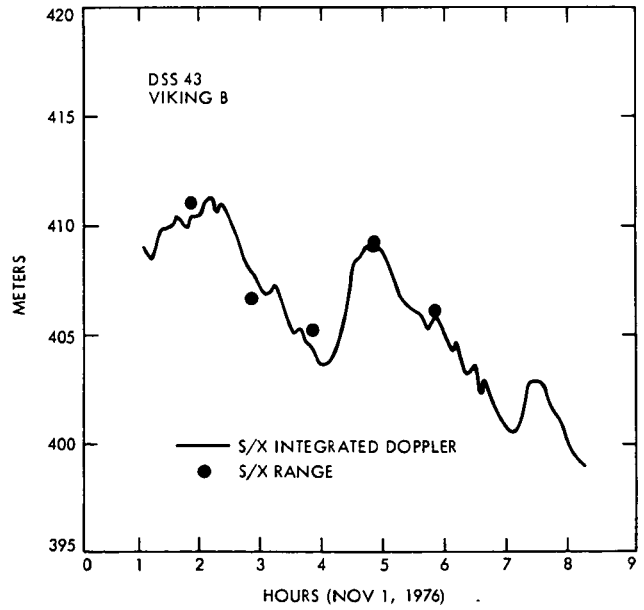
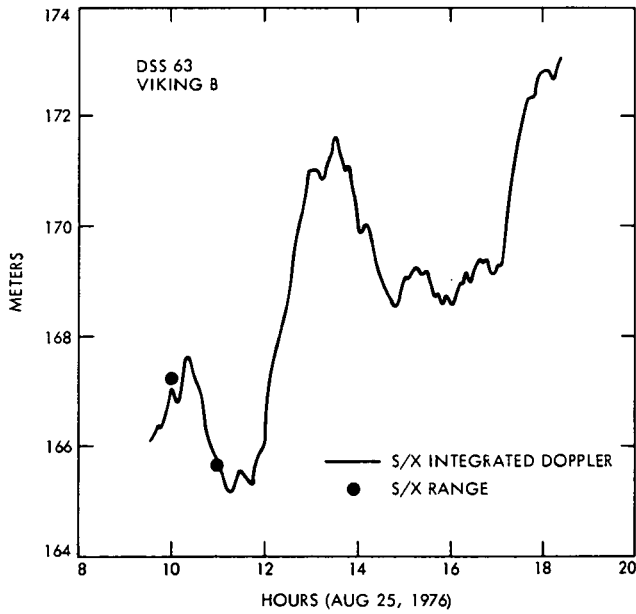
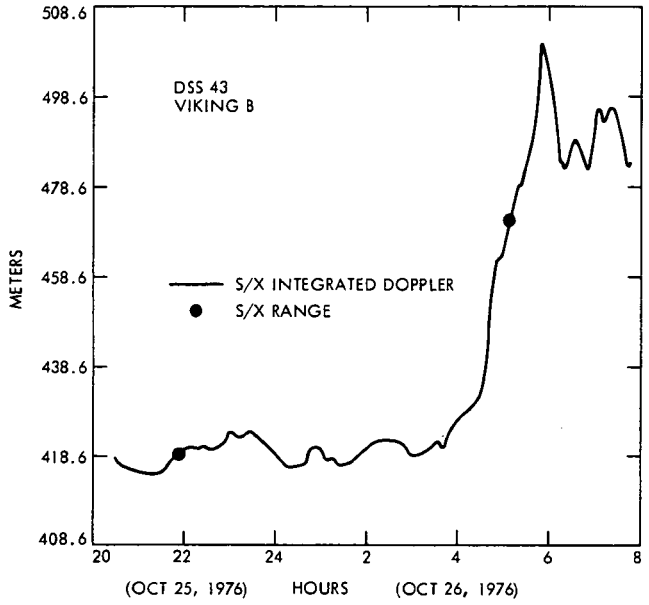
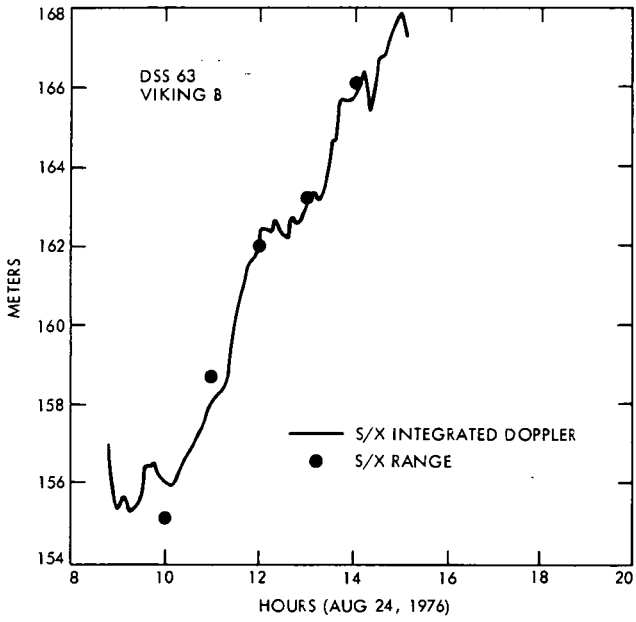


Fig. 9 (contd)

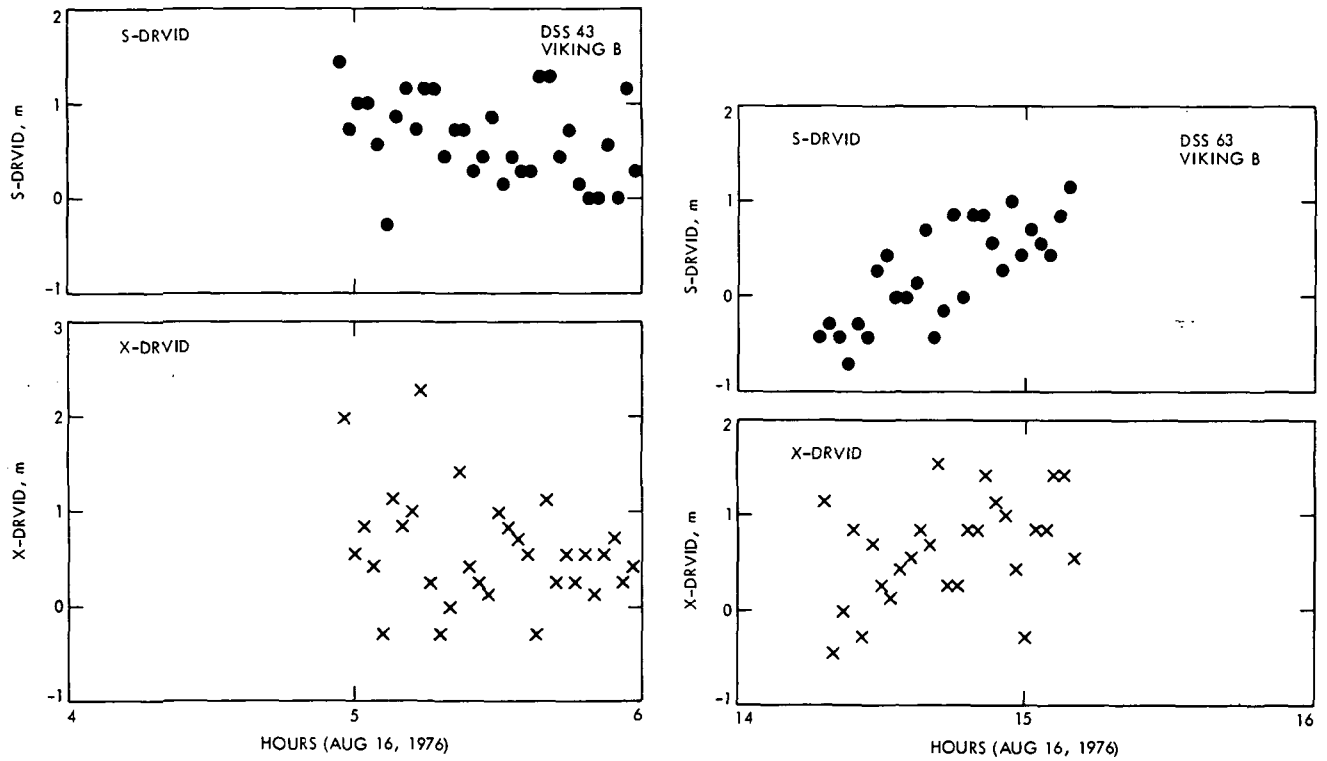


Fig. 10. S- and X-band DRVID

ORIGINAL PAGE IS
OF POOR QUALITY

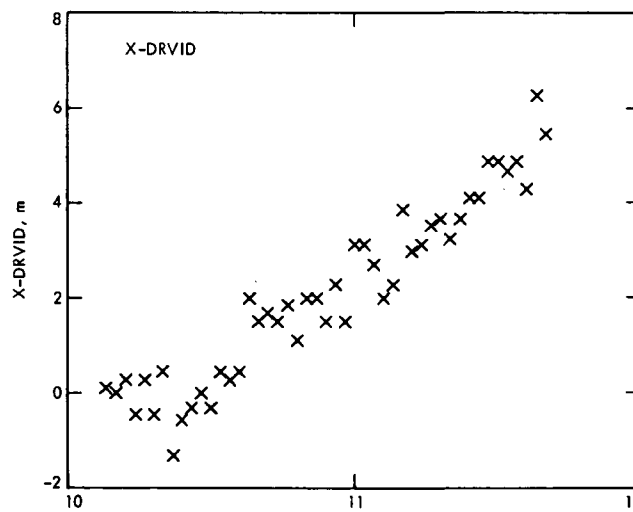
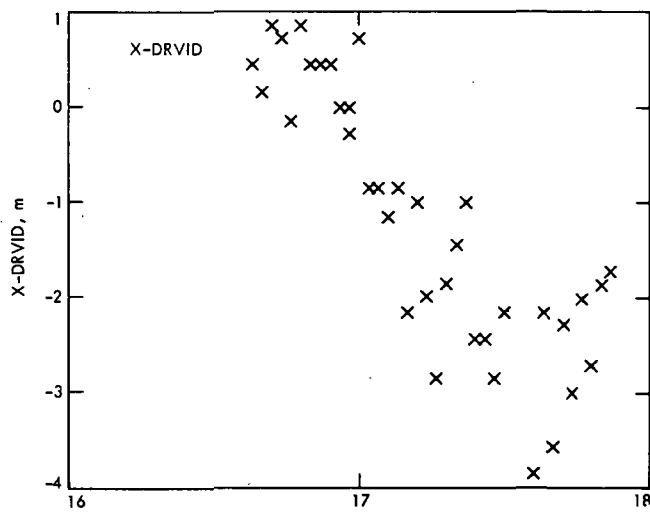
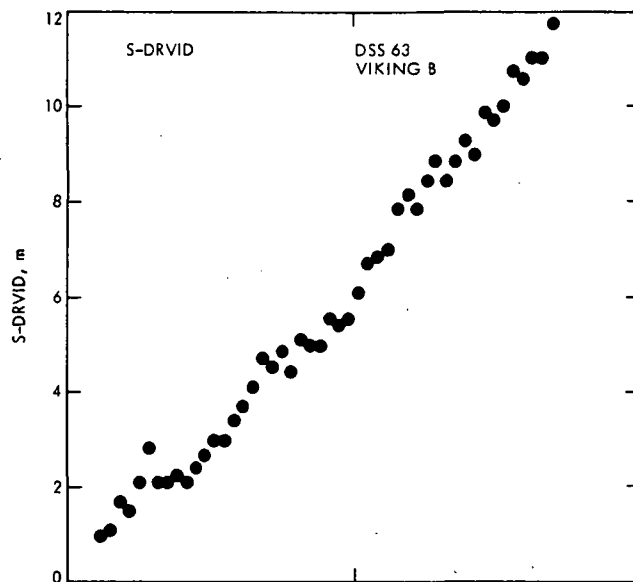
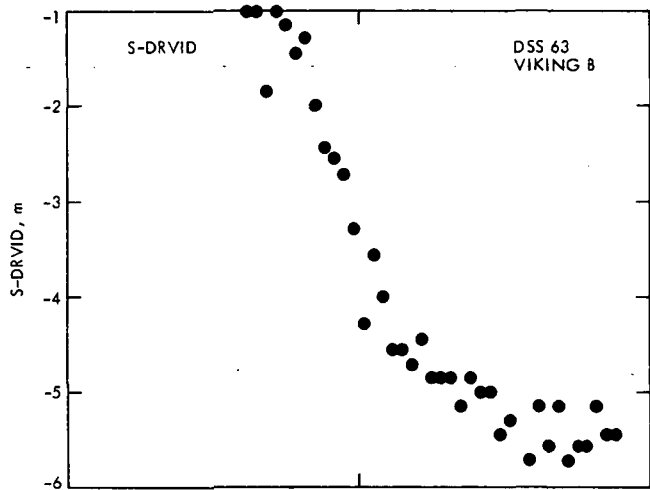


Fig. 10 (contd)

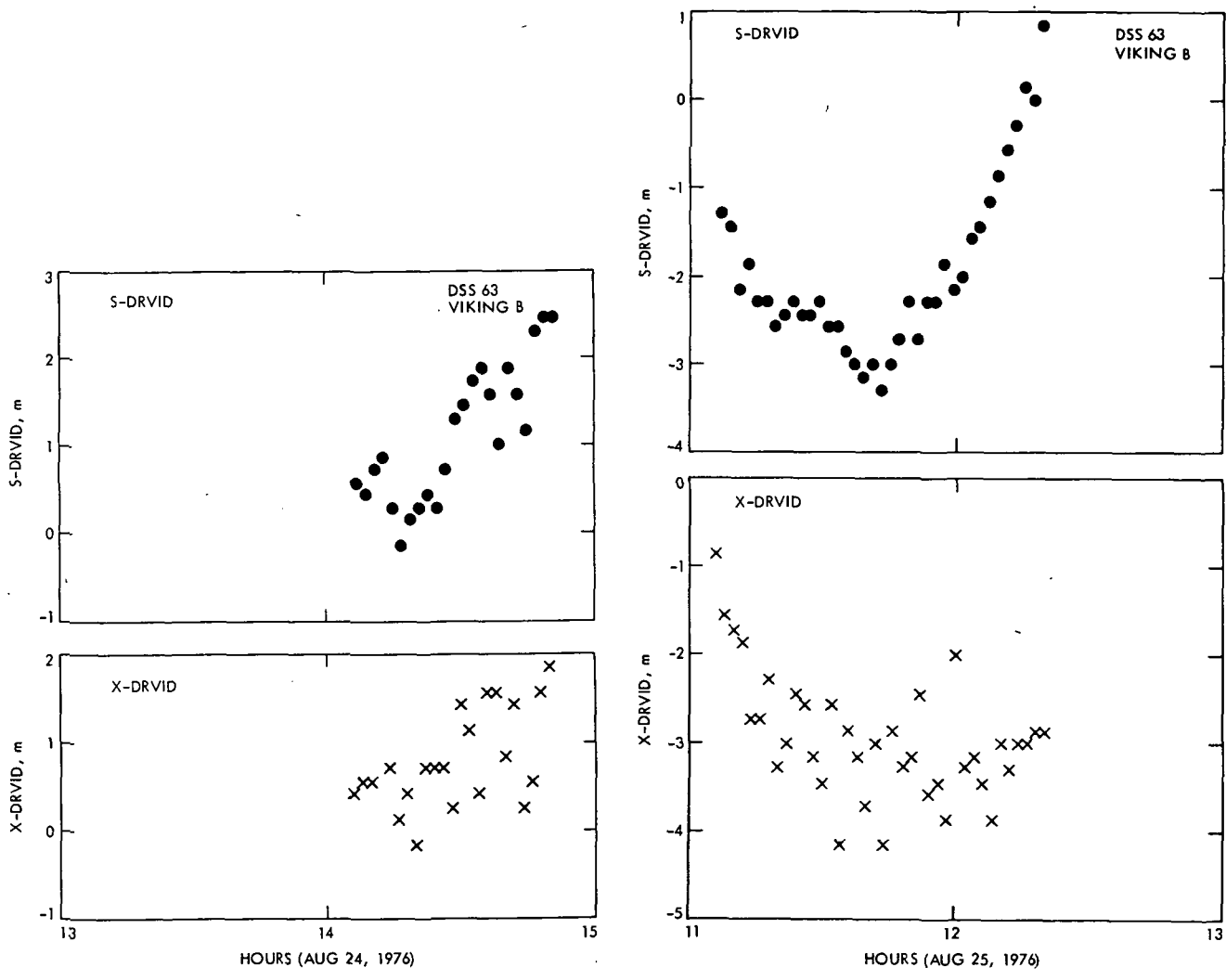


Fig. 10 (contd)

ORIGINAL PAGE IS
OF POOR QUALITY

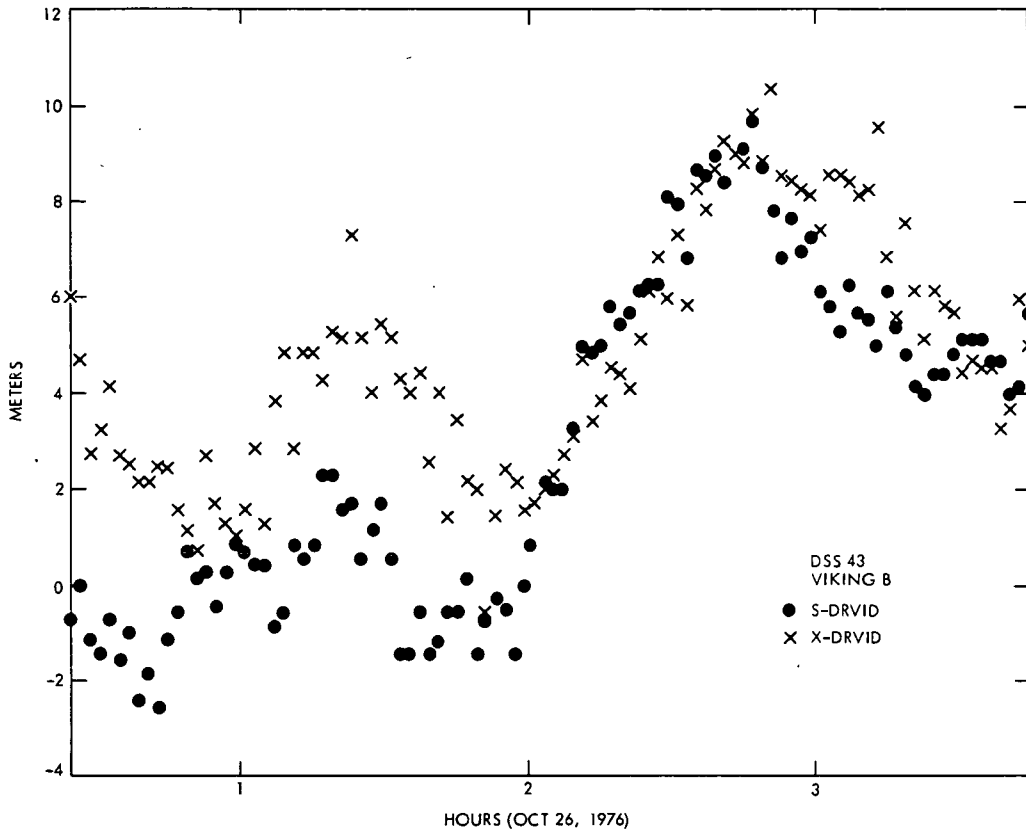


Fig. 10 (contd)

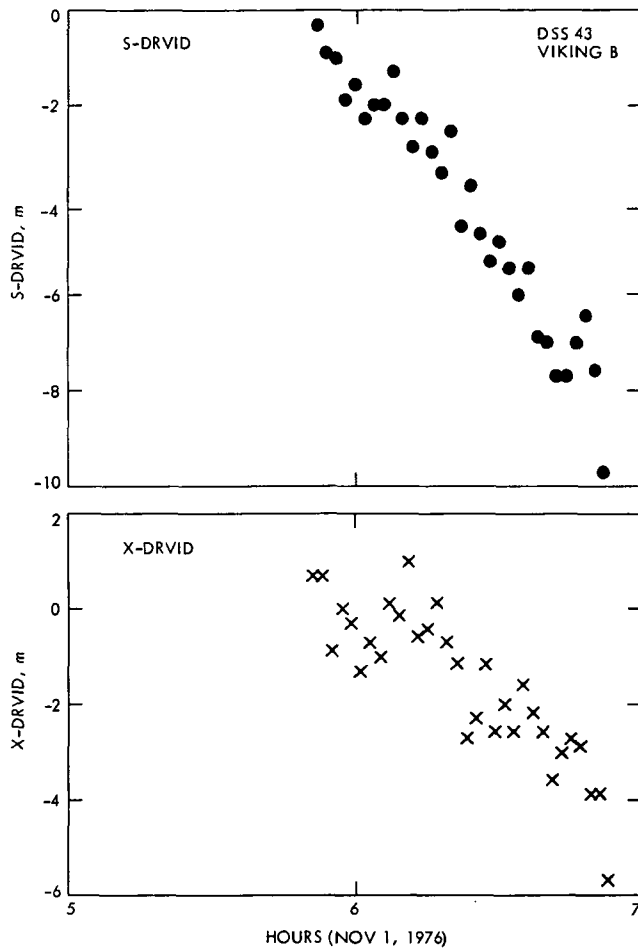


Fig. 10 (contd)

ORIGINAL PAGE IS
OF POOR QUALITY

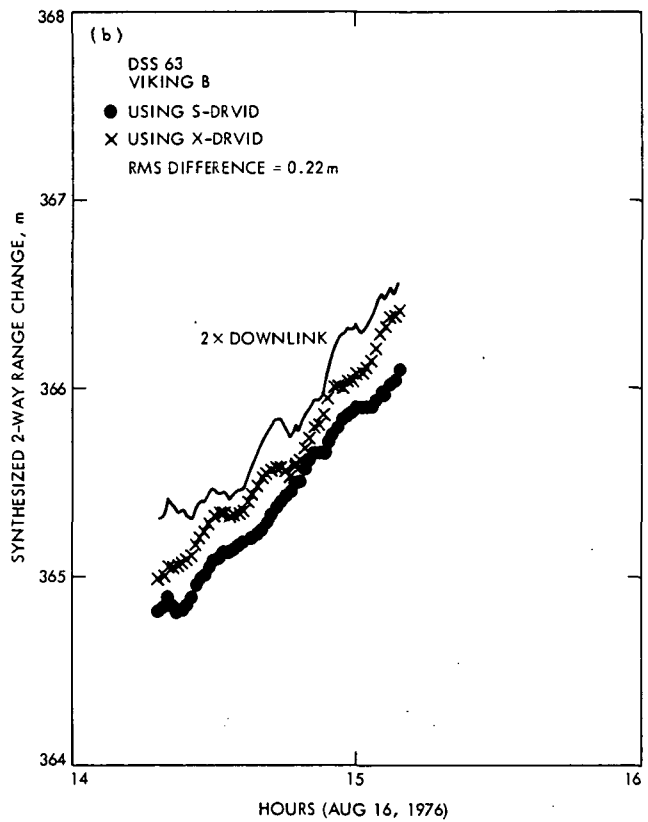
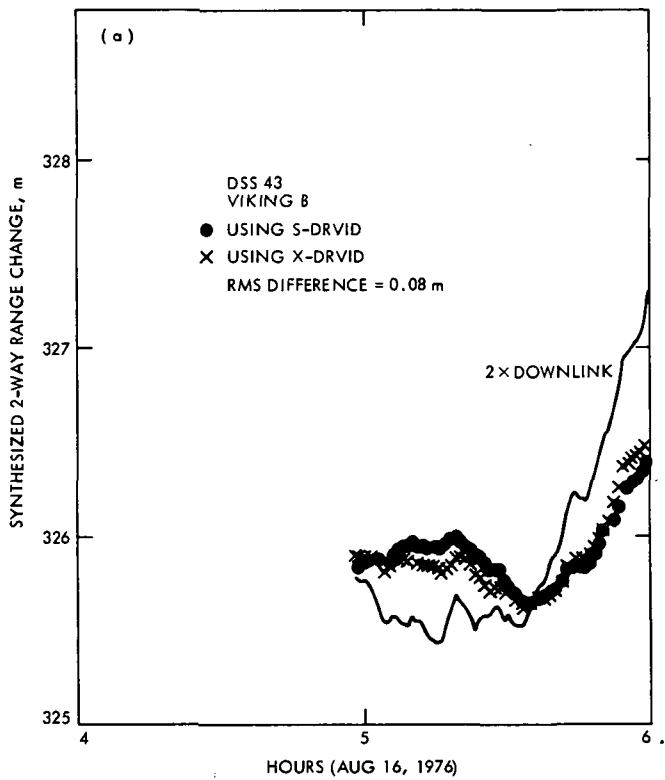


Fig. 11. Comparison of synthesized 2-way range changes and 2 x downlink range change

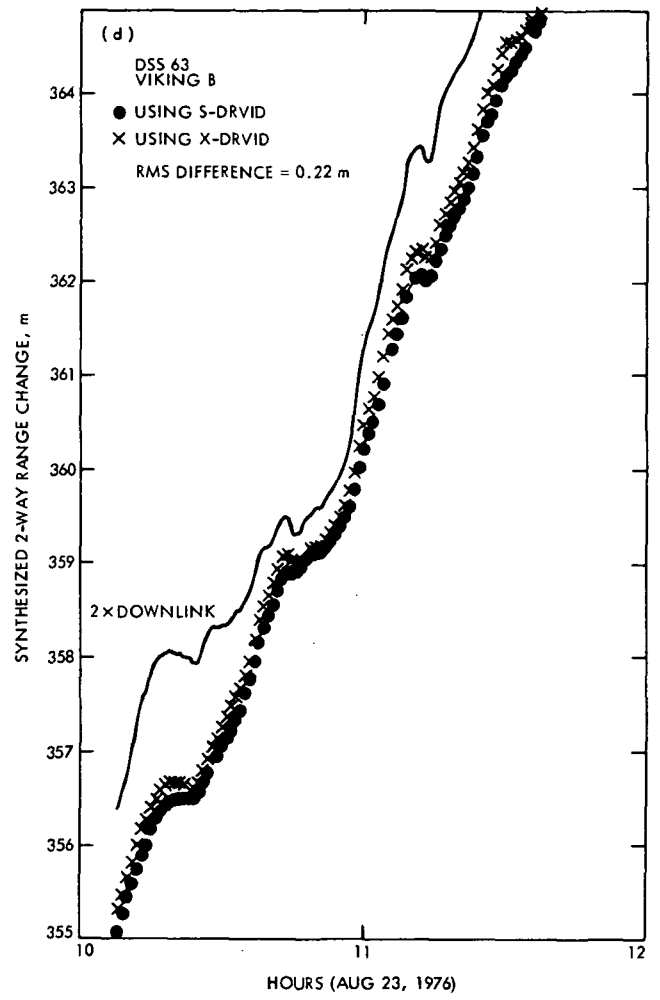
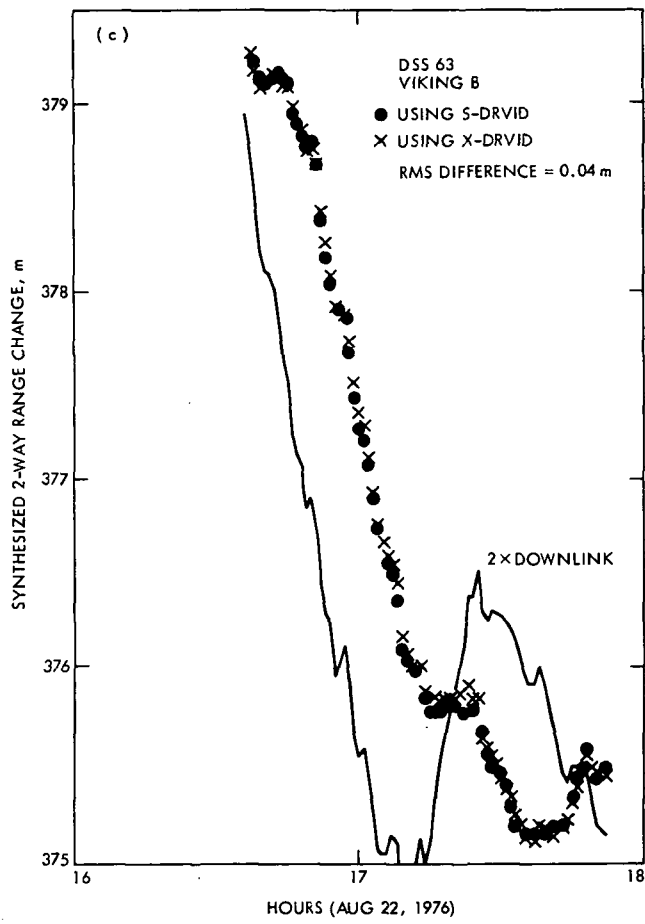


Fig. 11 (contd)

ORIGINAL PAGE IS
OF POOR QUALITY

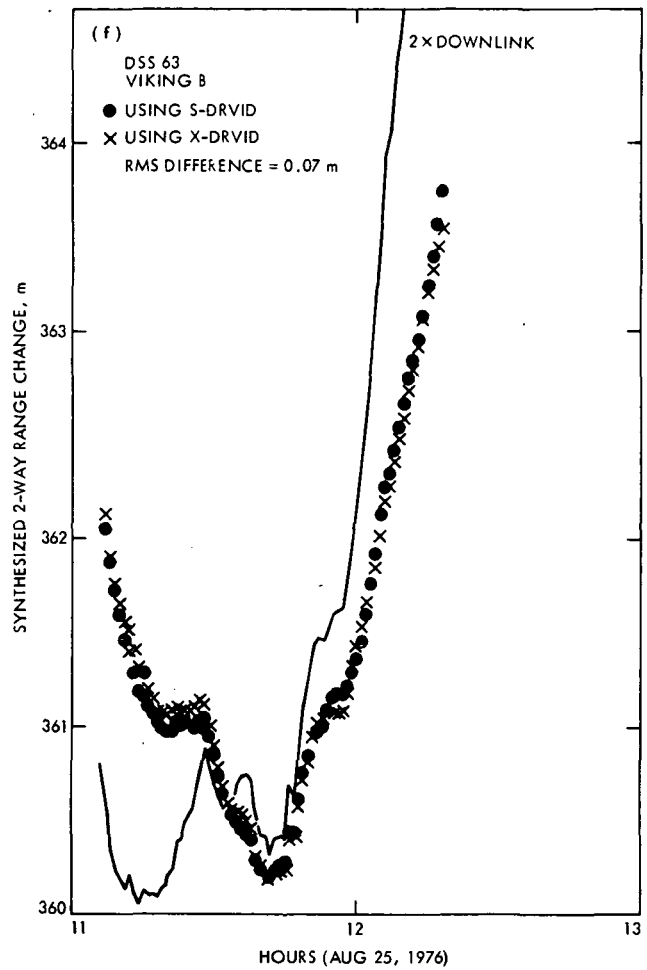
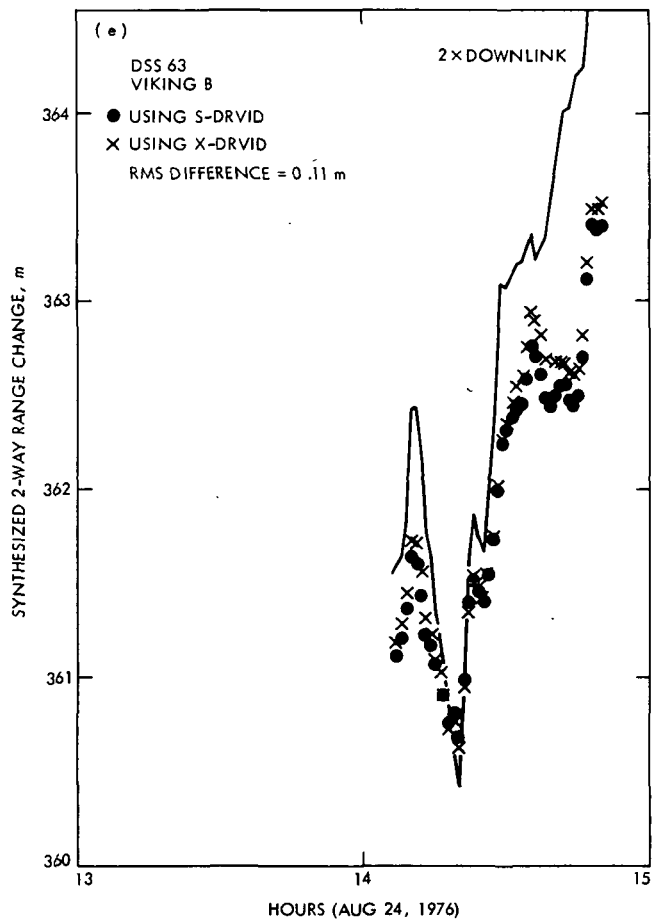


Fig. 11 (contd)

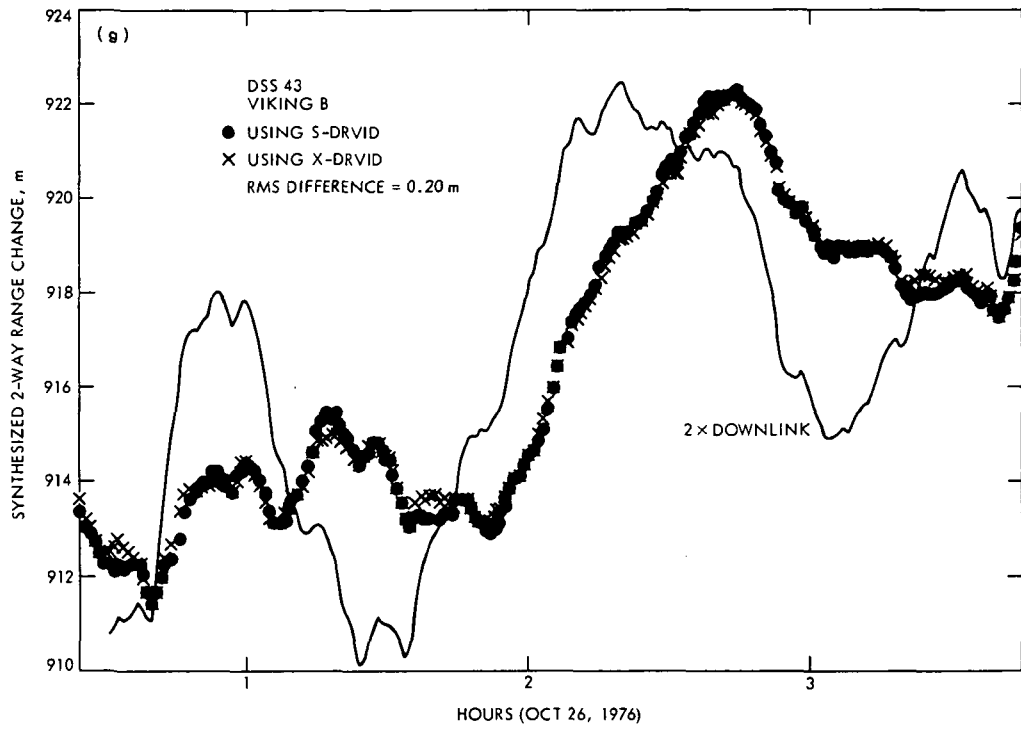


Fig. 11 (contd)

ORIGINAL PAGE IS
OF POOR QUALITY

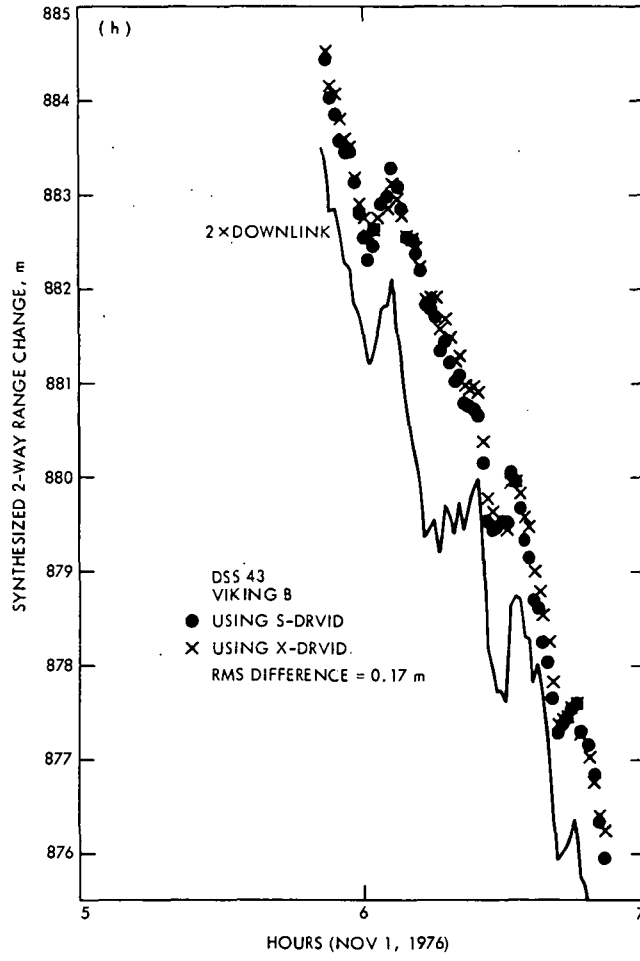


Fig. 11 (contd)

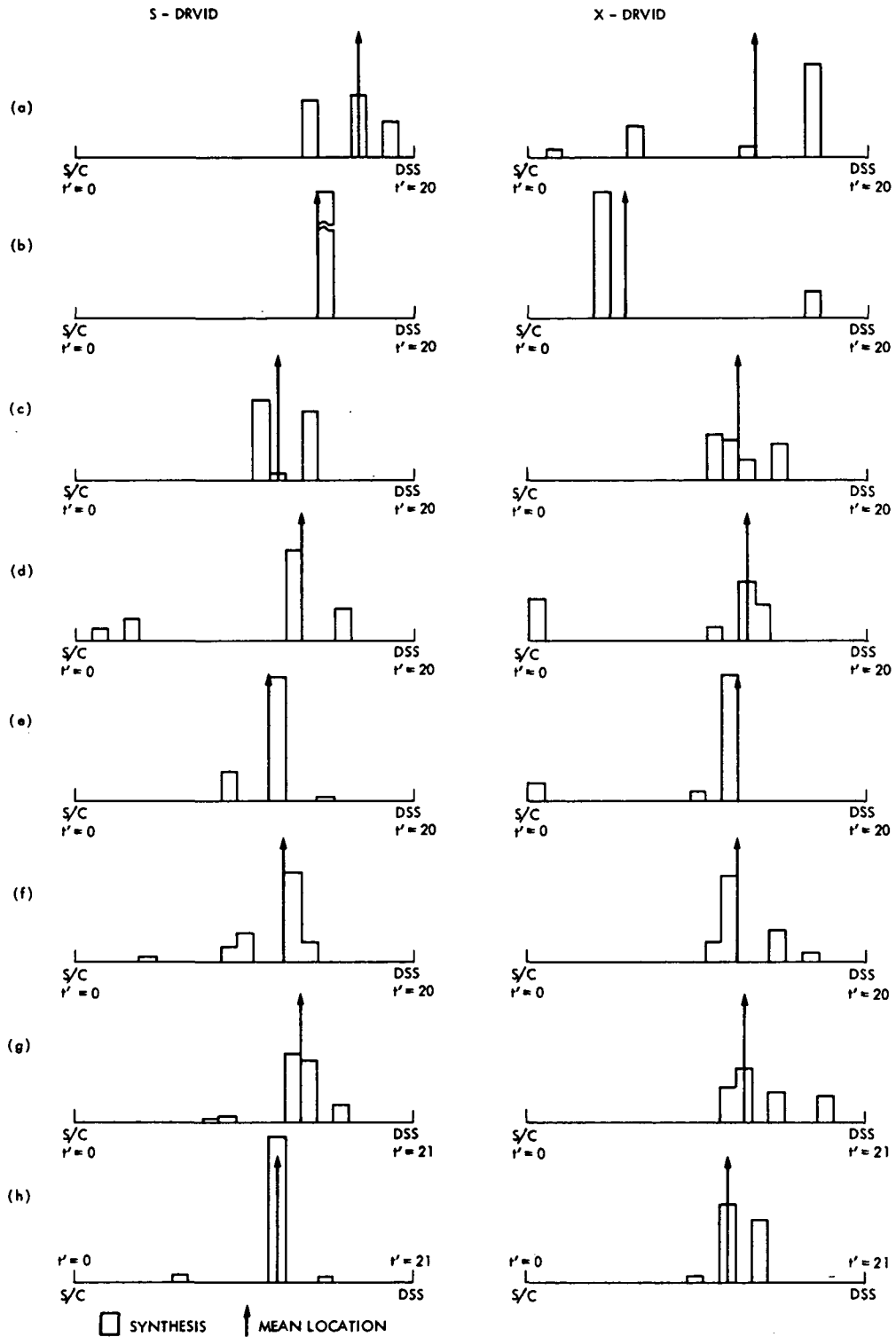


Fig. 12. Line-of-sight charged-particle relative distribution by synthesis and mean location by method of Ref. 3 (shown as arrows)

ORIGINAL PAGE IS
OF POOR QUALITY

N78-11142

Doppler Profiles of Radio Frequency Interference

B. K. Levitt

Communications Systems Research Section

The DSN is currently developing a wideband, digital system for monitoring the RF environment at each receive complex. One of the characteristics that will be used to identify observed RFI (radio frequency interference) sources is the time variations of their received spectra. As a preliminary aid to this effort, doppler profiles were computed for Earth satellites and planes with typical flight parameters. The analysis shows that doppler rates will be high enough to cause spectral smearing and degrade the surveillance station detection capability under certain conditions.

I. Introduction

The DSN is in the process of developing a wideband, digital surveillance system for on-site detection and identification of radio frequency interference (RFI) sources (Reference 1). When a signal of interest is detected, its identity will be determined on the basis of any or all of the following observations: time and duration of occurrence, source direction, polarization, spectral characteristics and their time variations, and a comparison of these data with previously identified RF signatures in storage. Focussing on the spectral variation aspect, doppler profiles were computed for Earth satellites and planes over a range of expected flight parameters, as a preliminary aid to the identification effort.

$$v = \frac{631.3}{\sqrt{r}} \quad (1)$$

Assume further that the satellite passes directly over the RF surveillance station, and that it is transmitting at S-band (center frequency $f_0 \cong 2295$ MHz). Also, neglect the Earth's rotation (a second-order effect) and use $r_0 = 6357$ km for the Earth's radius.

Referring to Fig. 1, the satellite-to-surveillance station range s is given by

$$\begin{aligned} s^2 &= r^2 \sin^2 \theta + (r \cos \theta - r_0)^2 \\ &= r^2 (1 + \gamma^2 - 2\gamma \cos \theta); \quad \gamma \equiv r_0/r \end{aligned} \quad (2)$$

II. Earth Satellites

To facilitate the doppler calculations, assume the satellite is in a circular orbit, with radius r (km), tangential velocity v (km/s), and constant angular velocity $d\theta/dt = v/r$ (rad/s). Equating gravitational and centrifugal forces on the satellite yields the relation (Reference 2)

$$2s \frac{ds}{dt} = 2r^2 \gamma \sin \theta \frac{d\theta}{dt} \quad (3)$$

The satellite signals received at the surveillance station exhibit a doppler shift

$$f = -\frac{f_0}{c} \left(\frac{ds}{dt} \right) = -\frac{f_0 v \gamma \sin \theta}{c \sqrt{1 + \gamma^2 - 2\gamma \cos \theta}}; |\theta| \leq \theta_0 \quad (4)$$

where $\theta_0 = \cos^{-1} \gamma$, and $c = 3 \times 10^5$ km/s is the velocity of light. The maximum doppler shift occurs at $\theta = \pm\theta_0$:

$$|f|_{\max} = \frac{f_0 v \gamma}{c} \propto r^{-3/2} \quad (5)$$

The doppler rate for the received satellite signals is given by

$$\frac{df}{dt} = -\frac{f_0 v^2 \gamma}{c r \sqrt{1 + \gamma^2 - 2\gamma \cos \theta}} \left(\cos \theta - \frac{\gamma \sin^2 \theta}{1 + \gamma^2 - 2\gamma \cos \theta} \right) \quad (6)$$

The maximum doppler rate occurs at $\theta = 0$:

$$\left| \frac{df}{dt} \right|_{\max} = \frac{f_0 v^2 \gamma}{c r (1 - \gamma)} \quad (7)$$

Note that $df/dt = 0$ at $\theta = \pm\theta_0$, so that the doppler shift reaches its maximum amplitude each time the satellite crosses the surveillance station horizon. The total observation time from horizon to horizon is

$$T = \frac{2\theta_0}{d\theta/dt} = \frac{2r}{v} \cos^{-1} \gamma \quad (8)$$

Earth satellite altitudes typically range between 300 and 1200 km (Ref. 2); doppler parameters are tabulated below for altitudes in this range.

$r - r_0$ km	v km/s	$ f _{\max}$ KHz	$\left \frac{df}{dt} \right _{\max}$ KHz/s	T min
300	7.74	56.5	1.46	8.63
750	7.49	51.2	.512	14.7
1200	7.26	46.7	.283	19.8

Normalized doppler profiles for satellite signals are plotted in Fig. 2.

III. Planes

As shown in Fig. 3, the simplification here is that the plane is travelling in a straight line at a constant velocity v as it passes within range of the surveillance station. The plane-to-surveillance station range s at closest approach is s_0 ; it need not be assumed that the plane is directly over the station at $s = s_0$. Letting time $t = 0$ at closest approach, the instantaneous range satisfies the relation

$$s^2 = s_0^2 + v^2 t^2 \quad (9)$$



$$2s \frac{ds}{dt} = 2v^2 t \quad (10)$$

The received doppler shift can be written in the form

$$f = -\frac{f_0 v \mu(t)}{c \sqrt{1 + \mu^2(t)}}; \mu(t) \equiv \frac{vt}{s_0} \quad (11)$$



$$|f|_{\max} = \frac{f_0 v}{c} \quad (12)$$

The doppler profile represented by Equations 11 and 12 is plotted in normalized form in Figure 4.

The doppler rate is given by

$$\frac{df}{dt} = -\frac{f_0 v^2}{c s_0} [1 + \mu^2(t)]^{-3/2} \quad (13)$$

$$\left| \frac{df}{dt} \right|_{\max} = \frac{f_0 v^2}{c s_0} \quad (14)$$

We might reasonably assume that plane velocities lie in the range between .05 and 1.00 km/s (110 to 2200 mi/hr). Furthermore, since faster planes typically fly at higher altitudes, assume the following ranges of s_0 for a given v :

v		min s_0		max s_0	
km/s	mi/hr	km	ft	km	ft
.05	110	1	3300	3	9800
.25	560	3	9800	12	39,000
1.00	2200	10	33,000	25	82,000

ORIGINAL PAGE IS
OF POOR QUALITY

Based on these flight parameter assumptions, for S-band signals ($f_0 \cong 2295$ MHz), the surveillance station would observe the following maximum doppler shifts and doppler rates:

ν km/s	$ f _{\max}$ kHz	s_0 km	$\left \frac{df}{dt}\right _{\max}$ KHz/s
.05	.38	1	.0190
		3	.0064
.25	1.9	3	.160
		12	.040
1.00	7.7	10	.77
		25	.31

IV. Conclusions

In the evolution of the digital RFI surveillance system design since Reference 1 was written, the current plan is to demonstrate a first-generation device at Goldstone in mid-1978, with the following features:

- 15 dB gain, 30° beamwidth, circular horn antenna with low, fixed elevation direction, rotating 360° in azimuth every minute

- generation of a 20 MHz spectrum, containing 2^{16} or about 65,000 lines (305 Hz resolution), every 5 seconds

Thus, an RF source in a given azimuthal direction will be observed for 5 seconds of each minute it is within range of the surveillance station.

Based on the assumptions in the previous two sections, upper and lower bounds for satellite and plane doppler profiles are compared in Figure 5. It is evident that the observed doppler signatures will be very useful for identifying mobile RF sources and their trajectories. Also, because the doppler rate is often very large near closest approach, there will be some spectral smearing in this region: obviously, this effect is more critical for line spectra than for wideband signals. This smearing effect can impact the RF detection capability: for example, a CW signal that barely exceeds the surveillance system threshold under zero doppler rate conditions will not be detected if the doppler rate is too high. As a benchmark, the 61 Hz/s doppler rate point is shown in the doppler profiles of Figure 5 (the doppler rate is always less than this level for the $\nu = .05$ km/s, $s_0 = 3$ km plane case): at this doppler rate, a CW spectral line will move through one resolution bin over the 5 second observation time.

References

1. Levitt, B. K., "Analysis of a Discrete Spectrum Analyzer for the Detection of Radio Frequency Interference," in *The Deep Space Network Progress Report 42-38*, Jet Propulsion Laboratory, Pasadena, California, pp. 83-98, April 15, 1977.
2. Levitt, B. K., "Radio Frequency Interference from Near-Earth Satellites," in *The Deep Space Network Progress Report 42-37*, Jet Propulsion Laboratory, Pasadena, California, pp. 69-77, February 15, 1977.

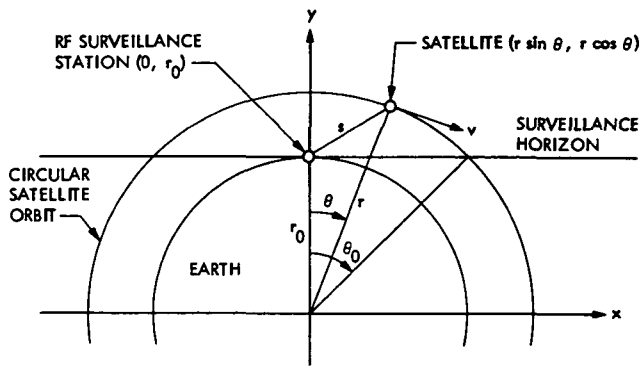


Fig. 1. Earth satellite doppler geometry

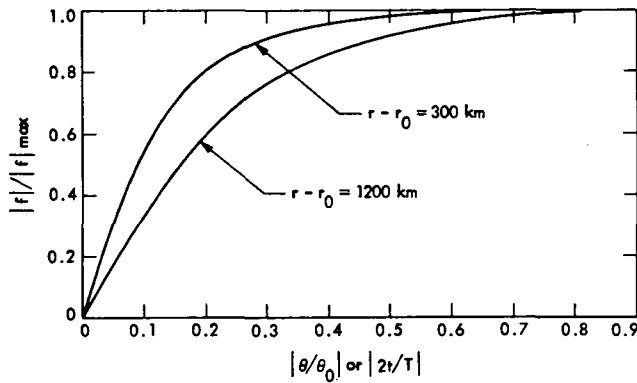


Fig. 2. Normalized doppler profiles for satellite signals received at a surveillance station. Zero-time reference ($t = 0$) is at closest approach ($\theta = 0$)

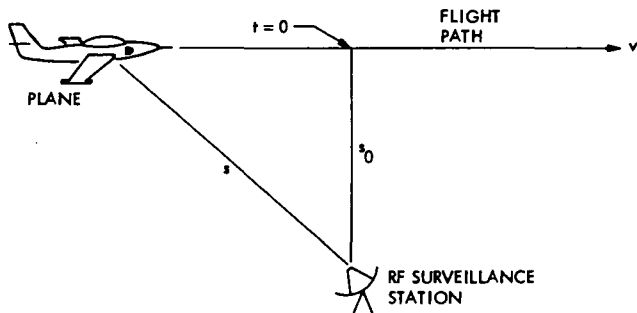


Fig. 3. Plane doppler geometry

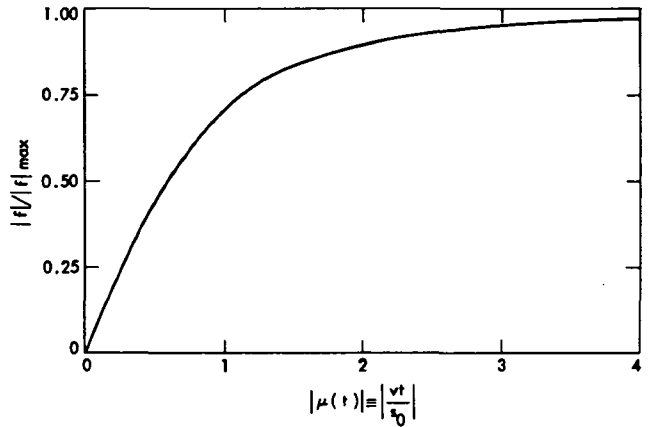


Fig. 4. Normalized doppler profile for plane. Time $t = 0$ at closest approach (range $s = s_0$)

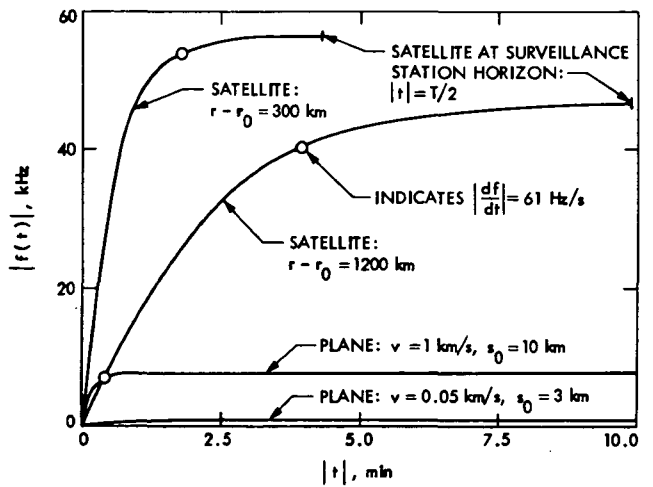


Fig. 5. Comparison of satellite and plane doppler profiles. (See text for explanation of symbols in this graph.)

ORIGINAL PAGE IS
OF POOR QUALITY.

N78-11143

Sampled-Data System Analysis of Antenna Conical-Scan Tracking

R. A. Winkelstein

Communications Systems Research Section

The conical-scan tracking system described by Ohlson and Reid in JPL Technical Report 32-1605 is analyzed as a sampled-data feedback control system. Tracking mode equations for both spacecraft signals and radio source signals are developed. In the case of spacecraft tracking, a rationale is presented for selection of parameters which minimizes the sum of the required scan radius and three times the RMS error jitter. With this criterion, reasonable system performance can be obtained with signals down to receiver threshold. For radio source tracking, the RMS error jitter is negligible, and a set of system parameters is recommended which allows conservative operation of the system.

I. Introduction

Conical-scan tracking is a method for automatically tracking a signal source with a Deep Space Network (DSN) antenna by continuously rotating the antenna a small fraction of the antenna beam width around the tracking boresight, a hypothetical line between the antenna and signal source. An error in the tracking angle results in a systematic variation in received signal strength which is used to derive correction offsets in the two antenna pointing axes. The DSN has used conical-scan tracking for several years, and an excellent report (Ref. 1.) has been written by J. E. Ohlson and M. S. Reid describing the system both theoretically and practically. Using Ohlson's work as a guide, the following report recasts the description of the conical-scan system into that of a sampled-data feedback control system with parameters perhaps more familiar to the control system analyst. In addition, a rationale is presented for optimum selection of system parameters as a function of desired system performance. Theoretical results derived from the sampled-data system model are in complete agreement with Ref. 1., and a correspondence of selected parameters is listed in Appendix B.

II. System Description

A block diagram of the conical-scan system is shown in Fig. 1. Two feedback channels are used, one for an assumed elevation axis correction and the other for an assumed cross-elevation axis correction. Both channels are identical except for the quadrature references, $\sin \omega_m t$ and $\cos \omega_m t$, and any necessary trigonometric transformations relating the assumed axes with the actual antenna axes driven by the antenna pointing subsystem. The radian rotation frequency, ω_m , is given by

$$\omega_m = \frac{2\pi}{P(1-A)} \quad (1)$$

where $P(1-A)$ is the rotation period of the antenna, typically 28 seconds. After each rotation, the antenna is stopped rotating for "AP" seconds, usually two, during which time new offset corrections are calculated and the antenna slightly repositioned as a result. Thus P represents the total time

between successive corrections, and is the basic timing quantity for the sampled-data model.

The receiver generates a signal strength voltage sampled at ten times a second by the correlator. Two types of signals can be tracked with the conical-scan system. One type is a space-craft signal, in which case the signal strength voltage is derived from the receiver automatic gain control (AGC) circuit. The second signal type comes from a wide band stellar radio source. In this case, the signal strength voltage comes from a wide band power detector.

The correlator block provides an estimate of the offset error by correlating the sampled signal strength voltage against the correlator input reference. An updated estimate is obtained for each rotation of the antenna. These estimates are filtered by the digital filter according to the relation

$$X_3(k) = X_3(k-1) + k_d X_2(k) \quad (2)$$

where $X_3(k)$ is the k th filter output, $X_3(k-1)$ is the previous filter output, $X_2(k)$ is the k th filter input, and k_d is a gain constant. The antenna pointing subsystem provides a constant offset proportional to X_3 for each channel throughout the succeeding rotation period. In addition to these fixed offsets and the conical-scan rotation signals, the antenna pointing subsystem provides position and velocity correction signals from predicted information called predicts which are generated external to the system before the track. Predicts are required by the conical-scan system because the system using Eq. (2) is a first order system that can only correct for small fixed offsets in the predict values. In this sense, the conical-scan system is used only to "fine tune" the pointing angle generated by the predicts.

III. System Analysis

In a properly operating conical-scan tracking system, the two feedback channels are identical but independent. Therefore an analysis will be made only of the elevation channel, with similar results applicable to the cross-elevation channel.

A. Laplace Block Diagram

A first step in the analysis is a formulation of the Laplace block diagram as shown in Fig. 2. $\phi(s)$ is the Laplace transform of the desired elevation angle, and $C(s)$ is the Laplace transform of the elevation angle generated by the conical-scan system. "s" is the Laplace transform complex frequency. The channel error, $E(s)$, given by

$$E(s) = \phi(s) - C(s) \quad (3)$$

is input to the correlator which acts as an integrate-and-dump circuit.

The correlator circuit can be represented as a pure integrator whose output is the difference between the present value and a past value occurring at the start of the integration time. This past value is obtained through a pure delay, $e^{-sP(1-A)}$, where e is the natural logarithm base, and $P(1-A)$ is as defined in Eq. (1). " k_i ", the integrator gain constant, is a function of the signal type received, and includes several receiver parameters. Detailed derivations of k_i for both space-craft and radio source signals will be made in later sections. The Z-transform, (Ref. 2), of the correlator output, $X_2(z)$, is the sum of the error estimate, $X_1(z)$, and a noise value, $N(z)$, generated by the receiver noise processes. z is equal to e^{sP} . $N(z)$ will be determined in later sections of this report, but in this section it will be considered as an input quantity.

The transmission function shown for the digital filter in Fig. 2 results from taking the Z-transform of Eq. (2). Since the offset correction calculated by the digital filter is maintained as a constant value, $C(s)$, throughout the antenna rotation period, the resulting model for this operation is known as a zero-order hold circuit with a gain constant of k_h as shown in Fig. 2.

B. Z-transform Equations

An analysis of the conical-scan sampled-data channel of Fig. 2 will be carried out in the z complex plane. It is desired to determine $E(z)$, the Z-transform of $E(s)$, as a function of the input and noise signals. The procedure is to take the Z-transform of the state equations of Fig. 2 and then generate the signal flow graph of Fig. 3. $E(z)$ is then calculated by means of Mason's general gain formula (Ref. 3). The relevant state equations for Fig. 2 are

$$E(s) = \phi(s) - X_3(z) \frac{z-1}{z} \frac{k_h}{s} \quad (4)$$

$$X_1(s) = \left(k_i \frac{\phi(s)}{s} - X_3(z) k_h k_i \frac{z-1}{z} \frac{1}{s^2} \right) (1 - e^{-sP(1-A)}) \quad (5)$$

$$X_2(z) = X_1(z) + N(z) \quad (6)$$

$$X_3(z) = X_2(z) k_d \frac{z}{z-1} \quad (7)$$

Taking the Z-transform of Eqs. (4) and (5) respectively gives

$$E(z) = \phi(z) - X_3(z) k_h \quad (8)$$

$$X_1(z) = k_i \left[\mathcal{F} \left(\frac{\phi(s)}{s} \right) - \mathcal{F}_m \left(\frac{\phi(s)}{s} \right)_{m=A} \right] - X_3(z) k_h k_i P(1-A) \frac{1}{z} \quad (9)$$

where $\mathcal{F}(\phi(s)/s)$ is the Z-transform of $\phi(s)/s$, and $\mathcal{F}_m(\phi/s)_{m=A}$ is the modified Z-transform of $\phi(s)/s$, with "A" substituted for the parameter "m". This latter expression results from the standard technique of using the modified Z-transform to find the Z-transform of a function with pure delay.

Eqs. (6) through (9) form the signal flow graph of Fig. 3, from which $E(z)$ is found to be

$$E(z) = \phi(z) - \frac{\left[\mathcal{F} \left(\frac{\phi(s)}{s} \right) - \mathcal{F}_m \left(\frac{\phi(s)}{s} \right)_{m=A} \right] k_i k_d k_h Z}{z - 1 + k_T} - \frac{N(z) k_h k_d Z}{z - 1 + k_T} \quad (10)$$

$$k_T = k_i k_d k_h P(1-A) \quad (11)$$

where k_T is the overall loop gain constant. The response to a constant step input, ϕ_c , is next found with $N(z)$ set equal to zero. For a constant step input

$$\phi(s) = \frac{\phi_c}{s} \quad (12)$$

$$\phi(z) = \frac{\phi_c z}{z - 1} \quad (13)$$

$$\mathcal{F} \left(\frac{\phi(s)}{s} \right) = \frac{\phi_c P z}{(z - 1)^2} \quad (14)$$

$$\mathcal{F}_m \left(\frac{\phi(s)}{s} \right) = \phi_c \left(\frac{AP}{z - 1} + \frac{P}{(z - 1)^2} \right) \quad (15)$$

Using the values from Eqs. (12) through (15) gives

$$E(z) = \frac{\phi_c z}{z - 1 + k_T} \quad (16)$$

and taking the inverse Z-transform of Eq. (16) gives

$$E_k = \phi_c (1 - k_T)^k \quad (17)$$

where E_k is the loop error after the k th period P of the input step. Eq. (17) shows that the loop error is driven to zero when the input signal is a constant step offset.

A system response time constant, τ , may be defined by using the exponential equivalent to the step offset response

$$(1 - k_T)^k = e^{-(kP)/\tau} \quad (18)$$

Solving Eq. (18) for τ gives

$$\tau = \frac{-P}{\ln(1 - k_T)} \quad (19)$$

τ is considered to be of prime interest to the conical-scan system user, and therefore is presented as one of the tradeoff parameters when selecting an optimum set of system constants.

C. Root Locus Analysis

Root locus theory is useful in predicting the region of feedback system stability. The root locus is a plot in the complex plane of closed loop gain poles as a function of an open loop gain constant. A sampled-data system is stable if its closed loop poles are within the unit circle in the z complex plane. From Fig. 2, the open loop gain, G_{ol} , is

$$G_{ol}(s) = - \frac{k_i k_h k_d}{s^2} (1 - e^{-sP(1-A)}) \quad (20)$$

Taking the Z-transform of Eq. (20) gives

$$G_{ol}(z) = - \frac{k_T}{z - 1} \quad (21)$$

Thus k_T is selected as the open loop gain constant of interest. The closed loop gain, G_{cl} , is

$$G_{cl}(z) = \frac{1}{1 - G_{ol}(z)} = \frac{z - 1}{z - (1 - k_T)} \quad (22)$$

From Eqs. (21) and (22) it is seen that the conical-scan channel has a single open loop pole at z equal to one. This pole travels to the left along the real axis as a closed loop pole in direct proportion to k_T . At a value of k_T equal to two, which is the upper limit for system stability, the closed loop pole leaves the unit circle. This result is in agreement with Eq. (17) which shows an increasing error for k_T greater than two.

Also predicted by root locus theory is that a sampled-data feedback system with only a single pole at z equal to one will have a steady state error to a ramp or velocity input. This then is the reason for the requirement on the predicts to compensate for the velocity and higher degree components of the pointing axis angle motion.

D. System Noise Gain

As shown in Fig. 3, system generated noise, $N(z)$, acts as an input to the channel loop and causes the loop error, $E(z)$, to jitter. The root mean square (RMS) value of the output jitter, σ_E , is proportional to σ_N , the RMS value of $N(z)$. Setting $\phi(s)$ equal to zero in Eq. (10) gives

$$E(z) = - \frac{N(z) k_h k_d z}{z - 1 + k_T} \quad (23)$$

The sampled-data power spectrum (Ref. 4) of $E(z)$ is $\Phi_E(z)$ and is given by

$$\Phi_E(z) = \langle E(z^{-1}) E(z) \rangle = - \frac{\Phi_N(z) k_h^2 k_d^2 z}{(1 - k_T) \left(z - \frac{1}{1 - k_T} \right) (z - 1 + k_T)} \quad (24)$$

where $\Phi_N(z)$, the power spectrum of $N(z)$, is

$$\Phi_N(z) = \langle N(z^{-1}) N(z) \rangle \quad (25)$$

Braces indicate ensemble average. For independent noise samples of $N(z)$, the sampled-data autocorrelation function, $\beta_N(kP)$, of $N(z)$ is

$$\beta_N(kP) = \begin{cases} \sigma_N^2 & k = 0 \\ 0 & k \neq 0 \end{cases} \quad (26)$$

The sampled-data power spectrum $\Phi_N(z)$ of $N(z)$ is the two-sided Z -transform of $\beta_N(kP)$. Therefore $\Phi_N(z)$ is

$$\Phi_N(z) = \sum_{k=-\infty}^{\infty} \beta_N(kP) z^{-k} = \sigma_N^2 \quad (27)$$

Eq. (24) becomes

$$\Phi_E(z) = - \frac{\sigma_N^2 k_h^2 k_d^2 z}{(1 - k_T) \left(z - \frac{1}{1 - k_T} \right) (z - 1 + k_T)} \quad (28)$$

The inverse Z -transform of $\Phi_E(z)$ is

$$\beta_E(kP) = \frac{1}{2\pi i} \oint dz \Phi_E(z) z^{k-1} \quad (29)$$

where the path of integration in the Z complex plane is counterclockwise on the unit circle, and i is the positive square root of minus one. $\beta_E(kP)$ is the sampled-data autocorrelation function of $E(z)$, and kP is correlation lag time. σ_E^2 is found from

$$\sigma_E^2 = \beta_E(0) = \frac{1}{2\pi i} \oint \frac{dz}{z} \Phi_E(z) \quad (30)$$

Substituting Eq. (28) into Eq. (30), and using the method of residues about the pole z equal to $1 - k_T$ gives

$$\sigma_E = \frac{\sigma_N}{k_T P (1 - A)} \sqrt{\frac{k_T}{2 - k_T}} \quad (31)$$

In agreement with root locus theory, and the response found for a step input, it is seen that the jitter on $E(z)$ becomes infinite as k_T approaches the value of two. Actual

selection of system parameters depends upon the specific relationship of k_i and σ_N to the receiving system being used.

IV. Spacecraft Tracking Correlation

A. Derivation of Equations

A simplified block diagram of the spacecraft tracking receiver system is shown in Fig. 4. The carrier, E_{ca} , of the spacecraft signal impinges on the DSN antenna with a value of

$$E_{ca} = \sqrt{2P_s} \cos \omega_0 t \quad (32)$$

where P_s is the power in the carrier, and ω_0 is the carrier radian frequency. The antenna has a power gain, G_p , given by

$$G_p = \left(\frac{1}{16}\right)^{\psi^2/W^2} \quad (33)$$

where ψ is the angle in degrees that the antenna points away from the tracking boresight, and W is the beamwidth in degrees between the half-power points. Thus the constant of $1/16$ is chosen to give a value of $1/2$ to G_p when the ratio of ψ to W is $1/2$. In terms of an exponential function, Eq. (33) becomes

$$G_p = e^{-\mu(\psi^2/W^2)} \quad (34)$$

$$\mu = 4 \ln 2 \quad (35)$$

where \ln is the natural logarithm. Eqs. (33), (34), and (35) assume a circular antenna pattern.

At point "a" of Fig. 4, the voltage V_a is

$$V_a(t) = E_a(t) + n_a(t) \quad (36)$$

$$E_a(t) = \sqrt{2P_s G_p} \cos \omega_0 t \quad (37)$$

Here, $E_a(t)$ is the signal voltage, and $n_a(t)$ is the noise voltage associated with the system temperature, T_s . $n_a(t)$ has a wide band two-sided power spectral density, $\Phi_a(\omega)$, given by

$$\Phi_a(\omega) = \frac{KT_s}{2} \quad (38)$$

where K is Boltzmann's constant equal to 1.3806×10^{-23} watt seconds/degree Kelvin. At point "b" in Fig. 4, the voltage V_b is

$$V_b(t) = E_b(t) + n_b(t) \quad (39)$$

$$E_b(t) = H \sqrt{2P_s G_p} \cos \omega_0 t \quad (40)$$

where $E_b(t)$, the signal portion of V_b , has a gain factor of H , the receiver voltage gain constant. $n_b(t)$, the noise portion of V_b , has a two-sided spectral density $\Phi_b(\omega)$ given by

$$\Phi_b(\omega) = \frac{H^2 KT_s}{2} \quad (41)$$

with a one-sided noise bandwidth of B hertz formed by the IF bandpass characteristic of the receiver. This type of noise, as derived in Appendix A of Ref. 5, can be represented as

$$n_b(t) = \sqrt{2} (n_1(t) \sin \omega_0 t + n_2(t) \cos \omega_0 t) \quad (42)$$

where Φ_1 and Φ_2 , the power spectral densities of $n_1(t)$ and $n_2(t)$, are

$$\Phi_1(\omega) = \Phi_2(\omega) = \Phi_b(\omega) \quad (43)$$

and $n_1(t)$ and $n_2(t)$ are independent baseband signals centered about zero frequency with a two-sided bandwidth of B hertz.

At point "c" in Fig. 4, the voltage V_c becomes

$$V_c(t) = E_c(t) + n_c(t) \quad (44)$$

$$E_c(t) = 2H \sqrt{P_s G_p} \cos^2 \omega_0 t \quad (45)$$

$$n_c(t) = 2(n_1(t) \sin \omega_0 t + n_2(t) \cos \omega_0 t) \cos \omega_0 t \quad (46)$$

High order trigonometric terms in Eqs. (45) and (46) are eliminated by the low pass filter in Fig. 4, and therefore the voltage V_d at point "d" becomes

$$V_d(t) = E_d(t) + n_d(t) \quad (47)$$

$$E_d(t) = H \sqrt{P_s G_p} \quad (48)$$

$$n_d(t) = n_2(t) \quad (49)$$

where $E_d(t)$ and $n_d(t)$ are respectively the signal and noise portions of $V_d(t)$. This voltage is sampled by the antenna pointing system computer and a correlation summation is formed which closely approximates the continuous integral shown in the correlator box of Fig. 4. At the end of each period P , an estimate, X_2 , of the offset error is provided such that the mean, X_1 , is

$$X_1 = k_c \int_{AP}^P dt E_d(t) \sin \omega_m t \quad (50)$$

and the noise portion, N , is

$$N = k_c \int_{AP}^P dt n_d(t) \sin \omega_m t \quad (51)$$

where k_c is the correlator gain constant. The cross-elevation correlator integral would contain the terms $\cos \omega_m t$ instead of the term $\sin \omega_m t$.

B. Correlation Gain Constant

G_p in Eq. (48) is a function of time t due to the action of the conical-scan motion of the antenna. This motion produces a time variation on ψ^2 in Eq. (34) given by

$$\psi^2(t) = (\theta + R \cos \omega_m t)^2 + (\phi + R \sin \omega_m t)^2 \quad (52)$$

where θ and ϕ are respectively the cross-elevation and elevation errors, assumed constant for $AP < t < P$. $R \cos \omega_m t$ and $R \sin \omega_m t$ are the conical-scan drive equations to the assumed cross-elevation and elevation antenna pointing axes. When the offsets are zero, the scan radius R is the constant angle with which the antenna rotates about the tracking boresight. Using Eqs. (52) and (34), Eq. (48) becomes

$$E_d(t) = H \sqrt{P_s} \exp \left[-\frac{\mu}{W^2} (R^2 + \theta^2 + \phi^2 + 2R(\theta \cos \omega_m t + \phi \sin \omega_m t)) \right] \quad (53)$$

From tables of integrals, Ref. 6, the following equation may be derived

$$\int_c^{c+2\pi} dx \exp (A \cos X + B \sin X) \sin X = \frac{2\pi B}{\sqrt{A^2 + B^2}} I_1 \left(\sqrt{A^2 + B^2} \right) \quad (54)$$

where I_1 is the modified Bessel function of order one, and c is an arbitrary constant. Using Eqs. (53) and (54), Eq. (50) is evaluated to be

$$X_1 = -k_c H \sqrt{P_s} P(1 - A) \exp -\frac{1}{2W^2} \left[(R^2 + \theta^2 + \phi^2) \right] \times \frac{\phi}{\sqrt{\phi^2 + \theta^2}} I_1 \left(\frac{R\mu}{W^2} \sqrt{\phi^2 + \theta^2} \right) \quad (55)$$

If the assumption is made that θ and ϕ are much smaller than R , the approximation of

$$e^{-X} \approx 1 - X \quad (56)$$

may be made to Eq. (50) before evaluation of the integral, or the approximation of

$$I_1(X) \approx \frac{X}{2} \quad (57)$$

may be made to Eq. (55). In either case, Eq. (55) simplifies to

$$X_1 = -k_c H \sqrt{P_s} P(1 - A) \frac{R\mu}{2W^2} \exp \left(-\frac{\mu R^2}{2W^2} \right) \phi, \quad \phi, \theta \ll R \quad (58)$$

From Fig. 2, the correlator response is

$$X_1(s) = E(s) \frac{k_i}{s} (1 - e^{-sP(1-A)}) \quad (59)$$

For a constant input, $-\phi$

$$E(s) = -\frac{\phi}{s} \quad (60)$$

Therefore Eq. (59) becomes

$$X_1(s) = -\frac{\phi k_i}{s^2} (1 - e^{-sP(1-A)}) \quad (61)$$

Taking the Z-transform of Eq. (61) gives

$$X_1(Z) = -\frac{\phi k_i P(1-A)}{z-1} \quad (62)$$

The inverse Z-transform of Eq. (62) is

$$X_n = -\phi k_i P(1-A) \quad n = 1, 2, \dots \quad (63)$$

Comparing Eqs. (58) and (63) gives the value for the correlation constant k_i .

$$k_i = k_c H \sqrt{P_s} \frac{R\mu}{2W^2} \exp\left(-\frac{\mu R^2}{2W^2}\right) \quad (64)$$

C. Noise Variance

The variance σ_N^2 of the noise N of Eq. (51) is

$$\sigma_N^2 = \langle N^2 \rangle \quad (65)$$

Substituting Eq. (51) into Eq. (65) gives

$$\sigma_N^2 = k_c^2 \int_{AP}^P dt \int_{AP-t}^{P-t} d\tau_1 \beta_d(\tau_1) \sin \omega_m t \sin \omega_m (t + \tau_1) \quad (66)$$

where $\beta_d(\tau_1)$ is the autocorrelation of $n_d(t)$ as a function of the lag time τ_1 . β_d is the inverse Fourier transform of Φ_d .

$$\beta_d(\tau_1) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \Phi_d(\omega) \exp(i\omega\tau_1) \quad (67)$$

From Eqs. (41), (43), and (49)

$$\Phi_d(\omega) = \frac{H^2 K T_s}{2} \quad (68)$$

The bandwidth of $n_d(t)$ may be taken as infinite since Eq. (51) acts as an integrator with a resulting bandwidth much less than the B hertz specified for Eq. (41). Evaluating Eq. (67) using Eq. (68) gives

$$\beta_d(\tau_1) = \frac{H^2 K T_s}{2} \delta(\tau_1) \quad (69)$$

where $\delta(\tau_1)$ is the Dirac delta function, (Ref. 7). From Eqs. (66) and (69), σ_N is found to be

$$\sigma_N = \frac{k_c H}{2} \sqrt{K T_s P(1-A)} \quad (70)$$

and therefore, Eq. (31) becomes

$$\sigma_E = \frac{W^2 \exp(\mu R^2/2W^2)}{\mu R} \sqrt{\frac{K T_s}{P_s P(1-A)}} \sqrt{\frac{k_T}{2 - k_T}} \quad (71)$$

D. Spacecraft Correlation S-Curve

The transfer function specified by Eq. (58) is true only in the region where ϕ is close to zero. As the offset error ϕ departs from zero, a closer approximation to Eq. (55) is found by including the ϕ^2 value in the exponential term.

$$X_1 = -k_c H \sqrt{P_s} P(1-A) \frac{R\mu}{2W^2} \times \exp\left(-\frac{\mu R^2}{2W^2}\right) \phi \exp\left(-\frac{\mu \phi^2}{2W^2}\right) \quad (72)$$

Normalizing Eq. (72) gives

$$y = -\frac{2X_1 W \exp(\mu R^2/2W^2)}{k_c H \sqrt{P_s} P(1-A) R\mu} = \frac{\phi}{W} \exp\left[-\frac{\mu}{2} \left(\frac{\phi}{W}\right)^2\right] \quad (73)$$

A plot of Eq. (73) is shown in Fig. 5. Also shown in Fig. 5 are values for the slope of Y . This slope is a multiplying factor to

k_i for a large deviation of ϕ . Examination of Fig. 5 shows that the slope error is less than 20% for values of ϕ less than 0.22.

V. Selection of Parameters

In this section, a rationale will be developed for selection of parameters which will provide in a defined sense, an optimum tradeoff between the system time constant τ , and the scan radius R . R results in a loss of signal strength called crossover loss due to the fact that the antenna is pointing R degrees away from the target. This loss is the value of G_p when R is substituted for ψ in Eq. (34). Thus the crossover loss, C_L , in dB is given by

$$C_L = \left(\frac{R}{W}\right)^2 40 \log_{10} 2 \quad (74)$$

Normally it is desirable for C_L to be in the order of 0.1 dB. Increasing the value of R will result in excessive crossover loss. However, Eq. (71) shows that for weak signal levels, low values of R will result in excessive jitter in the error. If some function, $F(R/W)$, can be found which combines the effects of both crossover loss and error jitter, then an optimum R can be determined which will minimize F . A recommended function for this purpose is

$$F\left(\frac{R}{W}\right) = \frac{R}{W} + 3 \frac{\sigma_E}{W} \quad (75)$$

Thus, statistically speaking, the antenna pointing loss for both R and error jitter will be less than F for 99.7% of the time. Using Eq. (71), Eq. (75) and its derivative become

$$F\left(\frac{R}{W}\right) = \frac{R}{W} + \frac{D \exp\left[\frac{\mu}{2}\left(\frac{R}{W}\right)^2\right]}{\frac{R}{W}} \quad (76)$$

$$F'\left(\frac{R}{W}\right) = 1 + \frac{D \exp\left[\frac{\mu}{2}\left(\frac{R}{W}\right)^2\right] \left[\mu\left(\frac{R}{W}\right)^2 - 1\right]}{\left(\frac{R}{W}\right)^2} \quad (77)$$

$$D = \frac{3}{\mu} \sqrt{\frac{KT_s}{P_s P(1-A)}} \sqrt{\frac{k_T}{2 - k_T}} \quad (78)$$

Solving for F and D in Eqs. (76) and (77) gives the minimum value of F and the optimum value of D as a function of R .

$$F\left(\frac{R}{W}\right) = \frac{R}{W} \left(1 + \frac{1}{1 - \mu\left(\frac{R}{W}\right)^2}\right) \quad (79)$$

$$D = \frac{\left(\frac{R}{W}\right)^2 \exp\left[-\frac{\mu}{2}\left(\frac{R}{W}\right)^2\right]}{1 - \mu\left(\frac{R}{W}\right)^2} \quad (80)$$

For selected values of T_s , P_s , P , and A , the optimum value of k_T may be found from Eqs. (78) and (80). The system time constant τ is then found from Eq. (19). Fig. 6 is a plot of these results and shows system time constant τ versus crossover loss in dB from Eq. (74) for selected values of T_s , and P_s in dBm. The dependence of τ on P is very weak and therefore a value of 30 seconds is selected as being satisfactory for all cases in practise. The value of AP equal to two seconds is also selected as being used in the current system (Ref. 1).

The values in Fig. 6 are actually the constraints imposed upon the conical-scan system by the weakness of the received signal. For example, if a 25^k receiver is to receive a signal at -175 dBm, with a system time constant less than 400 seconds, then the optimum crossover loss is 0.16 dB. Alternatively, if the crossover loss is constrained to be 0.10 dB, then the system time constant must be set to 950 seconds to avoid excessive jitter in the error offsets.

Setting the system time constant and crossover loss for a given weak signal, and then receiving a stronger signal will simply result in less error jitter if the proper procedure is followed. From Eqs. (11) and (64), it is seen that the system gain is proportional to the square root of the received signal power. Thus the system gain and hence the time constant τ should be adjusted for the actual signal strength being received. One method for such adjustment is to open the loop, put in a known offset ϕ and measure the average correlator output X_1 . The value of k_i can then be found from Eq. (63). The values of k_T and τ are then found from Eqs. (11) and (19).

In the case of strong signal reception, the recommended parameters for conservative operation of the conical-scan system are a time constant of 90 seconds and a crossover loss of 0.09 dB. A much smaller time constant, or a lower crossover loss, might have a negative effect on the antenna mechanisms and increase jitter caused by imperfect mechanical

components. After selecting the desired values of τ and cross-over loss, C_L , the corresponding values of k_T and R may be found by inverting Eqs. (19) and (74).

$$k_T = 1 - \exp\left(-\frac{P}{\tau}\right) \quad (81)$$

$$R = W \left(\frac{C_L}{40 \log_{10} 2} \right)^{\frac{1}{2}} \quad (82)$$

Since Eq. (79) shows a fixed relationship exists between F and R , independent of T_s and P_s , the value of F in dB is also given in Fig. 6 as a horizontal scale beneath C_L according to an equation similar to Eq. (74).

$$F_{dB} = F^2 40 \log_{10} 2 \quad (83)$$

VI. Radio Source Tracking Correlation

A. Correlation Gain Constant

The development of equations for radio source tracking follows along lines similar to those for spacecraft tracking. Figure 7 is a simplified block diagram of the radio source tracking receiver system. Wide band energy from the radio source impinges upon the antenna with an equivalent noise temperature of T_r degrees. At point "e" of Fig. 7, the voltage $V_e(t)$ is random noise

$$V_e(t) = n_e(t) \quad (84)$$

with Φ_e , the two-sided spectral density of $n_e(t)$, being

$$\Phi_e(\omega) = \frac{KT_i}{2} \quad (85)$$

$$T_i = T_s + G_p T_r \quad (86)$$

where T_i is the total input temperature, T_s is the system temperature, and G_p is the antenna power gain defined in Eq. (34). $V_f(t)$, the voltage at point "f" in Fig. 7, is also random noise

$$V_f(t) = n_f(t) \quad (87)$$

$$\Phi_f(\omega) = \frac{H^2 KT_i}{2} \quad (88)$$

The two-sided spectral density Φ_f is band limited by the one-sided IF receiver bandwidth of B hertz. $V_g(t)$, the voltage at point "g" of Fig. 7, is formed by the squaring operation of the power detector

$$V_g(t) = E_g(t) + n_g(t) \quad (89)$$

$$E_g(t) = \alpha H^2 K T_i B \quad (90)$$

$$\Phi_g(0) = \alpha^2 H^4 K^2 T_i^2 B \quad (91)$$

where $E_g(t)$ is a DC voltage proportional to the input power, α is the power detector gain constant, $n_g(t)$ is a noise voltage riding on $E_g(t)$, and $\Phi_g(0)$ is the spectral density of $n_g(t)$ about zero frequency. High frequency terms from the squaring process are filtered out by a low pass filter internal to the power detector.

After each period P , the correlator output X_2 is

$$X_2 = X_1 + N \quad (92)$$

$$X_1 = k_c \int_{AP}^P dt E_g(t) \sin \omega_m t \quad (93)$$

$$N = k_c \int_{AP}^P dt n_g(t) \sin \omega_m t \quad (94)$$

Using Eqs. (34), (52), (54), (86), (90), and (93), X_1 is found to be

$$X_1 = -k_c \alpha H^2 K B P (1 - A) T_r \exp \left[-\frac{\mu}{W^2} (R^2 + \phi^2 + \theta^2) \right] \\ \times \frac{\phi}{\sqrt{\phi^2 + \theta^2}} I_1 \left(\frac{2R\mu}{W^2} \sqrt{\phi^2 + \theta^2} \right) \quad (95)$$

For ϕ and θ much less than R , the approximation of Eq. (56) or of Eq. (57) may be used to linearize X_1

$$X_1 = -k_c \alpha H^2 K B P (1 - A) T_r \frac{R\mu}{W^2} \exp\left(-\frac{\mu R^2}{W^2}\right) \phi \quad (96)$$

Comparison with Eq. (63) gives the constant k_i

$$k_i = \frac{k_c \alpha H^2 K B T_r \mu}{W^2} \exp\left(-\frac{\mu R^2}{W^2}\right) \quad (97)$$

The S-curve constraint for larger values of ϕ is

$$y = -\frac{X_1 W \exp\left(\frac{\mu R^2}{W^2}\right)}{k_c \alpha H^2 K B P (1 - A) T_r R \mu} = \frac{\phi}{W} \exp\left[-\mu \left(\frac{\phi}{W}\right)^2\right] \quad (98)$$

and is also plotted in Fig. 5.

B. Radio Source Tracking Noise

The noise variance σ_N^2 is found by substituting Eq. (94) into Eq. (65) to give

$$\sigma_N^2 = k_c^2 \int_{AP}^P dt \int_{AP-t}^{P-t} d\tau_1 \beta_g(\tau_1) \sin \omega_m t \sin \omega_m (t + \tau_1) \quad (99)$$

where $\beta_g(\tau_1)$ is the autocorrelation of $n_g(t)$ as a function of the lag time τ_1 . Using the same reasoning as was done for spacecraft tracking, $\beta_g(\tau_1)$ is found to be

$$\beta_g(\tau_1) = \alpha^2 H^4 K^2 T_i^2 B \delta(\tau_1) \quad (100)$$

where T_i is now taken to be

$$T_i = T_s + \exp\left(-\frac{\mu R^2}{W^2}\right) T_r \quad (101)$$

Eqs. (99), (100), and (101) give

$$\sigma_N = k_c \alpha H^2 K \left(T_s + \exp\left[-\frac{\mu R^2}{W^2}\right] T_r\right) \sqrt{\frac{B P (1 - A)}{2}} \quad (102)$$

from which Eq. (31) becomes

$$\sigma_E = \frac{W^2 \left(T_s \exp\left[\frac{\mu R^2}{W^2}\right] + T_r\right)}{T_r R \mu \sqrt{2 B P (1 - A)}} \sqrt{\frac{k_T}{2 - k_T}} \quad (103)$$

For normal radio star sources, the large value of B , typically 10^7 hertz, makes σ_E insignificant. Therefore the conservative parameters recommended for strong signal spacecraft tracking are also recommended for radio source tracking.

VII. Conclusion

The conical-scan tracking system has been analyzed in terms of a sampled-data feedback control system. This can facilitate further investigation into such questions as response to ramp inputs, and design of a higher order loop. Moreover, a rationale has been developed for selection of system parameters in both the spacecraft and radio source tracking modes. In the case of spacecraft tracking it is shown that reasonable tracking performance can be obtained at signal levels down to receiver threshold. Appendix A lists the symbols used in this report and the equations in which they first appear. Appendix B gives the correspondence between selected symbols of this report and symbols of Ref. 1.

ORIGINAL PAGE IS
OF POOR QUALITY

References

1. Ohlson, J. E., and Reid, M. S., "Conical-Scan Tracking With the 64-m-diameter Antenna at Goldstone," Technical Report 32-1605, Jet Propulsion Laboratory, Pasadena, California, October 1, 1976.
2. Jury, E. I., *Theory and Application of the Z-Transform Method*, John Wiley & Sons, Inc., New York, 1964.
3. Mason, S. J., "Feedback Theory – Further Properties of Signal Flow Graphs," Proceedings of the Institute of Radio Engineers, Vol. 44, July 1956.
4. Nishimura, T., and Jury, E. I., "Contribution to Statistical Designing of Sampled-Data Control Systems," Electronics Research Laboratory Report Series 60, Issue No. 210, August 5, 1958, University of California, Berkeley.
5. Viterbi, A. J., *Principles of Coherent Communication*, McGraw-Hill Book Co., New York, 1966.
6. Gradshteyn, I. S., and Ryzhik, I. M., *Table of Integrals, Series, and Products*, #3.937, Academic Press, New York, 1965.
7. Davenport, W. B., and Root, W. L., *An Introduction to the Theory of Random Signals and Noise*, Appendix 1, McGraw-Hill Book Co., New York, 1958.

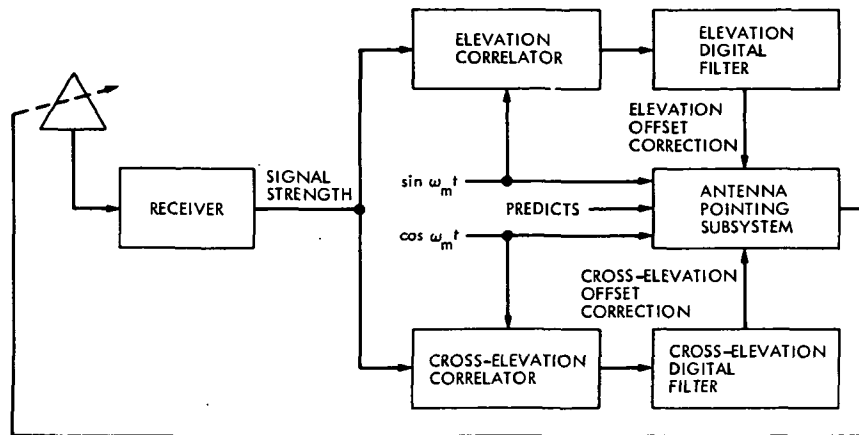


Fig. 1. System block diagram

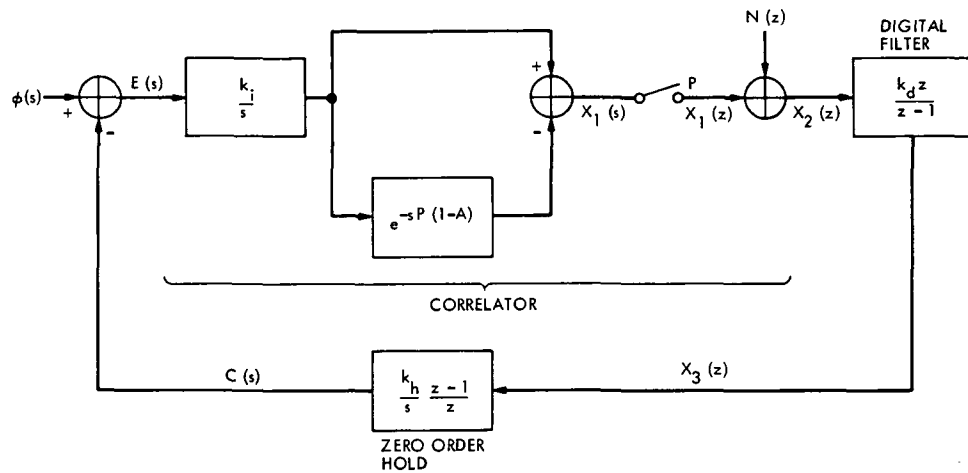


Fig. 2. Laplace block diagram

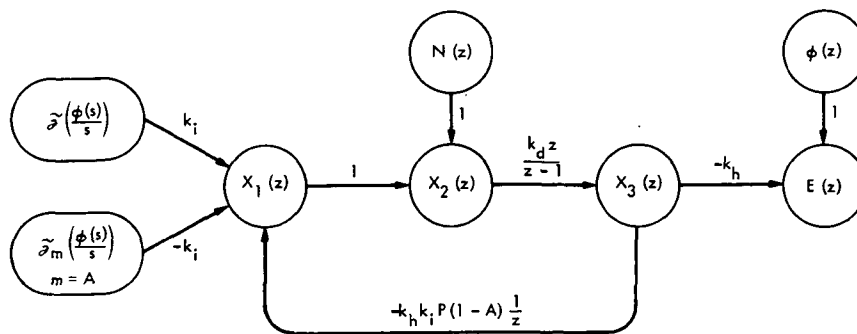


Fig. 3. Signal flow graph

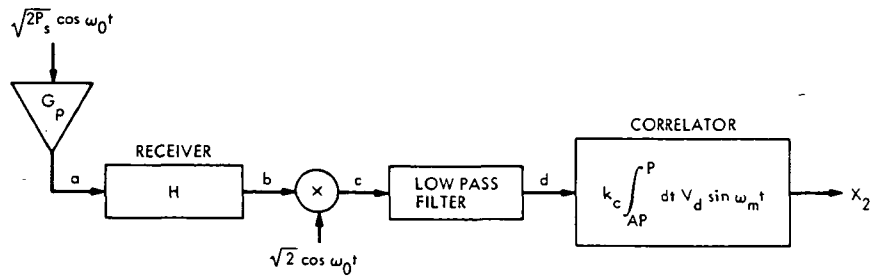


Fig. 4. Spacecraft tracking receiver system

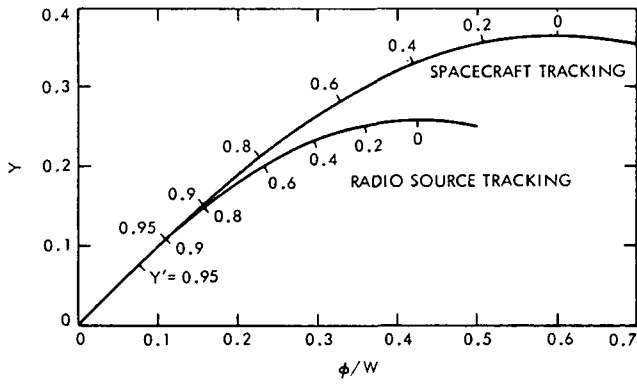


Fig. 5. Correlation S-curves

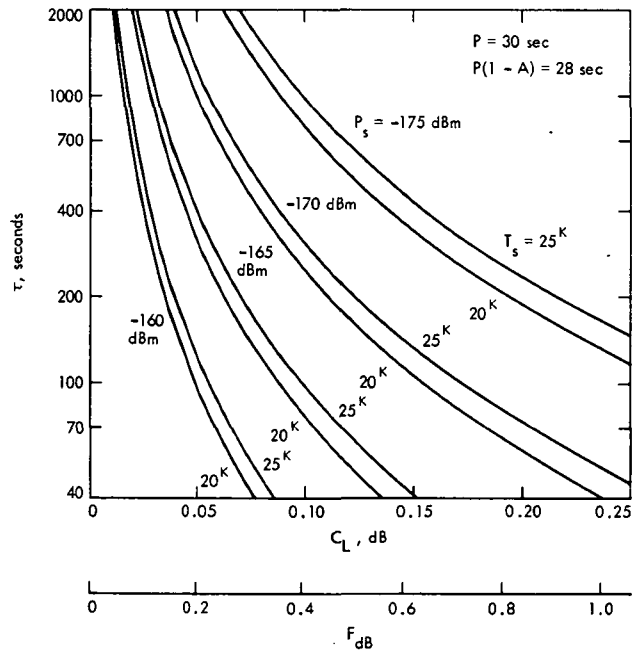


Fig. 6. Optimum conical-scan parameters for spacecraft tracking

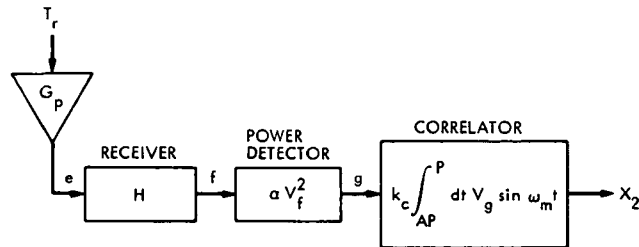


Fig. 7. Radio source tracking receiver system

Appendix A

List of Symbols

Following is a list of symbols used in this report and the equation number in which they first appear.

A	(1)	n_a	(36)	β_d	(66)
$C(s)$	(3)	n_b	(39)	β_E	(29)
C_L	(74)	n_c	(44)	β_g	(99)
D	(76)	n_d	(47)	β_N	(26)
$E(s)$	(3)	n_e	(84)	δ	(69)
$E(z)$	(8)	n_f	(87)	θ	(52)
E_a	(36)	n_g	(89)	μ	(34)
E_b	(39)	P^g	(1)	σ_E	(30)
E_c	(44)	P_s	(32)	σ_N	(26)
E_{ca}	(32)	R^s	(52)	τ	(18)
E_d	(47)	s	(3)	τ_1	(66)
E	(89)	T_i	(86)	ϕ	(52)
F^g	(75)	T_r	(86)	ϕ_c	(12)
F_{dB}	(83)	T_s	(38)	$\phi(s)$	(3)
G_{cl}	(22)	V_a	(36)	$\phi(z)$	(8)
G_{ol}	(20)	V_b	(39)	Φ_1	(43)
G_p	(33)	V^c	(44)	Φ_2	(43)
H	(40)	V_d	(47)	Φ_a	(38)
K	(38)	V^e	(84)	Φ_b	(41)
k	(2)	V^f	(87)	Φ_E	(24)
k_c	(50)	V^g	(89)	Φ_e	(85)
k_d	(2)	W^g	(33)	Φ_f	(88)
k_h	(5)	X_1	(5)	Φ_g	(91)
k_i	(5)	X_2	(2)	Φ_N	(24)
k_T	(11)	X_3	(2)	ω_m	(1)
N	(6)	Y	(73)	ω_0	(32)
n_1	(42)	z	(4)	ψ	(33)
n_2	(42)	α	(90)		

ORIGINAL PAGE IS
OF POOR QUALITY

Appendix B

Correspondence With TR 32-1605

A correspondence is given between selected symbols of Technical Report 32-1605 and the equivalent symbol or expression in this report. Numbers in parentheses are the equation numbers in TR 32-1605 in which the symbols first appear.

TR 32-1605	This Report
A (17)	$k_i P(1 - A)$
B (11)	B
C (11)	$k_c H^2 K$
g (10a)	G_p
h (15)	$k_h k_d$
K (73)	$k_c H$
k (99)	K
N_O (65)	$K T_s$
P (15)	$P(1 - A)$
P_S (61)	P_s
R (1)	R
r (27)	$1 - k_T$
T_i (10a)	T_i
T_{OP} (10a)	T_s
T_S (10a)	T_r
W (49)	W
β (3)	ψ
ϕ (2)	ϕ
σ_N (22)	σ_N
σ_ϕ (38)	σ_E
τ (35)	τ
ω_m (1)	ω_m

D/12
N78-11144

Antenna Pointing Subsystem Drive Tape Generator — Another New Interface

R. F. Emerson

Communications Systems Research Section

The Antenna Pointing Subsystem (APS) program was changed to meet new needs. To continue service to radio science targets, a new interface between the ephemerides of these targets and the APS program was required. This report describes briefly the program, called APSTAPE, written for this purpose. An important feature of the program is that it performs a read-after-punch operation, which ensures that the drive tapes generated will be error free. The program described has been successfully used in the field.

I. Introduction

Antenna pointing at a Deep Space Station (DSS) is performed by the Antenna Pointing Subsystem (APS). The APS consists of hardware, an XDS 910 computer and special peripherals, and software, the APS Program. Input to the APS Program is delivered by the Tracking Predicts System for most project requirements. For all scheduled tracking, the APS is operated by the station staff.

The tracking of both planetary and spacecraft targets for radio science activities requires an interface between the ephemeris and the APS Program. For radio science work, the ephemeris is delivered as a Chebychev polynomial while the APS Program requires a point-by-point type of ephemeris.

The APS Program was changed to meet new needs. One of the side effects of this change was that the drive tape interface (paper tape) was altered. A program for the conversion to the new format was needed. APSTAPE was written to perform the conversion using the XDS 930 computer, which supports radio

science activities. The following is a description of that program.

II. Interface Program Requirements

The requirements of the interface program can be broken into four sections: (1) input, (2) output, (3) operations, and (4) verification.

The input polynomial coefficients may be delivered to the station on cards, in the RADAR 11 format, or on magnetic tape, in the Ephemeris Delivery System (EDS) format. Since more than one pass or target may be delivered in such a file, parameters input by the station operators are required for the selection of the proper set. Additional entries are passed to the drive tape to be used by the APS Program operator. This data is used to verify that the correct drive tape is being used.

The primary output of APSTAPE is a valid drive tape for the APS XDS 910. While not necessary, it has been shown

desirable to provide the APSTAPE operator with listings of pertinent data. These listings permit the operator to monitor the APSTAPE operation and performance, and thereby provide the "warm feeling" necessary for effective operation.

Since the station staff for radio science is limited and is therefore highly "cross-trained", it is extremely important that the APSTAPE Program operation be straightforward and self-explanatory. Program-generated messages covering all normal and contingency operations are required. To the degree possible, the operating instructions for the program should be contained within the program itself. Satisfying this requirement leads to smooth operations even when the use of the program is occasional.

While paper tape itself is a reliable medium, its generation suffers from occasional failures. To eliminate the propagation of this type of error into the antenna pointing process, the output tape must be verified.

III. APSTAPE Program Description

The APSTAPE Program as written satisfies the above requirements. It is a "self-loading" program that guides an operator through the process by typed messages and requests. Verification of the output is done by reading the punched paper tape and comparing against the desired output. Since the paper tape punch does not have read after write, enough records are punched so that the tape reaches the paper tape reader. The operator is then requested to load the tape into the reader for the verification process. Should the tape fail verification, the operator is directed in a recovery procedure by the program. When a completed and verified tape has been produced, the operator is provided with a label to attach to the tape.

The program description is presented in three sections: (1) initialization, (2) processing, and (3) termination.

A. Initialization

At the start of the program execution, the variables are initialized and control and data information is accepted. Because of the separation of the punch and reader, a delay in the verification process must be set. A circular buffer, which contains the delayed data until verification, is cleared and its pointers are reset. At the completion of these actions, the operator is requested to enter the date, target, station, pass and predict set identifications, the step size, and the coefficient input medium.

With this information available, the program will scan the coefficient file for the proper data. The input data is checked

for consistency as it is entered. Errors result in requests for specific corrective action. A monitor output of the data to be used is listed. When this is complete, the execution automatically enters the processing phase.

B. Processing

In the processing part of the program, the drive tape points are computed, formatted into records, and punched into paper tape. The punched records are read and verified against a stored image.

The APS drive tape starts at the rise time of the polynomial and continues to the set time. Points are computed at intervals of the input step size. These values are formatted into records containing two sets of day, time, hour angle, and declination. It is not necessary that the step size divide the pass period evenly. If an odd number of points results, the last set is identically zero (day, time, and angles). This preserves the record format. After formatting, the record is punched. It is also saved in the circular buffer and listed for monitoring.

If verification is permitted (i.e., there is enough tape punched to reach from the punch to the reader), a record is read from the tape. Should it not match the contents of the circular buffer, the operator is directed to discard the offending record and all the tape between the reader and the punch. A new section is punched and the operator is directed to splice it to the verified tape. The splice is done by simply overlapping blank sections of tape and taping together. The program waits for each operator action.

This process continues until set time is encountered.

C. Termination

At the end of a pass, the drive tape must have an End-of-File record and some leader punched. Then the remaining records must be verified. As an aid to operations, a label is printed containing the necessary information about the drive tape generated. This label is attached to the tape by the operator.

Following a successful run, the program is reentered for a second one.

IV. Conclusions

The program has been successfully used to provide drive tapes for both the moon and a synchronous satellite. It is available from the DSN Program library under number DXI-5490-SP.

N78-11145

An MBASICTM User Profile

R. L. Schwartz
DSN Data Systems Section

An important aspect of any programming language development is an awareness of how the language is actually being used. This article presents the preliminary results of an empirical study of MBASIC user programs. The study, part of an ongoing effort, seeks to discover the features and capabilities of the MBASIC processor which are being utilized. This will aid in future decisions to be made concerning language extensions and the development of a batch compiler.

I. Introduction

Designers of languages and compilers usually have comparatively little information about the way in which programming languages are actually used by typical programmers. Designers have some view of what programmers using the language do, but this is rarely based on a representative sample of the programs which are actually being written and used.

There has been widespread realization that more data about language use is needed. The modification of programming languages and compilers for programming languages should take place with a clear picture of what elements of the language are being used. Too often in the history of programming language and compiler development, decisions have been based on a feeling on the part of the language designers or compiler writers of what features are most important. It is easy to fall in the trap of assuming that complicated constructions are the norm when in fact they are infrequently used. Such mistaken assumptions can lead to misguided optimization efforts and misinterpretation of user needs.

Empirical studies of user programs written in the programming language FORTRAN by Knuth (Ref. 1) and in the language ALGOL 60 by Wichmann (Ref. 2) have been previously made. The studies have analyzed user programs in vari-

ous ways. A *static* profile provided information about the frequency of occurrence of various statements and constructs, complexity of control flow, as well as other data which can be gathered from a program listing. A *dynamic* profile gave information related to frequency of execution of program statements, and how the execution time was distributed among the program statements. The overall conclusions of the studies showed that the vast majority of the statements in most programs are of a very simple nature, and that the bulk of the program execution time is spent executing a very small percentage of the program statements.

Such a profile of typical programs written in the MBASIC computer programming language has been initiated at JPL. The MBASIC language is an advanced version of BASIC designed at the Laboratory in the early part of 1971. Presently an MBASIC interpreter is implemented on the UNIVAC 1108 and DECsystem-10 computers. It has recently been recommended as the standard nonreal-time language for use in Deep Space Network program development. Intended to be a simple easy-to-learn general-purpose programming language, it is used extensively throughout the Laboratory for management, scientific, and light computing applications. References 3 and 4 give a general introduction to the features and capabilities of the MBASIC system.

II. The Study Performed

A preliminary *static* profile of MBASIC user programs analyzing a total of 114 programs stored in the MBASIC user library has been made, with programs chosen randomly from approximately 800 user program files. The statistics described in the following sections attempt to analyze the use of different statement types, modifiers, and the use of the backslash statement regenerator. These statistics were compiled automatically by a computer program written in the MBASIC language. Direct analysis of selected programs was used to aid in the interpretation of the results.

A. Overall Statement Usage

A top-level statement usage breakdown of the 14,908 statements appearing in the 114 MBASIC user files is given in Table 1. This statement breakdown does not include statements embedded in an IFTHENELSE statement (e.g., the GOTO statement embedded in IF B THEN GOTO 100). Statistics for this were compiled separately, and are reported in the next section. In the case where there are synonymous keywords (i.e., PRINT/WRITE, and STOP/END), the combined statistics are reported with the individual statistics given in a footnote. The assignment statement is somewhat of a special case: it may be initiated by a LET keyword for one or more simple assignments, by an EXCHANGE keyword for one or more variable exchanges, or without a keyword for a combination of assignments and exchanges. Each of these cases is shown separately in Tables 1 to 5.

A special analysis was made for the GOTO and GOSUB statements. The specification of the line number for the branching operation can, in general, be an arbitrary expression evaluating to a real number. A check was made to see how often a simple constant line number was used rather than an expression. This is also reported in Table 1.

B. IFTHENELSE Statement Usage

Tables 2 and 3 show the frequency of occurrence of the different program statements in the THEN and ELSE clauses of the IFTHENELSE statement, respectively. The preliminary results presented herein only analyze *one nesting level*. Thus in a statement such as

```
IF D
  THEN
    IF C
      THEN A = 14
      ELSE B = 7
    ELSE PRINT A
```

the assignments to the variables A and B would not be recorded. The statement in the THEN-clause would be noted

as an IF statement, and the statement in the ELSE-clause would be noted as a PRINT statement. Future empirical studies may analyze arbitrary IFTHENELSE nesting levels.

C. Modifier Usage

Table 4 summarizes the results concerning the use of the WHERE, IF, UNLESS, and FOR modifiers in program statements. Again, only a preliminary analysis of the top-level statements was performed (e.g., the WAIT statement in IF B THEN WAIT UNLESS X is not analyzed for modifier usage).

D. Backslash Usage

Table 5 reports the results concerning the use of the backslash statement regenerator for each of the statement types. Again, only the top-level statements have been analyzed.

III. Analysis of Results

Before analyzing the results given in Tables 1 to 5, the reader should note that Table 1, giving the statement usage breakdown, may tend to present a slightly misleading view of the relative usage of statements in programs. One might register surprise (or even joy) at the fact that the GOSUB statement is used more frequently than the GOTO statement. While this is indeed true at the top-level of statement nesting, the table does not show the extremely heavy usage of GOTO statements in IFTHENELSE statements. When the figures in Table 1 are modified to include the second level statements in the THEN and ELSE clauses of the IF statement, the results then show: ASSIGN 28.3% (4975), IF 12.8% (2243), GOTO 12.6% (2213), PRINT/WRITE 12.5% (2193), and GOSUB 7.5% (1325). Thus, while the GOSUB is used more frequently as an unconditional statement, the GOTO is used more overall. This breakdown is in contrast to the results reported by Knuth concerning FORTRAN programs written at Lockheed. He found the following statements most heavily used: Assignment 41%, IF 14.5%, GOTO 13%, CALL 8%, and WRITE 4%. The non-numeric applications for typical MBASIC programs would seem to account for the higher frequency of output statements. Generally, however, the frequency of usage of MBASIC statements follows the usage of FORTRAN statements rather closely.

The APPEND, PAUSE, REMOVE, DATA, WIDTH, READ, RENAME, EXIT and RANDOMIZE statements were very seldom used, while the EXCHANGE, TAPE, and REMARK statements were never used in the samples studied. The relatively high position for the LET statement (0.5%) was found to be due to one program which was translated from the BASIC language.

Table 2 shows that almost half of all THEN clauses in IF statements were GOTO statements, with an additional one quarter being assignment statements. The use of the GOTO, assignment, and PRINT/WRITE statements accounted for 82% of all THEN clause statements. Multiple levels of nesting of the IF statement in the THEN clause accounted for about 6% of the uses of the IF.

The use of the ELSE clause of the IF statement was somewhat different from that of the THEN clause. Assignment statements accounted for about a third of all ELSE clauses, with the GOTO accounting for about another quarter, the multi-level IF structures nested in the ELSE clause accounting for about another fifth. The high percentage of nested IF statements is probably due to the lack of a CASE-type construct in the MBASIC language.

The 46% GOTO and 26% ASSIGN for THEN clauses compared to 32% ASSIGN and 25% GOTO for the ELSE clauses can probably be explained in terms of a lack of ability to associate more than one statement with a THEN or ELSE clause. Multi-statement THEN and ELSE clauses are simulated by having a sequence of the form:

```

ℓ      IF B THEN GOTO ℓ + i
ℓ + 1  S1-ELSE      !ELSE CLAUSE
      .
      .
      .
ℓ + i - 1 GOTO ℓ + j  !BRANCH OUT OF IFTHENELSE
ℓ + j  S1-THEN      !THEN CLAUSE
      .
      .
      .
ℓ + j  !END OF IFTHENELSE

```

Thus the ELSE clause which appeared 43% of the time that the IF statement was used, is only used for single statement clauses. This statement was most frequently an assignment statement (32% of the time).

The modifier usage, given in Table 4, showed a rather limited use of modifiers. The UNLESS modifier was infrequently used, with only six statement types making any use of the modifier (approximately 50% of that usage was with the GOSUB statement). The highest percentage of statements with modifiers employed is the GOSUB statement, with about half of the GOSUB statements using modifiers. Of these statements, about 72% were WHERE modifiers. This is due to a lack of a procedure mechanism that allows the passing of parameters. This has been simulated with a statement of the form:

```
GOSUB line WHERE PARM1 = val1, PARM2 = val2 . . .
```

For procedures that return values (functions), the passing of the return value has been simulated with a statement of the form:

```
RETURN WHERE ANSWER = val
```

This was done about 12% of the time the RETURN statement was used.

The use of the FOR modifier is confined mostly to input/output statements and the assignment statement, with 67% of the FOR modifiers being used for the PRINT/WRITE and INPUT statements, and 30% being used for initialization of arrays in assignment statements. It is interesting to note that a GOTO statement was used with a FOR modifier three times (although an explanation of this use cannot be given here).

The IF modifier had 93% of its usage confined to three types of statements: 33% to branching statements (GOTO, and GOSUB), 28% to assignment statements, and 32% to input/output statements (PRINT/WRITE, and INPUT).

The use of the backslash, illustrated by Table 5, is confined to only 5 statement types. Approximately 36% of the PRINT/WRITE statements used a backslash. This accounted for 71% of the backslash usage, with an additional 21% of the backslashes being used for assignment statements (to limit the scope of modifiers).

Of the two pairs of synonymous keywords, PRINT/WRITE and STOP/END, both showed heavy favorites. PRINT was used about 84% of the time instead of WRITE, while END was used 79% of the time instead of STOP.

IV. Conclusions

Depending on the type of information desired, there are many different ways the data presented in Tables 1 to 5 could have been analyzed. Further efforts are needed to arrive at a more comprehensive evaluation of the data.

Overall, the data suggests that the fancier or more complex features of the MBASIC language are not being utilized extensively. The statistics lend credence to the hypothesis that the users of the MBASIC language are attempting to simulate some of the elementary structured programming constructs such as procedures with parameters, an IFTHENELSE statement with multi-statement THEN and ELSE clauses, and a CASE statement.

Users have attempted to introduce techniques for greater abstraction and problem reduction by simulating procedures

with parameters, and using non-numeric GOSUB expressions. Some 38% of all GOSUB line expressions have used meaningful variable names to serve as procedure names (with the variable initialized to the proper line number elsewhere). These conclusions suggest that the MBASIC structured programming extensions will eliminate much of the present use of modifiers.

MBASIC users have not taken advantage of the ability to use modifiers and backslashes as a general language feature; rather, its use has been confined mainly to input/output statements and assignment statements, places where other languages have allowed a similar capability. The exception to this is the use of the WHERE modifier, but in that case it was used to circumvent the lack of adequate abstraction mechanisms.

V. Future Work

This article has presented only a preliminary look at what needs to be an ongoing effort. Only a limited static profile has thus far been performed.

Future work should take place in three areas: a more comprehensive static profile, the development of a dynamic profile, and, at a later time, the integration of such information into the design considerations attendant to the planned MBASIC batch compiler.

The static profile of MBASIC programs can be enhanced in many ways. A larger sample of user programs would certainly contribute to more representative data. The automatic measurement program should be expanded to analyze the use of arbitrary IFTHENELSE nesting levels, the composition of

assignment statements and FOR-NEXT loops (as was done for FORTRAN and ALGOL 60), and special characteristics of programs. In addition, these statistics should also report on the mean usage (with variances) of each of the language features in individual programs.

A dynamic profile can be developed, which gives information about frequency of execution of statements and the distribution of the execution time over the program statements. A more ambitious, and important, study would determine if users are taking advantage of the dynamic scope rules of the MBASIC language. As the language presently exists only in interpretive form, with incremental parsing, the inter-statement relationships, such as the association between a FOR statement and the corresponding NEXT statement, must be determined on the basis of execution order rather than textual order. An important question is whether MBASIC language users are using this dynamic association rule. This has important consequences for the development of an MBASIC compiler, since it is intended that most MBASIC programs shall function exactly the same in both interpretive and compiled modes.

A long range goal is to incorporate, at user option, such dynamic and static profiles into the batch monitor, as part of the user program development cycle. Static information obtained during compile time, and dynamic information obtained at run time can be used to direct program optimization during the next compilation. Thus, the compiler effort can be concentrated on those portions of the code which are most important to program function. Conceptually, this information could be used, via a feedback loop between the execution phase and compile phase of program development, to have the program execution history direct a heuristic optimization of the object code produced by the compiler.

References

1. Knuth, D. E., "An Empirical Study of FORTRAN Programs," *Software Practice and Experience*, Vol. 1, pp. 105-133, 1971.
2. Wichmann, B., *Algol 60 Compilation and Assessment*, Academic Press, N. Y., 1973.
3. *MBASIC Fundamentals, Vol. 1*, Feb. 1974 (JPL internal document).
4. Schwartz, R., *Syntactic Description of the MBASIC Language*, October 1976 (JPL internal document).

Table 1. Top-level statement usage breakdown

Type	% of total (actual No.)	Type	% of total (actual No.)
ASSIGN	28.2 (4209)	STOP/END	0.5 (79)
IF	13 (1944)	LET	0.5 (74)
PRINT/WRITE	12.9 (1928)	COPY	0.4 (57)
GOSUB	8.2 (1227)	ON	0.3 (43)
GOTO	7.5 (1112)	APPEND	0.2 (26)
INPUT	6.7 (999)	PAUSE	0.2 (24)
RETURN	3.7 (553)	REMOVE	0.2 (23)
FOR	3.5 (528)	DATA	0.1 (21)
NEXT	3.5 (527)	WIDTH	0.1 (8)
STRING	3.3 (496)	READ	0.1 (8)
OPEN	1.9 (276)	RENAME	0 (5)
CLOSE	1.9 (229)	EXIT	0 (3)
DIM	1.5 (223)	RNDMZ	0 (1)
REAL	0.7 (101)	EXCHANGE	0 (0)
AT	0.7 (100)	TAPE	0 (0)
WAIT	0.6 (84)	REMARK	0 (0)

File count: 114

Statement count: 14,908

GOTO with constant line expression: 97% (1075)

GOSUB with constant line expression: 62% (756)

{ PRINT	10.8%	(1613)
{ WRITE	2.1%	(315)
{ STOP	0.1%	(17)
{ END	0.4%	(62)

**ORIGINAL PAGE IS
OF POOR QUALITY**

Table 2. THEN clause statement usage breakdown

Type	% Total THEN clauses (actual No.)	Type	% Total THEN clauses (actual No.)
GOTO	45.8 (891)	TAPE	0 (0)
ASSIGN	25.8 (501)	AT	0 (0)
PRINT/WRITE	10 (204)	DATA	0 (0)
IF	6.2 (121)	NEXT	0 (0)
RETURN	4.5 (87)	APPEND	0 (0)
GOSUB	4 (77)	OPEN	0 (0)
INPUT	1.6 (31)	REAL	0 (0)
END/STOP	0.7 (13)	REMARK	0 (0)
FOR	0.5 (9)	EXIT	0 (0)
COPY	0.3 (6)	READ	0 (0)
ON	0.1 (2)	STRING	0 (0)
REMOVE	0.1 (1)	WIDTH	0 (0)
CLOSE	0.1 (1)	EXCHANGE	0 (0)
LET	0 (0)	RNDMZ	0 (0)
PAUSE	0 (0)	DIM	0 (0)
RENAME	0 (0)	WAIT	0 (0)

THEN Clause count: 1944

}	PRINT	7.8%	(152)
	WRITE	2.7%	(52)
}	END	0.6%	(12)
	STOP	0.1%	(1)

Table 3. ELSE clause statement usage breakdown

Type	% Total ELSE clause (actual No.)	Type	% Total ELSE clauses (actual No.)
ASSIGN	31.9 (265)	PAUSE	0 (0)
GOTO	25.3 (210)	RENAME	0 (0)
IF	21.4 (178)	TAPE	0 (0)
PRINT/WRITE	9 (75)	AT	0 (0)
COSUB	4 (21)	DATA	0 (0)
INPUT	3.7 (31)	NEXT	0 (0)
RETURN	3 (25)	WAIT	0 (0)
ON	0.8 (7)	OPEN	0 (0)
COPY	0.8 (7)	REAL	0 (0)
REMOVE	0.5 (4)	REMARK	0 (0)
FOR	0.4 (3)	EXIT	0 (0)
CLOSE	0.2 (2)	READ	0 (0)
STOP/END	0.2 (2)	STRING	0 (0)
APPEND	0.1 (1)	WIDTH	0 (0)
DIM	0 (0)	EXCHANGE	0 (0)
LET	0 (0)	RNDMZ	0 (0)

If statements with ELSE Clause: 43% (831)

{ PRINT	7.3%	(61)
{ WRITE	1.7%	(14)
{ STOP	0.1%	(1)
{ END	0.1%	(1)

**ORIGINAL PAGE IS
OF POOR QUALITY**

Table 4. Modifier usage by statement type

Type	% Using modifiers (actual No.)		% Total WHERE	% Total IF	% Total UNLESS	% Total FOR
GOSUB	51.2	(628)	72.5	13	13.5	.9
WAIT	40.5	(34)	0	0	100	0
GOTO	32.3	(359)	12.4	86.5	.3	.8
WRITE/PRINT	30.9	(596)	1.5	36.6	3.6	58.3
APPEND	30.8	(8)	0	25	62.5	12.5
WIDTH	25	(2)	0	100	0	0
INPUT	24.1	(241)	2.3	35.5	0	62.2
FOR	21.8	(115)	99.1	.9	0	0
REMOVE	21.7	(5)	0	80	0	20
READ	12.5	(1)	0	0	0	100
COPY	12.3	(7)	0	71.4	0	28.6
RETURN	11.8	(65)	66.2	33.8	0	0
ASSIGN	11.7	(493)	8.6	47.3	4.7	39.3
STRING	10.3	(51)	55.8	23.4	0	20.8
CLOSE	10	(23)	84	12	0	4
REAL	9.9	(10)	58.3	41.7	0	0
STOP/END	8.9	(7)	14.3	85.7	0	0
OPEN	8.3	(23)	0	48.3	0	51.7
PAUSE	4.2	(1)	100	0	0	0
ON	2.3	(1)	0	100	0	0
AT	2	(2)	0	100	0	0
DIM	1.8	(4)	0	100	0	0
NEXT	.6	(3)	100	0	0	0
RNDMZ	0	(0)	-	-	-	-
TAPE	0	(0)	-	-	-	-
EXCHANGE	0	(0)	-	-	-	-
LET	0	(0)	-	-	-	-
RENAME	0	(0)	-	-	-	-
<hr/>						
{ WRITE	50.5%	1.1%	47%	0%	51.9%	
{ PRINT	27.1%	1.7%	31.9%	5.2%	61.1%	
{ END	6.5%	0%	100%	0%	0%	
{ STOP	17.6%	33.3%	66.7%	0%	0%	

Table 5. Backslash usage by statement type

Type	% Using backslash (actual No.)	Type	% Using backslash (actual No.)
PRINT/WRITE	35.9 (693)	TAPE	0 (0)
STOP/END	8.9 (7)	AT	0 (0)
ASSIGN	6.1 (256)	CLOSE	0 (0)
INPUT	1.7 (17)	DIM	0 (0)
GOSUB	.2 (2)	EXCHANGE	0 (0)
APPEND	0 (0)	GOTO	0 (0)
COPY	0 (0)	LET	0 (0)
NEXT	0 (0)	OPEN	0 (0)
PAUSE	0 (0)	READ	0 (0)
REAL	0 (0)	RENAME	0 (0)
RNDMZ	0 (0)	WIDTH	0 (0)
REMOVE	0 (0)	WAIT	0 (0)
STRING	0 (0)		
{ PRINT	37% (597)		
{ WRITE	30.5% (96)		
{ STOP	11.8% (2)		
{ END	8.1% (5)		

ORIGINAL PAGE IS
OF POOR QUALITY

N78-11146

CRISPFLOW, A Structured Program Design Flowcharter

R. C. Tausworthe and K. C. Landon

DSN Data Systems Section

This article describes the design, prototype implementation, and use of an automated tool, called CRISPFLOW, for generating design flowcharts of structured programs. The user of this tool provides a file of English-like textual statements which directs the drawing of the flowcharts. Layout, symbol sizing, location of text, scaling, and other routine tasks are performed automatically.

I. Introduction

CRISPFLOW is a software design aid which automatically generates ANSI-standard (and DSN Standard) flowcharts in response to directives in a structured program design language given to a computer. The name derives from Control-Restricted Instructions for Structured Program FLOWcharting. Because CRISPFLOW creates flowcharts at design time (from a program design language), rather than at implementation time (from a computer language), it serves as a tool which encourages top-down development of software, relieving the designer of the hand-drawn flowchart burden. Flowcharts are easily available and economically produced, edited, and redrawn during the design phase of a software development. The charts produced are graphic representations of the specifications for coding, not after-the-fact documentation of what was coded.

The user of CRISPFLOW creates a file of statements grouped into modules; each such statement represents one of the symbols on a flowchart to be drawn. The syntax of CRISPFLOW statements is taken from a subset of CRISP (Ref. 1), except that English phrases can be used within the syntax, as well as computer-language constructs, since no code

is being generated from CRISPFLOW statements. Each module flowcharted occupies one standard page.

CRISPFLOW produces flowcharts for a canonic subset of structures most commonly used for structured programming. Less common structures, which include the extended IF form, the indexed looping structure, paranormal exits, and other components of CRISP described in Ref. 1 are not included in the CRISPFLOW processor as of this report.

The CRISPFLOW processor reads a module from a user file and converts this to an internal representation which describes the flowchart symbols to be drawn, the sizes of these symbols, where these symbols are to appear on a page, and what text is to be printed in and around each symbol. The user needs only to prepare a structured source program as described in the next Section of this report. All topological considerations — layout, symbol sizes, text locating, scaling, etc. — are performed automatically by the CRISPFLOW processor.

The CRISPFLOW processor, as of this report, operates as an MBASICTM program on the Univac 1108 which computes the placement of the flowchart symbols on the page and generates the points for plotting. When this is complete, a

FORTRAN subprogram draws the chart on either CALCOMP or COM plotter, the latter on microfilm or microfiche.

II. The CRISPFLOW Language

As described in Reference 1, the CRISP concept is one whereby a control restrictive syntax is superimposed on some base language of the user's choice. Whenever the base language is an abbreviated form of technical English, the constructs form what is known as a procedure (or program) design language. This CRISPFLOW prototype recognizes certain of the CRISP constructs and generates structured flowcharts in response. These constructs hereafter will be referred to as the CRISPFLOW language.

This report will not detail this language, but, will rather illustrate its capabilities. Further information on the CRISPFLOW language and processor is available from the authors.

CRISPFLOW permits the flowcharting of PROGRAM, PROCEDURE(TO), and SUBROUTINE modules containing IFTHENELSE, CASE, LOOP-REPEAT-IF, and LOOP-WHILE-REPEAT flowchart structures. Each CRISPFLOW language statement begins with a keyword which signifies the flowchart symbol to be drawn. Each CRISPFLOW statement results in a specific flowchart response.

A. Base Language Statements

Text which does not begin with one of the recognized CRISPFLOW keywords is transferred directly onto the chart, inserted into a standard process flowchart symbol (a rectangle). For example, the text

BASE LANGUAGE STATEMENT

in the CRISPFLOW input file creates the flowchart image shown in Figure 1a because the word "BASE" is not a recognized keyword.

CRISPFLOW breaks the text string at spaces as needed to fit, and centers the resulting substrings within the rectangle. If the substring cannot fit inside the default-size rectangle, CRISPFLOW selects a rectangle sufficiently large to contain all of the text and then centers the text inside it.

B. CRISPFLOW Comment Annotation

Comments in the CRISPFLOW source images are denoted by the delimiters "<*" and ">", which enclose the comment. Comments may appear in any part of any statement. When a comment appears in a statement, CRISPFLOW usually

discards it, and the comment does not appear on the flowchart. For example, the input statement

BASE <*COMMENT*>LANGUAGE STATEMENT

again results in the same flowchart image as in Figure 1a.

The comment text is not always discarded, however. In particular, the comment field in PROGRAM, TO, PROCEDURE, SUBROUTINE, DO, CALL, and CALLX statements *do* appear on the flowchart in a manner described later.

C. Box Numbers and Cross-Reference Annotation

Statements in CRISPFLOW have optional additional fields available for annotating the ANSI standard box number and cross-reference notations (see Ref. 2) on each of the flowchart symbols. For example, the form

.5/X1 BASE LANGUAGE STATEMENT

results in the flowchart response shown in Figure 1b. The text string between "." and "/" is placed at the upper right of the flowchart symbol, while the text string following "/" up to the first space is placed at the upper left of the flowchart symbol; the entire text message beyond (if not beginning with a keyword) is placed inside the box as described previously.

D. Source Cosmetics

Certain optional characters preceding the CRISP keyword or base language statement are considered "cosmetic", and are ignored by CRISPFLOW. These characters are: "!", ":", "-", ">", "_", and ".", when "." is not in column 1. Cosmetic characters may be used to enhance readability of the input file and such usage will be illustrated in subsequent examples. For the present, observe that

.5/X1 !: : . . . BASE LANGUAGE STATEMENT

results in the same flowchart image as shown in Figure 1b, previously described.

E. Module Delimiters

A *module*, as used here, is a set of statements that CRISPFLOW flowcharts as a single page. Each module flowcharted must begin and end with one of the following pairs of delimiter keywords:

PROGRAM PROCEDURE TO SUBROUTINE
ENDPROGRAM ENDPROCEDURE ENDTO ENDSUBROUTINE

ORIGINAL PAGE IS
OF POOR QUALITY

A sample short module and its flowchart are shown in Figure 2. In the upper right-hand corner, CRISPFLOW has entered the text following MOD#, then, on the next line, the text field following the keyword and up to, but not including "MOD#"; next, the comment (date) text, and finally; "PAGE OF ". Any of the first three of these strings may be absent. The fourth always appears. CRISPFLOW also supplies a signature box for the designer, coder, and auditor; below this appears the date the flowchart was drawn.

CRISPFLOW draws the same entry symbol in response to the module-header keywords PROGRAM, PROCEDURE, TO, and SUBROUTINE, and the exit symbol in response to ENDPROGRAM, ENDPROCEDURE, ENDT0, and ENDSUBROUTINE. Text following MOD# on the module header is drawn above and left of the entry symbol, and text following the keyword is printed within the entry symbol. CRISPFLOW prints "RETURN" in the exit symbol of SUBROUTINE modules.

F. Striped Symbols

The ANSI standard convention (Ref. 2) for indicating that a more detailed description or representation of a function is to be found elsewhere is by means of "striping" the chart symbol horizontally or vertically. CRISPFLOW uses statements having the keywords DO, CALL, and CALLX for this purpose. The text following the keyword appears at the top of each such symbol, and the comment field, if there is one, appears in the lower part of the symbol. Multiple comment fields are merged together. For example, the statements

```
.8      DO PARSE <*CONVERT MODULE TO A TREE &
        STRUCTURE.*>
.4/S2   CALL INVMTX<*INVERT A MATRIX.*>
.12/X3  CALLX PRTLIN<*PRINT A LINE*>&
        <*ON THE TERMINAL.*>
```

cause CRISPFLOW to draw Figures 3a, 3b, and 3c, respectively. CRISPFLOW stripes DO and CALL statements horizontally, and CALLX statements vertically. Note the use of the ampersand at the end of a line to continue the statement.

The CRISP DO statement is used to reference a procedure flowchart by the same name elsewhere in the current documentation set; the CALL statement represents a linkage to a subroutine documented elsewhere in the current documentation set. The CALLX form denotes reference to a procedure or subroutine *not* in the current documentation set, but documented at the place cited by the cross-reference identifier (X3 above). CALLX statements, for example, are commonly used to denote invocations of library subroutines.

G. The IFTHENELSE Structure

Binary decisions are plotted using the IFTHENELSE structure, which requires a block of statements of the form illustrated by

```
.4      IF (DELTA<0)
.5      :   Z=X+Y
        :->(ELSE)
.6      :   Z=X-Y
        :..ENDIF
```

Note the use of cosmetics to enhance readability of the CRISPFLOW source. The result is shown in Figure 4.

On the flowchart, CRISPFLOW places the condition under test in the diamond followed by a question mark; "THEN" and "ELSE" clauses are drawn to the left and right of the diamond, respectively, and are labelled "T" and "F". The user may have as many statements as desired in either the THEN or the ELSE clauses. CRISPFLOW recognizes either ELSE or (ELSE) as the beginning of the ELSE-clause. The ENDIF statement is required to terminate the IF-block.

H. Looping Structures

CRISPFLOW provides two of the more common iterative structures: both are bounded within LOOP...REPEAT statements in the source module, as illustrated in the following two examples, which produce Figures 5 and 6, respectively.

```
PROCEDURE: NEWTON'S METHOD <* 11 Aug 77*> &
        MOD# 1
```

```
<* THIS ITERATIVE PROCEDURE COMPUTES THE
<* ROOT OF SOME UNSPECIFIED FUNCTION OF X
<* STARTING WITH AN APPROXIMATION OF THE
<* ROOT. A LIMIT IS PLACED ON THE NUMBER
<* OF ITERATIONS TO FORCE LOOP TERMINA-
<* TION IF CONVERGENCE FAILS OR IS TOO
<* SLOW.
```

```
.1      X=BEST GUESS OF ROOT, LIMIT=100, &
        T=THRESHOLD
.2      LOOP
.3      ! DELTA=FUNCTION(X)/DERIVATIVE(X)
.4      ! X=X-DELTA, LIMIT=LIMIT-1, &
        ! ENOUGH=(ABS(DELTA)>T AND LIMIT>0)
.5      ! _REPEAT IF (NOT ENOUGH)
        ENDPROCEDURE
```

```

SUBROUTINE: SEARCH <*29 AUG 76*> MOD# S29
  <* ARGUMENTS ARE VAL: INTEGER,
  <* FOUND: BOOLEAN.
  <* SEARCH THROUGH INTEGER ARRAY FOR
  <* VAL, AND SET
  <* FOUND ACCORDINGLY.
.1 SELECT INITIAL PORTION TO BE SEARCHED
.2 LOOP WHILE (PORTION SIZE>1)
.3 ! REDUCE PORTION TO BE SEARCHED
  !_REPEAT
.4 EXAMINE CHOSEN PART, AND SET FOUND
  ENDSUBROUTINE

```

I. The CASE Structure

When branching in a program may take more than two alternate paths, the CASE structure may be used. The format of the CASE structure is illustrated in the next example, which is a module that invokes one of three different subroutines depending on the discriminant of a quadratic equation ($b^2 - 4ac$):

```

TO FINDROOT <*11 AUG 77*> MOD# 1.5.7.3
.1 DISCRIMINANT =B**2-4*A*C
.2 CASE (DISCRIMINANT)
  :->(=0)
.3 : CALL ONEROOT <*PRINT SINGLE VALUE*>
  :->( >0)
.4/T7: CALL REALROOTS <*PRINT 2 REAL &
  : VALUES*>
  :->( <0)
.5/N3: CALL COMPROOTS <*PRINT 2 COMPLEX &
  : VALUES*>
  :..ENDCASES
  ENDT0

```

The decision condition follows the keyword CASE in parentheses, parenthesized case labels signal the beginnings of clauses, and the keyword ENDCASES terminates the structure. The flowchart draws the decision-text string in a diamond, draws the appropriate number of outcome flowpaths below the decision symbol, and labels the flowpaths with the various outcomes. Figure 7 shows the flowcharted result.

J. Nested Structures

CRISPFLOW automatically configures substructures to be flowcharted within structures. The following SAMPLE procedure illustrates the nesting of structures in both source and flowchart (Figure 8) forms:

```

PROCEDURE: SAMPLE<*18 SEPT 75*> MOD# 1.3.5
  <* THIS SAMPLE MODULE DEMONSTRATES THE
  <* CRISPFLOW SYNTAX WITH A HYPOTHETICAL
  <* MESSAGE TRANSMISSION SYSTEM. STATE-
  <* MENTS DENOTED BY ST1 THROUGH ST4 ARE
  <* UNSPECIFIED HERE, AND THE READ ROUTINE
  <* IS EXTERNAL TO THE SET OF DOCUMENTA-
  <* TION FOR WHICH THE FLOWCHART IS BEING
  <* PRODUCED.
.1 IF (UNALLOCATED)
.2 CASE (MODE)
  (1)
.3 ST1
  (2)
.4 ST2
.5 ST3
  (3)
.6 ST4
  ENDCASES
.7/S1 CALL OPEN (MODE) <*CHANNEL IS MODE &
  NUMBER.*>
  (ELSE)
.8 LOOP WHILE (AVAILABLE)
.9/XS2 CALLX READ (CHR) <*READ CHARACTER*>
.10/S3 CALL WRITE(CHR) <*WRITE CHARACTER*>
  REPEAT
  ENDIF
.11 DO CLOSE <*MESSAGE SENT*>
.12 DO RELEASE <*DISCONNECT CHANNEL*>
  ENDPROCEDURE

```

Note that comments after the procedure declaration do not transfer to the chart. In this case, the flowchart is less explanatory than the source module, and should therefore be accompanied by supplementary explanatory narrative.

III. The CRISPFLOW Processor Design

The top-level CRISPFLOW control structure repeats for each module in the source file the following three steps:

```

PARSE "the module into a tree"
LAYOUT "the nodes of the tree on a plotter page"
DRAW "the boxes and flowlines on the chart"

```

The parser scans each line of the source input stream to pick up "tokens" (Refs. 3, 4) and uses a bottom-up parse, based on the first "keyword" token of each source line, to construct an internal representation of the module's parse tree. The SAMPLE of Figure 8 generates the tree structure shown in Figure 9, for example. The phrase structure of the CRISP grammar is designed so that there exists a close correspondence between nodes in this tree and "super-boxes"

(discussed below) on the chart. Each node in the tree corresponds to a string of terminals in the grammar, and ultimately is also made to reference all the information needed to locate, draw, fill, and connect the box it represents.

The box-arrangement algorithm in the CRISPFLOW system, LAYOUT, is responsible for translating between the procedural and flowchart media in a manner which yields both elegant and homomorphic mappings between source statements and flowchart structures. The importance of the former criterion cannot be over-emphasized, for the utility of the system to potential users is greatly diminished if the boxes and flowlines on a chart are not arranged in a straightforward, well-formatted and easily-readable fashion. At the same time, however, economic and human engineering considerations dictate that such machine constraints as processing time and memory requirements not be ignored.

The algorithm herein adopted for LAYOUT is a compromise, therefore, between a "dynamic programming" approach and one which forestalls repetitive, brute-force search schemes to allow rapid chart layout with only modest memory demands. The "super-box" approach yields very well-arranged charts without attempting to pack boxes as densely (or as "cleverly"), as might be expected.

The layout routine scans the parser-generated tree in a postorder, or bottom-up, traverse (Ref. 5), and for each node encountered, forms a transparent "super-box" which encloses the superboxes of all its (nested) subtrees. The size of each superbox is determined by the type of node (i.e., statement) to which it corresponds. The postorder walk permits the grammatical structure of the tree to be utilized as a "reduction system", or means to layout the chart in a single, comprehensive bottom-up pass through the tree. As the nodes at each level in the tree are scanned and "reduced" to super-boxes, they are used, in turn, to construct the super-boxes of the nodes at the next higher level. This procedure continues until the top-level statements of a module are combined into the super-box for the entire chart.

The DRAWing routine is activated after the layout tree traversal, and it first centers the super-box of the module on the plotter page and scales it down to fit, if necessary. It then performs a preorder, or top-down, tree scan, during which it establishes coordinates, draws the boxes, fills them with text, and connects them with flowlines, arrows, and collecting nodes.

A. Super-box Definitions

The figures of this section illustrate the super-box formats necessary to represent the flowchart symbols in this design, along with their structured conglomerates. Figure 10 shows

the atomic boxes which are actually drawn in and annotated. Figure 11 illustrates how "then" and "else" sub-superboxes stand in relation to the superbox of the IFTHENELSE structure. Superboxes for LOOP, CASE and other structures are similarly defined.

In Figures 10 and 11, DH is a unit of horizontal spacing (normally about 3 mm) and DV is that for vertical spacing (normally about 5 mm). These allow for box separation, arrow heads, etc. SBH and SBW (possibly with distinguishing subscripts) are acronyms for Super-Box Height and Width, respectively, and are shown in Figure 11 with the equations used to compute them from the known dimensions of their constituent sub-super-boxes (either atomic or conglomerate). The coordinates of all chart symbols within a super-box are given relative to the axes of that super-box; the origin of a super-box is located at its top dead center (i.e., at the point bisecting its top edge).

Super-box outlines never intersect, and it is this property which makes LAYOUT as straightforward as it is. To be more precise, one super-box may enclose or be enclosed by others, or two super-boxes may share a common edge, but no two super-boxes may ever overlap. No super-box spacing is needed since the super-box which directly encloses a chart symbol includes sufficient space around its perimeter to give an eye-pleasing format. Thus, the layout algorithm simply "pushes" together the super boxes it "picks up" in a postorder walk according to the "blueprints" such as that in Figure 11. Once the conglomeration of boxes comprises a complete module, the chart origin and scale are determined to fit the chart on the page.

B. Data Structures

The two primary structures needed for the above-sketched algorithms are a binary tree (containing the parsed module descriptions) and a pushdown stack (to enable walks of the tree). Each node in the tree has identical structure, with sufficient fields to keep box number, cross-reference, type, size, x-y coordinates, and text information. Two pointer fields are included in addition, so that the module can be mapped onto a standard (leftmost descendant, sibling) binary representation (thus allowing a constant number of fields for each node packet), as shown in Figure 9. An entry in the stack consists of a pointer to a node in the tree.

C. The Tree

The attributes of each flowchart symbol are organized into the structure TREE. It is declared as a vector of records, and each record, representing one node, has the fields described below. The nodes are linked together into the tree structure by two pointer fields in each node, denoted TREE.BROTHER

and TREE.SON. These pointers permit flowchart representation via a binary mapping of the tree nesting levels.

The fields (not all applicable to each node class) within each node of the TREE are:

<u>NAME</u>	<u>DESCRIPTION</u>
CLASS	the type of node (e.g., IF, CASE, . . .).
SON, BROTHER	Pointers
BOXNUM, XREF	The box number and cross-reference identification strings for the chart symbol.
TEXT, COMMENT	The texts of the CRISP or base language statement and any comment on the same line.
SBW, SBH	The width and height of the super-box associated with this node.
W	The width of the chart symbol to which this node corresponds.
X,Y	The chart coordinates of this superbox.

IV. Operational Characteristics

The Univac 1108 current implementation of CRISPFLOW is coded in the MBASIC language (Refs. 6 and 7), and therefore the source files to direct plotting must be MBASIC processor compatible (standard data file format). Such files may be created using the MBASIC processor or the JPL text editor.

CALCOMP flowcharts are normal (21.6 × 28 cm or 8-1/2 × 11 inch) size but lack the drawing quality of the COM plotter charts. The COM plotter produces high-quality charts on either 35mm film or 105mm microfiche; 35mm output

also is accomplished by a 23-cm (9-inch) photostat furnished automatically. Full-page reproductions are available then from the Central Reproductions Facility.

The cost of a flowchart, such as SAMPLE in Figure 8, is about \$4-\$6 at prime-shift computer rates; during weekends, the rate drops to about \$.40-\$.60 per chart. When the MBASIC batch compiler becomes available, however, these costs are expected to drop by a factor of 5 to 10. Recoding the CRISPFLOW processor in a non-interpretive language would also provide approximately the same operational cost savings.

V. Conclusion

The use of CRISPFLOW as a design tool brings the power of a computer-based system to relieve many of the objections often levied against flowcharting. Editing and updates are as easily accomplished on charts as they are on any other data files. The charts are accurately and quickly drawn, and displayed in excellent drafting quality. Further, due to the high level of the CRISPFLOW language and the 1-1 correspondence between statements and chart symbols, programmers themselves quickly learn to use the source versions of modules, and order the drawing of flowcharts only when they become satisfied with their designs and desire documentation of a more graphic form for others.

The processor described here is only a prototype; undoubtedly it will evolve and improve as users interact with it and suggest improvements. Usage will probably increase when an MBASIC compiler becomes available to reduce costs. Even in its current state, however, it is demonstrating the benefits of computer-based graphic documentation of computer programs.

References

1. Tausworthe, Robert C., *Standardized Development of Computer Software; Part 1, Methods*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1977, Chapter 7.
2. "American National Standard Flowchart Symbols and Their Usage in Information Processing", American National Standards Institute, Inc., ANSI X3.5-1970, Sept. 1, 1970.
3. Donovan, John J., *Systems Programming*, McGraw-Hill, Inc., 1972, pp. 227-240.
4. Aho, A. V., and Ullman, J. D., *The Theory of Parsing, Compiling, and Translation*, Vols. I and II, Prentice-Hall, Inc., Englewood Cliffs, N.J.
5. Knuth, D., *Fundamental Algorithms*, Vol. 1, Addison-Wesley Pub. Co., Reading, Mass., 1969.
6. "MBASIC, Vol. 1, Fundamentals," Jet Propulsion Laboratory, Pasadena, Ca., March 1973 (JPL internal document).
7. "MBASIC, Vol. 2, Appendices", Jet Propulsion Laboratory, Pasadena, Ca., Oct. 1973 (JPL internal document).

ORIGINAL PAGE IS
OF POOR QUALITY

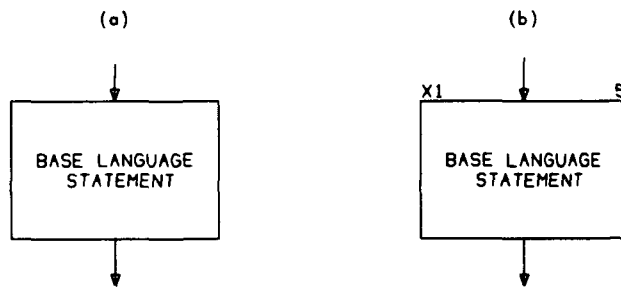


Fig. 1. Response to base language statements: (a) without box number or cross-reference annotations; (b) with annotations

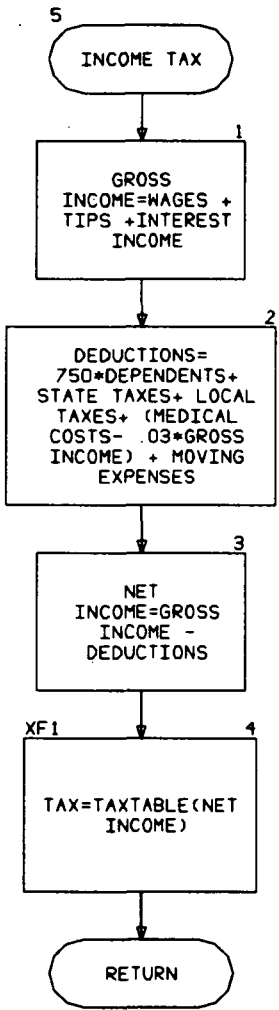
ORIGINAL PAGE IS
OF POOR QUALITY

```

SUBROUTINE INCOME TAX <*11 AUG 77*>          MOD# 5
.1      GROSS INCOME=WAGES + TIPS +INTEREST INCOME
.2      DEDUCTIONS= 750*DEPENDENTS+ STATE TAXES+ LOCAL TAXES+&
        (MEDICAL COSTS- .03*GROSS INCOME) + MOVING EXPENSES
.3      NET INCOME=GROSS INCOME - DEDUCTIONS
.4/XF1  TAX=TAXTABLE(NET INCOME)
        ENDSUBROUTINE

```

5
INCOME TAX
11 AUG 77
PAGE OF



D:	
C:	
A:	

16 AUG 77

Fig. 2. A sample flowchart showing entry and exit symbols, plus other annotations supplied by CRISPFLOW for convenience in finding the chart among a set of charts on other documentation and for design control signatures

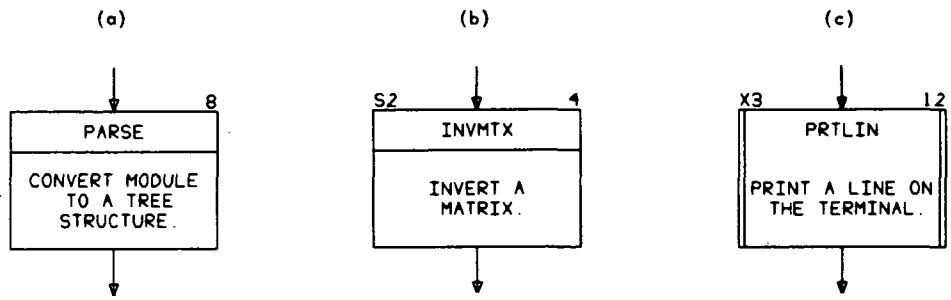


Fig. 3. Striped symbols used for displaying procedure, internal subroutine, and external subroutine documentation

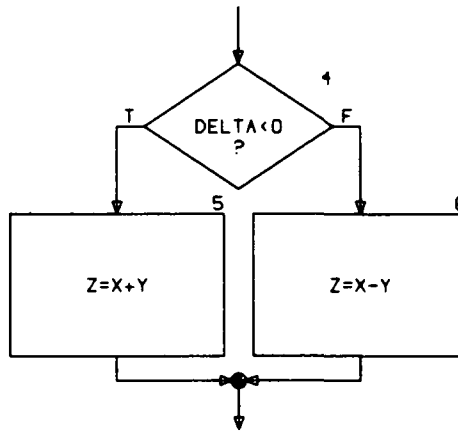
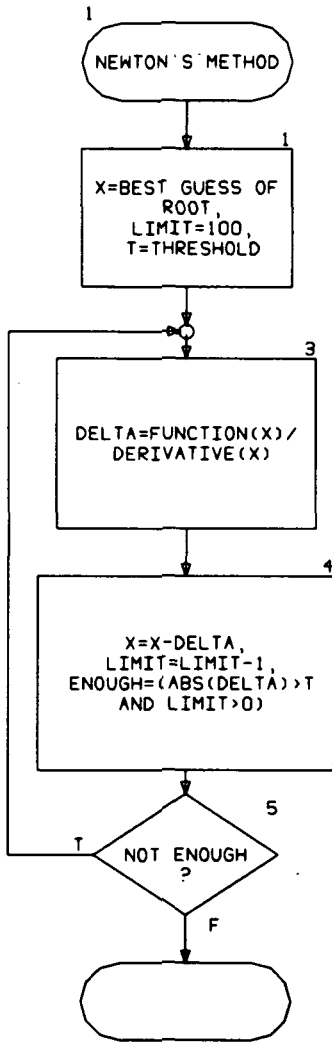


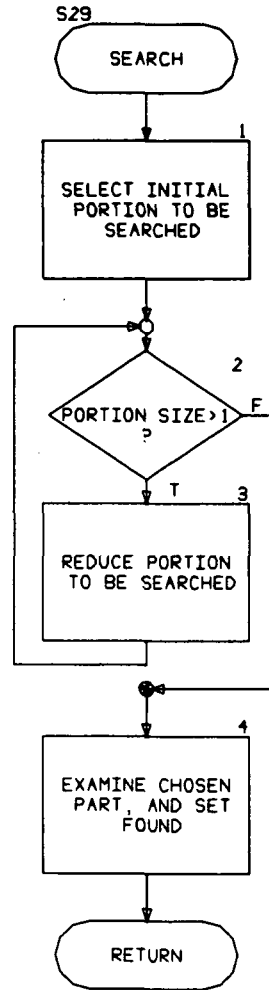
Fig. 4. Flowcharted response to an IFTHENELSE block of statements in CRISPFLOW

ORIGINAL PAGE IS
OF POOR QUALITY



D:	
C:	
A:	

16 AUG 77



D:	
C:	
A:	

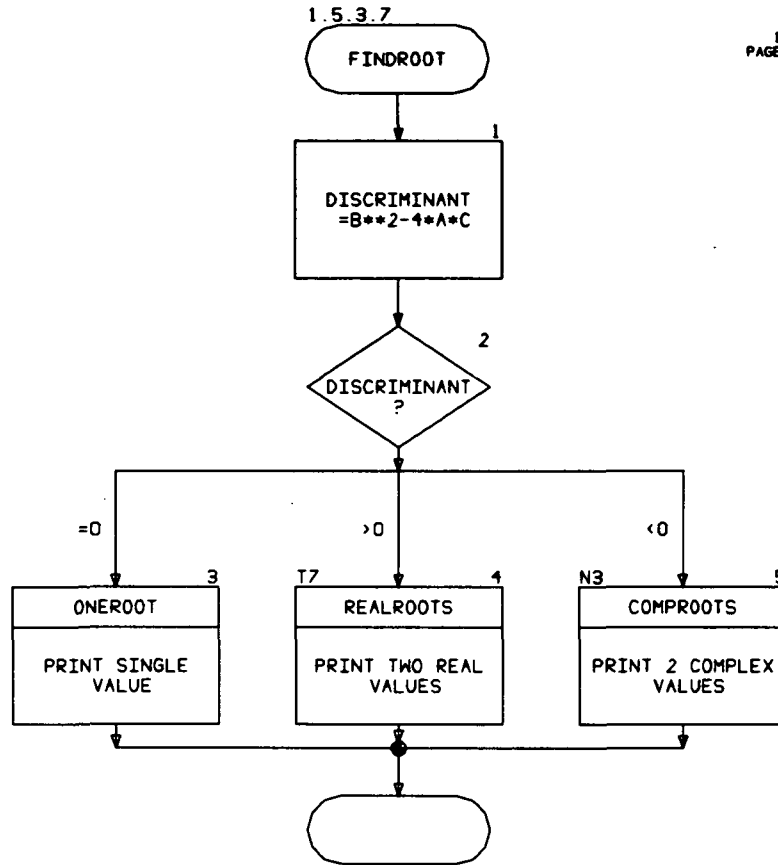
16 AUG 77

Fig. 5. Procedure module demonstrating LOOP . . REPEAT IF block of statements

Fig. 6. Flowchart of a subroutine module illustrating a LOOP WHILE . . REPEAT block of statements

1.5.3.7

1.5.3.7
FINDROOT
11 AUG 77
PAGE OF

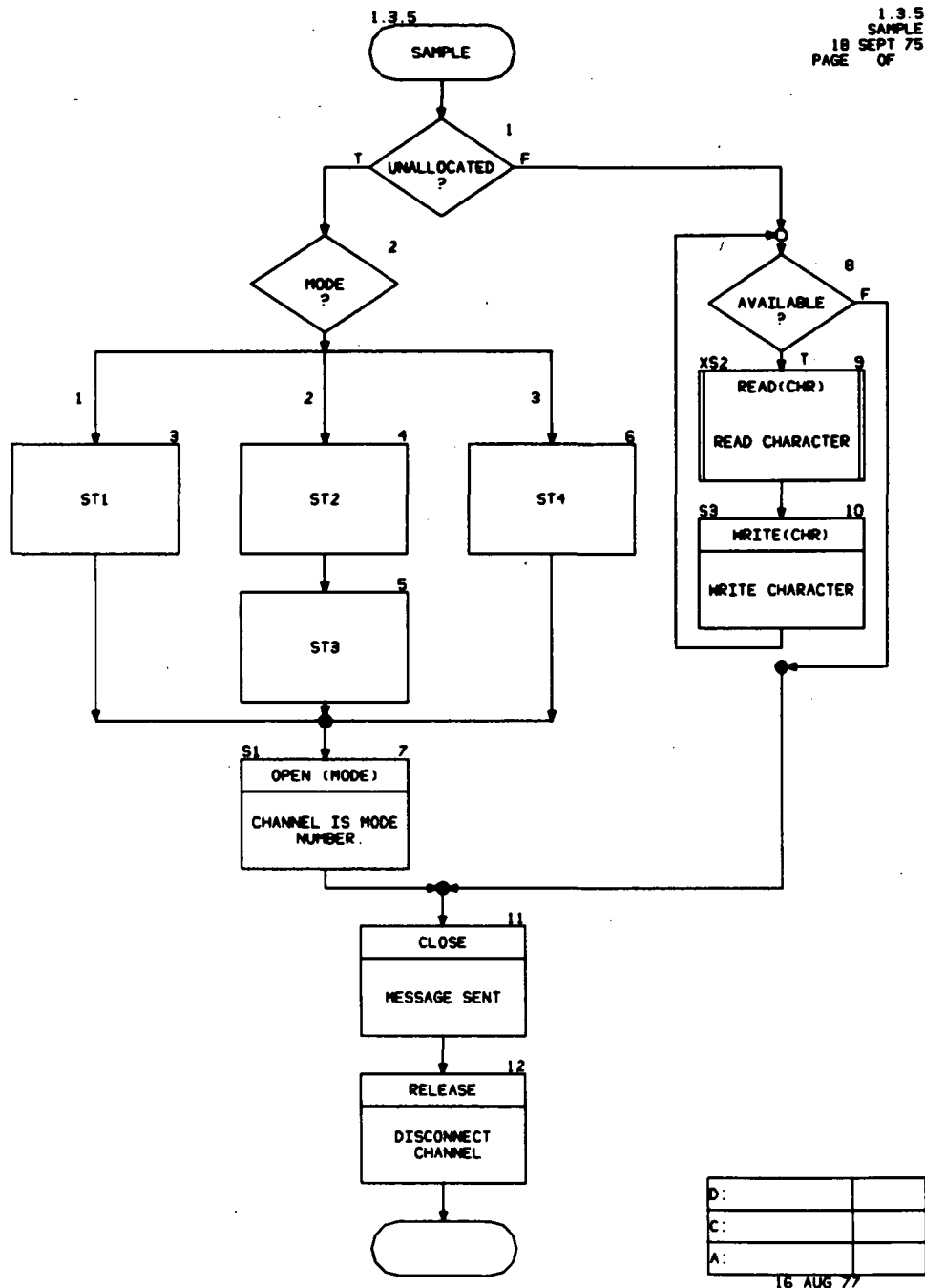


D:	
C:	
A:	

16 AUG 77

Fig. 7. Flowcharted result of a TO module containing a CASE structure

ORIGINAL PAGE IS
OF POOR QUALITY



D:	
C:	
A:	

16 AUG 77

Fig. 8. SAMPLE procedure illustrating nesting of structures

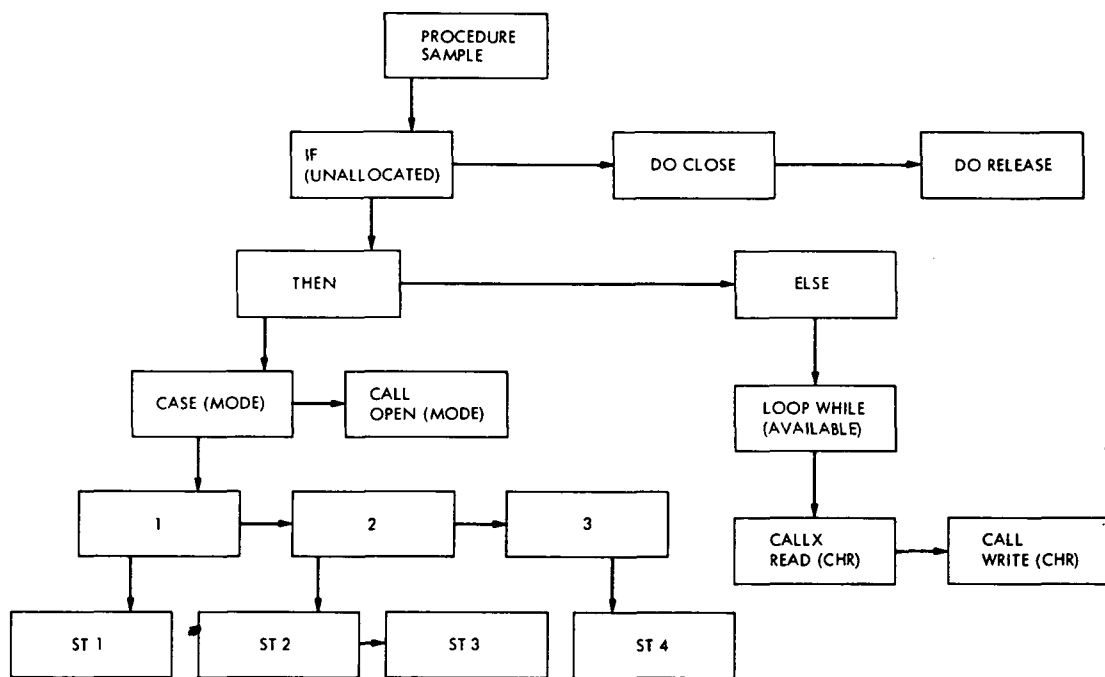


Fig. 9. The Binary Parse Tree of Procedure "SAMPLE". SON pointers exit the bottoms of boxes, BROTHER pointers leave the sides

ORIGINAL PAGE IS
OF POOR QUALITY

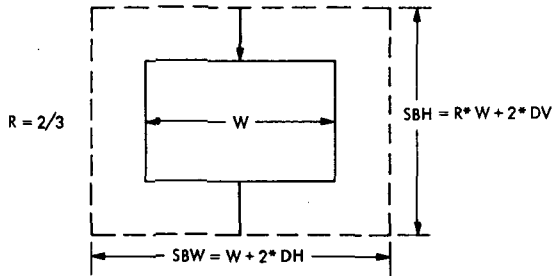


Fig. 10a. The standard rectangle for CALL, CALLX, DO, and target statements

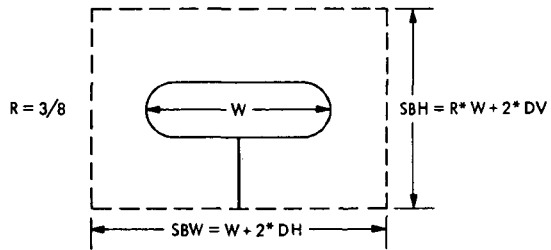


Fig. 10c. The oval for module entrances

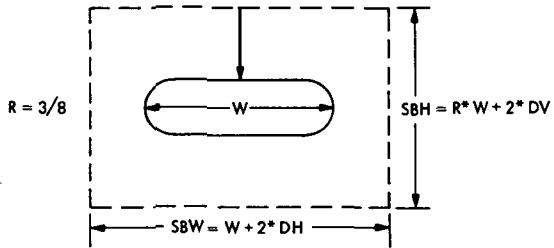


Fig. 10b. The oval for module exits

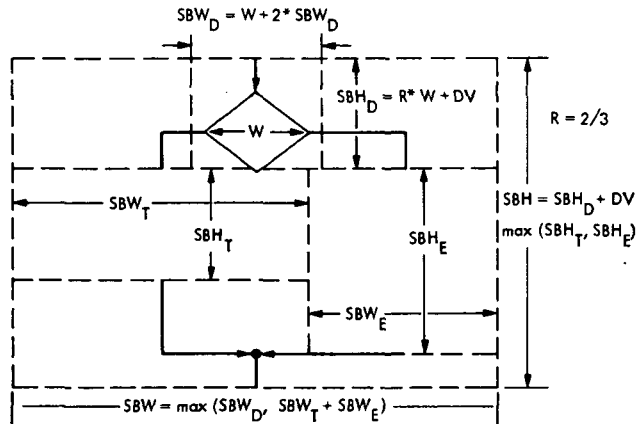


Fig. 11. The super-box of an IF structure with Decision super-box and Then and Else clause statement list super-boxes as constituent parts

D15
N78-11147

The CTA 21 Radio Science Subsystem — Non-Real-Time Bandwidth Reduction of Wideband Radio Science Data

A. L. Berman and J. I. Molinder
TDA Engineering Office

The concept of a centralized facility at JPL to reduce the bandwidth of wideband digital recordings originated with the Pioneer Venus Differential Long Baseline Interferometry (DLBI) Experiment. This article presents a functional description of the resulting facility — the CTA 21 Radio Science Subsystem (CRS), located within JPL at the Compatibility Test Area (CTA-21). Particularly emphasized is a mathematical derivation and description of the digital filter process which comprises the core of the CRS.

I. Introduction

In December 1978, the Pioneer Venus Multi-probe Mission Spacecraft will encounter the planet Venus, and at that time the Differential Long Baseline Interferometry Experiment will attempt to measure the wind velocities in the atmosphere of Venus as four probes descend through the atmosphere. The fundamentals of the experiment are described in Reference 1. The end product at each of the involved tracking stations will be a wide bandwidth digital recording (magnetic tape) covering the entire bandwidth of interest, which includes five spacecraft signals (the spacecraft bus and the four descending probes). Plans to mid-1976 had assumed that the deliverable to the experimenter, located at the Massachusetts Institute of Technology (MIT), would be the actual wideband recordings produced at these tracking stations. This plan involved the Project procuring for the experimenter recorders capable of playing back these wideband recordings, and required that the experimenter implement and operate a system to convert the wideband recordings into a usable computer-compatible form. Estimates by MIT of the resources required to implement and operate this conversion process turned out to be a very signifi-

cant portion of the total processing costs at MIT for the experiment. This problem resulted in the evolution of the concept of a centralized facility at JPL ("CTA 21 Radio Science Subsystem") to reduce the bandwidth of the wideband digital recordings and supply the experimenter with the bandwidth reduced data on computer-compatible magnetic tape. In the following sections, the functional characteristics and operation of the CTA 21 Radio Science Subsystem will be described.

II. Functional Description of the CTA 21 Radio Science Subsystem

A. Definition

The CTA 21 Radio Science Subsystem, a dedicated and integral element of the DSN Radio Science System, performs the following functions:

- (1) Reduces the digitally recorded output of the wideband open-loop receiver to a narrow bandwidth centered about a predicted frequency profile.

- (2) Interprets the timing information included with the digitally recorded wideband open-loop receiver data, and appropriately time-tags the bandwidth-reduced output data.
- (3) Provides the time-tagged bandwidth-reduced output data and the corresponding predicted frequency profile on computer-compatible magnetic tape for distribution to the appropriate Radio Science Experimenters.

CTA 21 Radio Science Subsystem functions and interfaces are presented in Figure 1.

B. Key Characteristics

The key characteristics of the CTA 21 Radio Science Subsystem are as follows:

- (1) Hardware and software are compatible with the DSN Mark III configuration.
- (2) Processing is fully digital and automated:
 - (a) A posteriori frequency profile obtained from project navigation-DSN prediction system via high-speed data (HSD) interface.
 - (b) Processing proceeds at wideband tape sample rate.
 - (c) Playback of output tape detects signal presence.
- (3) Bandwidth reduction filter is an integrate-and-dump, with dump performed by computer.
 - (a) Output sample rate is adjustable from 1/200 to 1/4000 of input sample rate.
 - (b) Maximum tracking rates accommodated by variation of output sample rate.
- (4) Timing information recorded with the wideband input data is interpreted and used to time-tag the output data.
- (5) Output tape quantity is dependent upon user-selected sample rate.
- (6) A computer-compatible tape, which includes the bandwidth-reduced data and the corresponding a posteriori frequency profile, is provided to the Experimenters.

C. Functional Operation

The CTA 21 Radio Science Subsystem (CRS) functions and data flow are presented in Figure 2. A magnetic tape, containing wideband open-loop receiver data which have been recorded by the Digital Recording Assembly of the DSS Radio

Science Subsystem, is received at CTA 21 in nonreal time via Network Information Control (NIC). Additionally, an a posteriori frequency profile is received via high speed data line (HSDL) from the NOCC Support Subsystem, and is stored. The wideband open-loop receiver data is played back via the Digital Recording Assembly within the CRS, and the Bandwidth Reduction Assembly performs the bandwidth reduction process, using an integrate-and-dump digital filter which is preset to the starting sample and the predicted frequency and phase (presetting the filter with a predicted starting phase allows the processing to occur in segments, if necessary). The processing proceeds at the wideband tape sample rate, and results in a bandwidth reduction of 1/200 to 1/4000 of the input bandwidth. The integrate-and-dump routine provides an effective filter of $\sin(N\pi f/2f_0)/\sin(\pi f/2f_0)$ where N is the decimation ratio and f_0 is the input bandwidth. Figure 3 is a functional block diagram of the Bandwidth Reduction Assembly.

The appropriately time-tagged, bandwidth-reduced data and the a posteriori phase information are written on computer-compatible tape for delivery to the Experimenters. As a final step, the output tape is played back for signal detection and validation.

III. Analysis of the Bandwidth Reduction Assembly Signal Processing

A mathematical analysis of the signal processing performed by the bandwidth reduction assembly is given in this section. Effects of noise and quantization are not considered.

A simplified block diagram of the signal processing involved in the bandwidth reduction is shown in Figure 4, where

$$\phi_a(t) = \text{actual (probe) phase}$$

$$\phi_p(t) = \text{predicted (probe) phase}$$

$$\Delta t = (\text{input) data sample interval} = 1/2(2 \ 1/12 \times 10^6) = 0.24 \ \mu\text{sec}$$

The digital filter input sequences $x(n\Delta t)$ and $y(n\Delta t)$ are given by

$$\begin{aligned} x(n\Delta t) &= A \sin \phi_a(n\Delta t) \cos \phi_p(n\Delta t) \\ &= \frac{A}{2} \left\{ \sin [\phi_a(n\Delta t) - \phi_p(n\Delta t)] \right. \\ &\quad \left. + \sin [\phi_a(n\Delta t) + \phi_p(n\Delta t)] \right\} \end{aligned}$$

$$\begin{aligned}
y(n\Delta t) &= A \sin \phi_a(n\Delta t) \sin \phi_p(n\Delta t) \\
&= \frac{A}{2} \left\{ \cos [\phi_a(n\Delta t) - \phi_p(n\Delta t)] \right. \\
&\quad \left. - \cos [\phi_a(n\Delta t) + \phi_p(n\Delta t)] \right\}
\end{aligned}$$

Note that the first term in the final expressions for $x(n\Delta t)$ and $y(n\Delta t)$ above is a sampled signal of frequency $\omega_a(t) = \dot{\phi}_a(t) - \dot{\phi}_p(t)$ while the second term is a sampled signal of frequency $\omega_s(t) = \dot{\phi}_a(t) + \dot{\phi}_p(t)$. If $\phi_p(t)$ is properly computed $\omega_a(t) \ll \omega_s(t)$. The low frequency term is the desired output of the digital filters.

The frequency response of the digital filters can be determined by noting that if

$$x(n\Delta t) = e^{j2\pi f n \Delta t}$$

then

$$u(n\Delta t) = \bar{H}(f) e^{j2\pi f n \Delta t}$$

where $\bar{H}(f)$ is the frequency response of the filter. Computing $u(n\Delta t)$ yields

$$\begin{aligned}
u(n\Delta t) &= \sum_{k=0}^{N-1} e^{j2\pi f(n-k)\Delta t} \\
&= \left(\sum_{k=0}^{N-1} e^{-j2\pi k f \Delta t} \right) e^{j2\pi n f \Delta t}
\end{aligned}$$

and thus

$$\bar{H}(f) = \sum_{k=0}^{N-1} e^{-j2\pi k f \Delta t}$$

As is to be expected, $\bar{H}(f)$ is periodic with period $1/\Delta t$. Summing the series yields

$$\bar{H}(f) = e^{-j\pi(N-1)f\Delta t} \frac{\sin \pi N f \Delta t}{\sin \pi f \Delta t}$$

the magnitude of which is sketched as a function of $f\Delta t$ for $N = 7$ in Figure 5.

Note that in general the first zero of the frequency response occurs at $f\Delta t = 1/N$. If $\phi_p(n\Delta t)$ is properly computed $f_a\Delta t =$

$\omega_a\Delta t/2\pi < 1/2N$. Since the sampling interval $\Delta t = 1/2B$ where B is the input bandwidth (Nyquist rate)

$$f_a < \frac{2B}{2N} = \frac{B}{N}$$

corresponding to a bandwidth reduction factor of at least N . This in turn means that the output samples need only be computed at intervals of $N\Delta t$ (corresponding to an "integrate and dump" filter) resulting in a significant reduction of data. The ratio N of the input rate to the output sample rate is called the "Decimation Ratio." Of course, care must be taken that $f_s\Delta t = \omega_s\Delta t/2\pi$ corresponds to a point of high attenuation in the filter response relative to $f_a\Delta t$.

As an example, the actual bandwidth reduction assembly has the following specifications:

$$\text{Input bandwidth} = 2 \text{ 1/12 MHz} = B = \frac{1}{2\Delta t}$$

$$\Delta t = \frac{1}{2(2 \text{ 1/12} \times 10^6)} = 0.24 \text{ } \mu\text{sec (Nyquist rate)}$$

Decimation Ratio N adjustable from 200 to 4000 (corresponding to the number of samples summed)

Computing the first zero of the filter response for $N = 200$ and 4000 yields output bandwidths of approximately 1 kHz to 20 kHz as shown below:

$N=200$

$$f = \left(\frac{1}{200} \right) \left(\frac{1}{\Delta t} \right) = \frac{1}{200} \times (4 \text{ 1/6} \times 10^6) = 20.83 \text{ kHz}$$

$N=4000$

$$f = \left(\frac{1}{4000} \right) \left(\frac{1}{\Delta t} \right) = \frac{1}{4000} \times 4 \text{ 1/6} \times 10^6 = 1.04 \text{ kHz}$$

Note, as discussed above, that f_a must be less than half the value of the first zero of the filter response if the Nyquist rate for the output signal is to be maintained.

Finally, it is of interest to compare the digital filter with the corresponding analog filter employing the same integration time. It is easily shown that the frequency response of the running integration filter shown in Figure 6 is given by

$$H(f) = e^{-j\pi f T} T \frac{\sin \pi f T}{\pi f T}$$

Note that this filter (or its digital counterpart) need not be strictly realizable since processing is not done in real time. Setting $T = (N - 1) \Delta t$ yields

$$H(f) \cong (N - 1) \Delta t e^{-j\pi(N-1)f\Delta t} \frac{\sin \pi(N-1)f\Delta t}{\pi(N-1)f\Delta t}$$

$$\frac{H(f)}{\Delta t} = e^{-j\pi(N-1)f\Delta t} \frac{\sin \pi(N-1)f\Delta t}{\pi f \Delta t}$$

Concentrating on the region bounded by the first zero of the response $0 \leq f\Delta t \leq 1/N$ and assuming $N \gg 1$ so that $\pi f \Delta t \approx \pi/N \ll 1$ allows use of the approximation $\sin \pi f \Delta t \approx \pi f \Delta t$ and thus over this region

$$\frac{H(f)}{\Delta t} \approx \bar{H}(f)$$

Alternatively note that

$$\bar{H}(f) = \sum_{k=0}^{N-1} e^{-j2\pi k f \Delta t} = F \left\{ \sum_{k=0}^{N-1} \delta(t - k\Delta t) \right\}$$

where $F \{ - \}$ denotes the Fourier transform of the quantity contained within the braces and $\delta(t)$ is the impulse or Dirac delta function.

Defining

$$p(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq (N - 1) \Delta t \\ 0 & \text{otherwise} \end{cases}$$

$H(f)$ can be written in the form

$$\bar{H}(f) = F \left\{ p(t) \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t) \right\}$$

Using the fact that the Fourier transform of a product is the convolution of the Fourier transforms and the relation

$$F \left\{ \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t) \right\} = \frac{1}{\Delta t} \sum_{k=-\infty}^{\infty} \delta \left(f - \frac{k}{\Delta t} \right)$$

yields

$$\begin{aligned} \bar{H}(f) &= (N - 1) \Delta t e^{-j\pi(N-1)f\Delta t} \frac{\sin \pi(N-1)f\Delta t}{\pi(N-1)f\Delta t} \\ &* \frac{1}{\Delta t} \sum_{k=-\infty}^{\infty} \delta \left(f - \frac{k}{\Delta t} \right) \end{aligned}$$

where * indicates convolution. Performing the convolution gives the result

$$\begin{aligned} \bar{H}(f) &= (N - 1) \sum_{k=-\infty}^{\infty} e^{-j\pi(N-1)(f-k/\Delta t)\Delta t} \\ &\times \frac{\sin \pi(N-1) \left(f - \frac{k}{\Delta t} \right) \Delta t}{\pi(N-1) \left(f - \frac{k}{\Delta t} \right) \Delta t} \end{aligned}$$

Thus $H(f)$ is simply the superposition of an infinite series of $\sin x/x$ functions shifted by integral multiples of the sampling frequency $1/\Delta t$.

IV. Planned Implementation Schedule

The planned implementation schedule to meet the Pioneer-Venus DLBI experiment requirements is as follows:

CTA 21 Radio Science Subsystem	
First demonstration	15 November 1977 to 15 January 1978
Transfer to operations	1 April 1978
CTA 21 Radio Science Subsystem Interface with Experimenter	
Output tape in correct format	1 October 1977
Output tape generated from a tone	1 January 1978
Output tape generated from an ALSEP ¹ signal	7 February 1978

¹ Apollo Lunar Surface Experiments Package

Reference

1. Miller, R. B., "Pioneer Venus 1978 Mission Support," in *The Deep Space Network Progress Report 42-31*, pp. 11-14, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1976.

ORIGINAL PAGE IS
OF POOR QUALITY

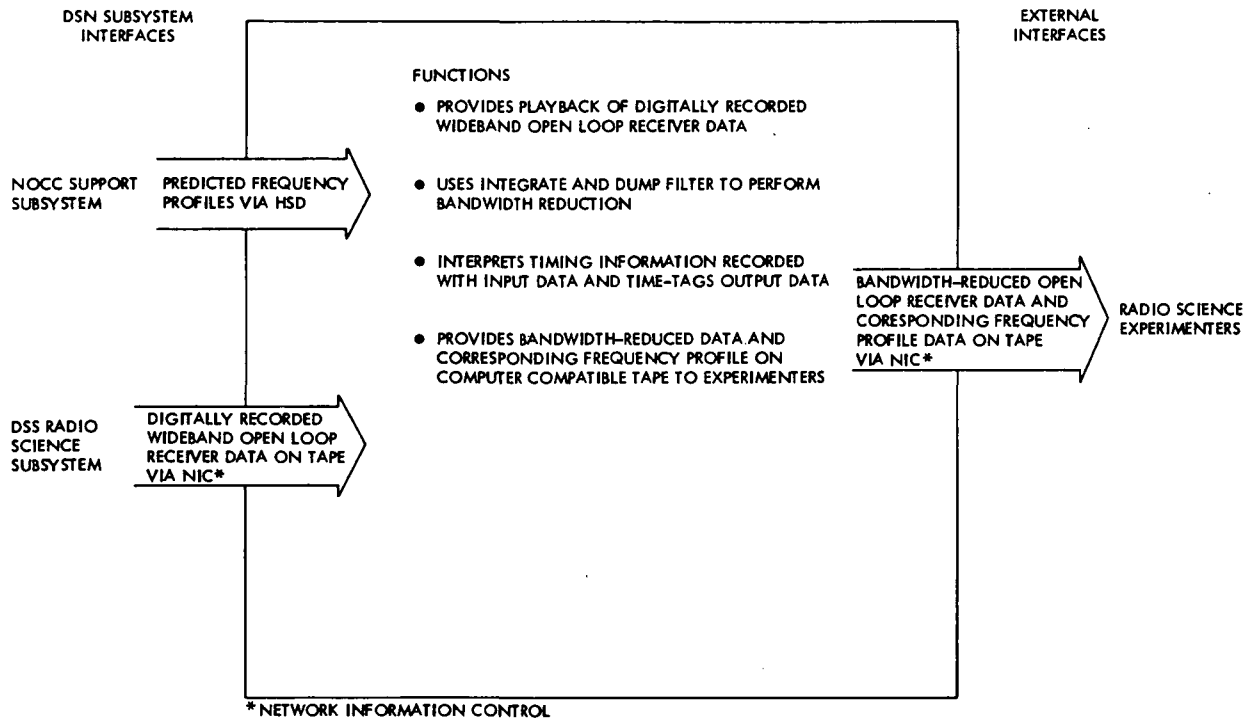


Fig. 1. CTA 21 radio science subsystem functions and interfaces

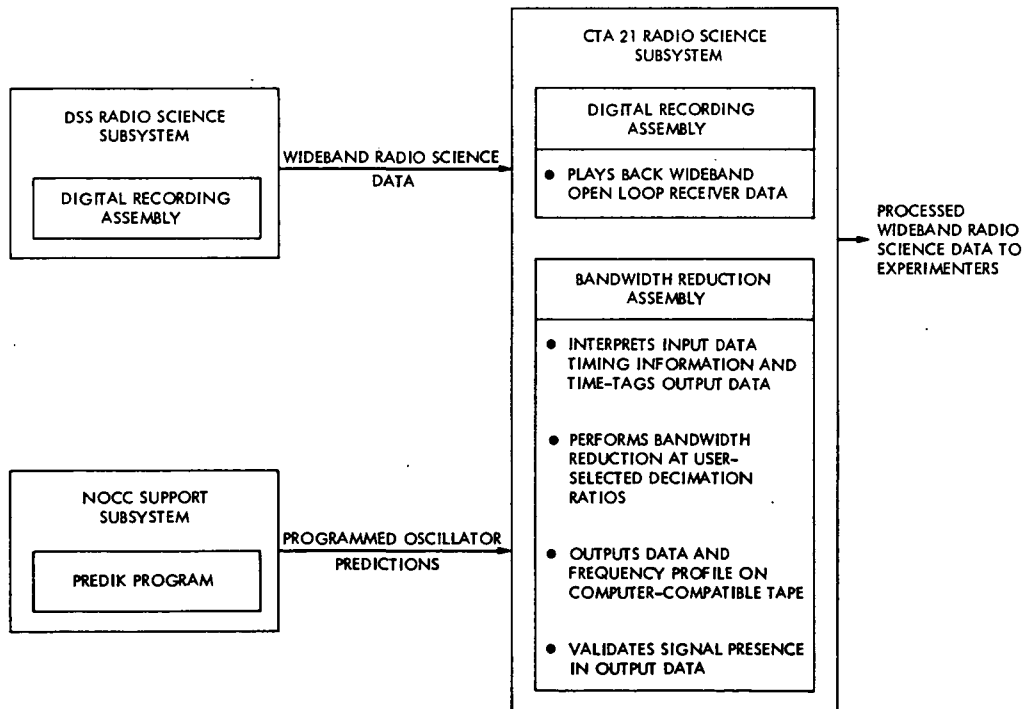


Fig. 2. CTA 21 radio science subsystem functions and data flow

ORIGINAL PAGE IS
OF POOR QUALITY

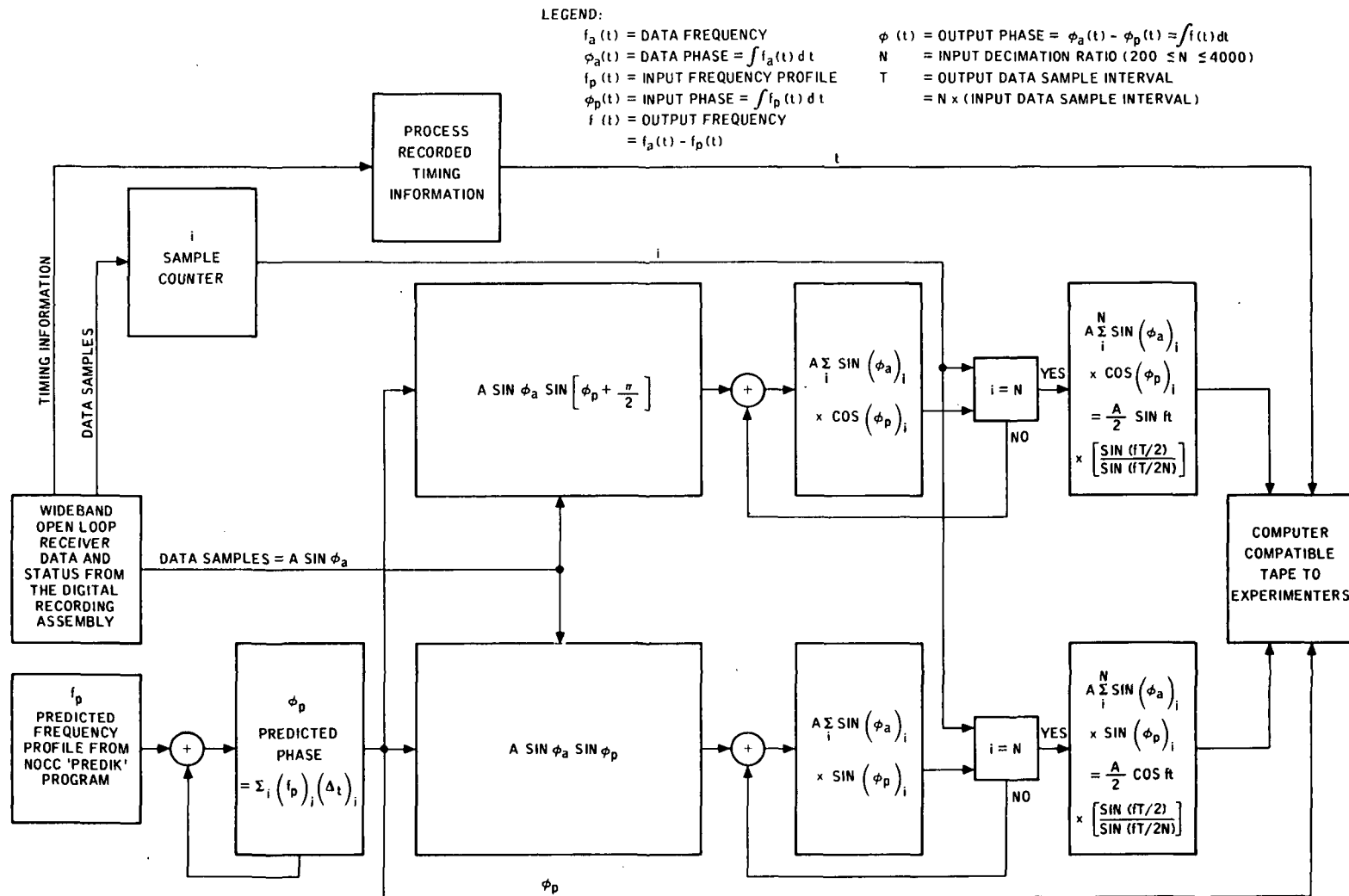


Fig. 3. Bandwidth reduction assembly functional block diagram

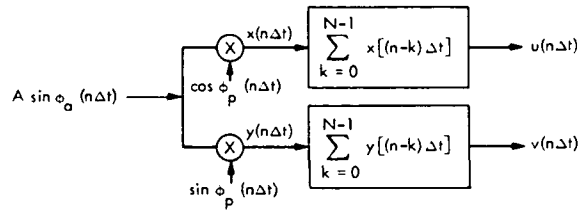


Fig. 4. Block diagram of bandwidth reduction signal processing

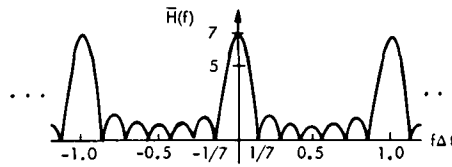


Fig. 5. Frequency response of the digital filter for $N = 7$

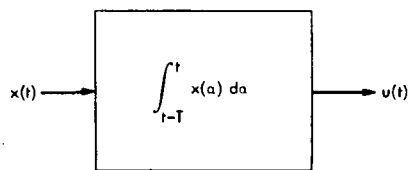


Fig. 6. Running integration filter

N78-11148

Electron Density in the Extended Corona — Two Views

A. L. Berman
TDA Engineering Office

Recent analyses of Viking and Mariner solar conjunction radio metric data have led to two significantly different views of the average radial dependence of electron density in the extended corona ($5r_{\odot} \leq r \leq 1 \text{ AU}$):

$$N_e(r) \propto r^{-2}$$

and

$$N_e(r) \propto r^{-2.3}$$

This article compares the two models and concludes that the "steeper" model ($r^{-2.3}$): (1) is in excellent agreement with other experimental observations of coronal electron density, (2) is consistent with the predicted and observed radial dependence of the solar wind velocity, and (3) augments the case for a turbulence scale that expands linearly with radial distance, when considered in combination with recent observations of the radial dependence of RMS phase fluctuations.

I. Introduction

Recent analyses of Viking and Mariner radio metric data acquired during solar conjunction have led to two significantly differing views of the equatorial coronal electron density function. In a series of recent articles (e.g., Refs. 1-3) Berman, using Viking S-band doppler noise, has shown that the radial dependence of RMS phase (ϕ) in the extended corona¹ is:

$$\phi \propto a^{-1.30}$$

¹Here to be considered as approximately $5r_{\odot} \lesssim r \lesssim 1 \text{ AU}$, where r = radial distance and r_{\odot} = solar radius.

where a = signal path closest approach point, and emphasizes that this result can be explained as the signal path integration of the following nominal electron density (N_e) function:

$$N_e(r) \propto r^{-2.3}$$

where r = heliocentric distance.

More recently, Callahan (Ref. 4), analyzing Viking S-X doppler, has also concluded that phase fluctuations are of the form:

$$\phi \propto a^{-1.3}$$

However, Callahan infers from the above relationship the following electron density fluctuation (n) radial dependence

$$n \propto r^{-1.8}$$

A very common assumption made in coronal investigations is:

$$n/N_e = \epsilon; \epsilon \neq \epsilon(r)$$

Thus implying (for the Callahan Inference)

$$N_e(r) \propto r^{-1.8}$$

Muhleman (Ref. 5), analyzing Mariner 6 and Mariner 7 S-band range data, has recently reported the following electron density models:

$$N_e(r) \propto r^{-2.05} \text{ (Mariner 6)}$$

$$N_e(r) \propto r^{-2.08} \text{ (Mariner 7)}$$

Although the Muhleman and Callahan results are not completely complementary (in terms of a radially constant ratio $\epsilon = n/N_e$), they do provide a composite picture of a significantly less "steep" corona in the way of radial dependence.

That the difference between

$$N_e(r) \propto r^{-2}$$

and

$$N_e(r) \propto r^{-2.3}$$

is substantial is easily seen by assuming the commonly accepted average in situ measured value of approximately 7.5 electrons/cm³ (Refs. 6 and 7) and extrapolating back to 5 solar radii ($5 r_\odot$). One has

Radial dependence	N_e @ 1 AU, electrons/cm ³	N_e @ $5 r_\odot$, electrons/cm ³
r^{-2}	7.5	13,900
$r^{-2.3}$	7.5	42,900

so that the difference in the models at $5 r_\odot$ is seen to be a factor of approximately 3. It will thus be the purpose of the following sections to: (1) explore the theoretical basis of the phase fluctuation – electron density relationship, and (2) ascertain whether one or the other of the two proposed models for N_e is more consistent with the many other experimental observations of the corona made over the last decade or so.

II. The Phase Fluctuation—Electron Density Fluctuation Relationship

In 1968 Hollweg (Ref. 8), using a statistical ray analysis based on geometric optics similar to that originally formulated by Chandrasekhar in 1952 (Ref. 9), derived the following expression for RMS phase induced by electron density fluctuations in the solar corona:

$$\phi^2 \simeq 2 \sqrt{\pi} r_e^2 \lambda^2 \epsilon^2 \int_a^\infty [N_e(r)]^2 L_t(r) \frac{rdr}{\sqrt{r^2 - a^2}}$$

where

- λ = signal wavelength
- r_e = classical electron radius
- r = radial distance
- a = signal closest approach point
- ϵ = fluctuation to density ratio
- L_t = transverse scale of fluctuations
- N_e = electron density

Functionally similar expressions are given in Jokipii, (1969, Ref. 10), and Little (1970, Ref. 11). Little notes that this expression is valid for different functional forms of the fluctuation spectrum, with only a slight change in the numerical factor. Hollweg (1970, Ref. 12), subsequently derives the expression *specifically* for the (now commonly accepted) power law fluctuation spectrum, as follows:

$$\phi^2 \simeq 3\pi \left(\frac{\alpha - 1}{\alpha}\right) r_e^2 \lambda^2 \epsilon^2 \int_a^\infty [N_e(r)]^2 L(r) \frac{rdr}{\sqrt{r^2 - a^2}}$$

with

- $\alpha + 2$ = exponent of the three dimensional spatial spectrum $\simeq 3.5$
- L = outer scale of turbulence

Hollweg (Ref. 13, 1968) considers the relationship:

$$L_t(r) \propto r$$

to be a result of inhomogeneities expanding with a radially out-flowing solar wind. In Ref. 8, Hollweg treats the most common assumptions of a constant transverse scale and one linear with radial distance:

$$L_t(r) = 200 \text{ km}$$

and

$$L_t(r) = 30(r/r_\odot) \text{ km}$$

where

$$r_\odot = \text{solar radius}$$

Substitution of these assumptions for the transverse scale produces the following results² (with $N_e \propto r^{-(2+\xi)}$ and $n/N_e = \epsilon$):

(1) Scale constant with radial distance

$$\phi^2 \propto \int_a^\infty \frac{1}{(r^2+\xi)^2} \frac{rdr}{\sqrt{r^2-a^2}} \propto (a^{-(1.5+\xi)})^2$$

(2) Scale linear with radial distance

$$\phi^2 \propto \int_a^\infty \frac{1}{(r^2+\xi)^2} \frac{r^2 dr}{\sqrt{r^2-a^2}} \propto (a^{-(1.0+\xi)})^2$$

²The dependence on closest approach distance (a) is obtained by noting that the integral:

$$\int_a^\infty \frac{r^\alpha dr}{\sqrt{r^2-a^2}}$$

is transformed via the substitution $r = a(\cos x)^{-1}$ to

$$a^\alpha \int_0^{\pi/2} (\cos x)^{-(1+\alpha)} dx$$

It is thus seen that the constant scale produces the relationship between RMS phase and electron density fluctuations inferred by Callahan (as described in Sect. I), and similarly, usage of the linear transverse scale produces the functional relationship argued by Berman.

The case for a linear scale is made by Little (1970, Ref. 11), and Houminer (1973, Ref. 14), among others. Their data in support of a linear scale is reproduced here in Figs. 1 and 2. More recently, Jokipii (1973, Ref. 15), and Woo (1977, Ref. 16), among others, have considered a constant correlation scale on the order of:

$$L \sim 2 \times 10^6 \text{ to } 1 \times 10^7 \text{ km}$$

Scales on this order are obtained from spacecraft measured correlation times (τ_c) at approximately 1 AU of:

$$\tau_c \sim 6 \times 10^3 \text{ to } 3 \times 10^4 \text{ s}$$

with

$$L \sim v_r \times \tau_c$$

where

$$v_r = \text{solar wind velocity } (\sim 350 \text{ km/s})$$

What is thus obtained is a *radial* correlation length at approximately 1 AU; more appropriate for usage with a columnar phase measurement such as doppler noise would be a linear *transverse* correlation length as described by Hollweg and experimentally observed by Little and Houminer.

Using Viking S-Band doppler noise and near simultaneous Viking S-X range data, Berman (1977, Ref. 3), derived the following scale for 60-s sample interval doppler noise (time scale of the observations $\sim 15 \times 60$ s):

$$L(a) = \left[\frac{0.43}{\epsilon^2} \right] (a/r_\odot), \text{ km}$$

or, assuming nominal bounding values of ϵ :

$$L(a) = 43(a/r_\odot), \text{ km; } \epsilon = 0.1$$

$$L(a) = 4300(a/r_\odot), \text{ km; } \epsilon = 0.01$$

ORIGINAL PAGE IS
OF POOR QUALITY

Factoring into the scale the fluctuation frequency (ν) dependence (Berman, 1977, Ref. 17), one has:

$$L(a) = \frac{0.43}{\epsilon^2} \left(\frac{a}{r_\odot} \right) \left(\frac{\tau}{60} \right)^{1.4}, \text{ km}$$

where

τ = doppler sample interval, s

$$\nu \approx (2 \times 15 \times \tau)^{-1}$$

It is thus seen (i.e., given the *well determined* radial dependence of phase fluctuations and the Hollweg derived relationship) that if other experimental observations of the radial dependence of electron density support a $r^{-2.3}$ corona, the case for a linear transverse scale is considerably strengthened.

III. Experimental Observations of Electron Density in the Extended Corona

Over the last decade, a sizeable number of experiments, utilizing a variety of techniques, have been performed to measure and determine electron density in the extended corona. Table 1 is a comprehensive listing of (the partial results of) these experiments. Assuming an electron density of the form:

$$N_e(r) \propto r^{-(2+\xi)}$$

the table presents the determined (or calculated) values of ξ . Of the thirteen values listed, the mean value is:

$$\bar{\xi} = 0.298$$

with a one standard deviation of

$$1\sigma(\xi) = 0.17$$

These results would certainly appear to argue strongly for an average corona of:

$$N_e(r) \propto r^{-2.3}$$

Many of these determinations of electron density utilized data whose closest approach points were in a region $\ll 1$ AU. A slightly different method of proceeding would be to select electron density values from some "interior" region approxi-

mating the region of observations, and then apply the powerful boundary condition of the *average* electron density at 1 AU:

$$N_e(215 r_\odot) \approx 7.5 \text{ electrons/cm}^3$$

which is (reasonably) well known from in situ spacecraft measurements. In this regard, the following electron density values at $10 r_\odot$ are selected:

Newkirk (1967, Ref. 24; a compilation of techniques	9800 electrons/cm ³
Saito (1970, Ref. 23; photometry and coronameter measurements with heliographic latitude solved for)	8100 electrons/cm ³
Counselman (1972, Ref. 20; pulsar data with heliographic latitude solved for)	8400 electrons/cm ³
Weisberg (1976, Ref. 19; pulsar data with heliographic latitude solved for)	8000 electrons/cm ³

The above yield an average electron density at $10 r_\odot$ of:

$$N_e(10 r_\odot) = 8575 \text{ electrons/cm}^3$$

It is encouraging that these diverse values for $r = 10 r_\odot$ are nicely clustered about the mean value. Solving for a corona of the form:

$$N_e \propto r^{-(2+\xi)}$$

where

$$N_e(10 r_\odot) = 8575 \text{ electrons/cm}^3$$

$$N_e(215 r_\odot) = 7.5 \text{ electrons/cm}^3$$

yields:

$$\xi = 0.30$$

and

$$N_e(r) \propto r^{-2.3}$$

or, a similar value to the average computed from Table 1.

IV. Relationship Between Solar Wind Velocity and Density

One writes the condition for constant mass efflux (Hollweg, 1968, Ref. 13), as:

$$F = N_e(r)v_r(r)r^2$$

where

$$F = \text{constant}$$

$$v_r(r) = \text{radial component of solar wind velocity}$$

Hence, one might expect:

$$N_e = \frac{F}{r^2 v_r(r)}$$

Now at 1 AU, the average solar wind velocity is reasonably well known (Hundhausen, 1972, Ref. 26; several years of Vela spacecraft data):

$$v_r(215 r_\odot) \approx 400 \text{ km/s}$$

At $r = 10 r_\odot$, one can use values from Models³ of Hartle and Barnes (Ref. 26), Wolff, Brandt, and Southwick (Ref. 26), and Brandt, Wolff, and Cassinelli (Ref. 27), as follows:

$$v_r(10 r_\odot) \sim 170 \text{ km/s (Hartle and Barnes)}$$

$$v_r(10 r_\odot) \sim 185 \text{ km/s (Wolff, Brandt, and Southwick)}$$

$$v_r(10 r_\odot) \sim 160 \text{ km/s (Brandt, Wolff, and Cassinelli)}$$

or, an average value at $r = 10 r_\odot$ of:

$$v_r(10 r_\odot) \sim 172 \text{ km/s}$$

Assuming a power law⁴ model for the solar wind velocity

³The various models depicting an increasing solar wind velocity with radial distance are substantiated by experimental observations (e.g., Ekers, 1970, Ref. 28).

⁴It is recognized that a power law assumption for the solar wind velocity in the extended corona is only approximate, similar to the power law assumption for density in the extended corona.

and solving for the resultant *radial dependence* of the solar wind velocity, one has:

$$v_r(r) \propto r^{0.28}$$

hence, the condition of constant mass efflux predicts:

$$N_e(r) = F \frac{1}{r^2 v_r(r)}$$

$$= F \frac{1}{r^2 r^{0.28}}$$

$$\propto r^{-2.28}$$

or, once again, the familiar value.

One thus notes that for a coronal electron density of the form

$$N_e \propto r^{-(2+\xi)}$$

the parameter ξ can be simply identified as the radial dependence of the solar wind velocity.

V. Comparison of RMS Phase to the Scintillation Index

Using dual frequency Pioneer 9 spacecraft data, H. Chang (1976, Ref. 29), was able to make simultaneous observations of the integrated electron density (I) and the scintillation index⁵ (m). He found the scintillation index to be proportional to the integrated electron density, $m \propto I$.

Both the scintillation index and RMS phase are derived from the integrated temporal columnar fluctuation spectrum (P):

$$m^2(a) \sim \int P(a, \nu) d\nu$$

$$\phi^2(a) \sim \int P(a, \nu) d\nu$$

⁵A measure of received signal level variations induced by electron density fluctuations along the signal path.

Hence, Chang's data should certainly imply that RMS phase is also proportional to integrated electron density. The condition of $\phi(a) \propto I$ combined with

$$n/N_e = \epsilon; \epsilon \neq \epsilon(r)$$

requires (as shown in Sect. II):

$$L_t(r) \propto r$$

Thus, Chang's findings additionally substantiate the case for a linear transverse scale.

VI. Conclusions

It is here concluded that $N_e(r) \propto r^{-2.3}$ is a very reasonable assumption for the average radial dependence of electron density in the extended corona, based on the very favorable comparisons to:

- (1) Other experimental observations of the radial dependence of electron density.
- (2) The predicted and observed behavior of the solar wind velocity.
- (3) The observed relationship between the scintillation index and integrated electron density.

Accepting this conclusion, the following observations and assumptions form a self-consistent set in the region $5r_\odot \leq r \leq 1$ AU:

$$\phi(a) \propto a^{-1.3}$$

$$N_e(r) \propto r^{-2.3}$$

$$L_t(r) \propto r$$

$$n(r) \propto r^{-2.3}$$

$$n/N_e = \epsilon; \epsilon \neq \epsilon(r)$$

$$v_r(r) \propto r^{0.3}$$

On the other hand, if one combines the Callahan inference with the *least* steep corona experimentally reported (Muhleman; Mariner 6, 1977) and the radial dependence of RMS phase:

$$\phi(a) \propto a^{-1.3}$$

$$n(r) \propto r^{-1.8}$$

$$N_e(r) \propto r^{-2.05}$$

then it is required that

$$L(r) \propto L_0$$

$$n/N_e = \epsilon(r)$$

$$\propto r^{0.25}$$

$$v_r(r) \propto r^{0.05}$$

In regard to these required conditions, it is difficult to accept that the turbulence per unit density (ϵ) increases with radial distance; even more difficult to reconcile is the required near constancy ($\propto r^{0.05}$) of the solar wind velocity for $5r_\odot \leq r \leq 1$ AU, in contradiction to the predicted and observed *average* radial dependence ($\propto r^{0.3}$) of the solar wind velocity in this region.

Acknowledgments

I would like to thank Dan Cain, who first drew my attention to the work of J. V. Hollweg; and Lee Brunn, who kindly provided me with a copy of H. Chang's dissertation.

References

1. Berman, A. L., and Wackley, J. A., "Viking S-Band Doppler RMS Phase Fluctuations Used to Calibrate the Mean 1976 Equatorial Corona," in *The Deep Space Network Progress Report 42-38*, Jet Propulsion Laboratory, Pasadena, Calif., 15 April 1977.
2. Berman, A. L., Wackley, J. A., Rockwell, S. T., and Kwan, M., "Viking Doppler Noise used to Determine the Radial Dependence of Electron Density in the Extended Corona," in *The Deep Space Network Progress Report 42-38*, Jet Propulsion Laboratory, Pasadena, California, April 15, 1977.
3. Berman, A. L., "Proportionality Between Doppler Noise and Integrated Signal Path Electron Density Validated by Differenced S-X Range," in *The Deep Space Network Progress Report 42-38*, Jet Propulsion Laboratory, Pasadena, California, April 15, 1977.
4. Callahan, P. S., "A Preliminary Analysis of Viking S-X Doppler Data and Comparison to Results of Mariner 6, 7, and 9 DRVID Measurements of the Solar Wind Turbulence," in *The Deep Space Network Progress Report 42-39*, Jet Propulsion Laboratory, Pasadena, California, June 15, 1977.
5. Muhleman, D. O., Esposito, P. B., and Anderson, J. D., "The Electron Density Profile of the Outer Corona and the Interplanetary Medium from Mariner 6 and Mariner 7 Time Delay Measurements," *Astrophysical Journal*, 211, February 1, 1977.
6. Hundhausen, A. J., Bame, S. J., Asbridge, J. R., and Sydoriak, S. J., "Solar Wind Proton Properties: Vela 3 Observations from July 1965 to June 1967," in *The Journal of Geophysical Research, Space Physics*, Vol. 75, No. 25, September 1, 1970.
7. Ness, N. F., Hundhausen, A. J., and Bame, S. J., "Observations of the Interplanetary Medium: Vela 3 and Imp 3, 1965-1967," in *The Journal of Geophysical Research*, Vol. 76, No. 28, October 1, 1971.
8. Hollweg, J. V., "A Statistical Ray Analysis of the Scattering of Radio Waves by the Solar Corona," in *The Astronomical Journal*, Vol. 73, No. 10, Part 1, December 1968.
9. Chandrasekhar, S., "A Statistical Basis for the Theory of Stellar Scintillation," *The Monthly Notices of the Royal Astronomical Society*, 5, 1952.
10. Jokipii, J. R., and Hollweg, J. V., "Interplanetary Scintillations and the Structure of Solar Wind Fluctuations," in *The Astrophysical Journal*, Vol. 160, May 1970.
11. Little, L. T., "Small Scale Plasma Irregularities in the Interplanetary Medium," in *Astronomy and Astrophysics*, 10, 1971.
12. Hollweg, J. V., "Angular Broadening of Radio Sources by Solar Wind Turbulence," in *The Journal of Geophysical Research, Space Physics*, Vol. 75, No. 19, July 1, 1970.
13. Hollweg, J. V., and Harrington, J. V., "Properties of Solar Wind Turbulence Deduced by Radio Astronomical Measurements," in *The Journal of Geophysical Research, Space Physics*, Vol. 73, No. 23, December 1, 1968.
14. Houminer, Z., "Power Spectrum of Small Scale Irregularities in the Solar Wind," *Planetary Space Science*, Vol. 21, 1973.

ORIGINAL PAGE IS
OF POOR QUALITY

15. Jokipii, J. R., "Turbulence and Scintillations in the Interplanetary Plasma," in *The Annual Review of Astronomy and Astrophysics*, 1973.
16. Woo, R., Yang, F., Yip, W. K., and Kendall, W. B., "Measurements of Large Scale Density Fluctuations in the Solar Wind Using Dual Frequency Phase Scintillations," in *The Astrophysical Journal*, Vol. 210, No. 2, Part 1, December 1, 1976.
17. Berman, A. L., "Phase Fluctuation Spectra – New Radio Science Information to Become Available in the DSN Tracking System, Mark III-77, in *The Deep Space Network Progress Report 42-40*, Jet Propulsion Laboratory, Pasadena, California, August 15, 1977.
18. Edenhofer, P., Esposito, P. B., Hansen, R. T., Hansen, S. F., Lueneburg, E., Martin, W. L., Zygielbaum, A. I., "Time Delay Occultation Data of the Helios Spacecrafts for Probing the Electron Density Distributions in the Solar Corona," to be published in *The Journal of Geophysics*.
19. Weisberg, J. M., Rankin, J. M., Payne, R. R., and Counselman III, C. C., "Further Changes in the Distribution of Density and Radio Scattering in the Solar Corona," *The Astrophysical Journal*, Vol. 209, October 1, 1976.
20. Counselman III, C. C., and Rankin, J. M., "Density of the Solar Corona from Occultations of NP0532," *The Astrophysical Journal*, Vol. 175, August 1, 1972.
21. Muhleman, D. O., Anderson, J. D., Esposito, P. B., Martin, W. L., *Radio Propagation Measurements of the Solar Corona and Gravitational Field: Applications to Mariner 6 and 7*, Technical Memorandum 33-499, Jet Propulsion Laboratory, Pasadena, California, 1971.
22. Muhleman, D. O., Ekers, R. D., and Fomalont, E. B., "Radio Interferometric Test of the General Relativistic Light Bending Near the Sun," *Physical Review Letters*, Vol. 24, No. 24, 15 June 1970.
23. Saito, K., "A Non-Spherical Axisymmetric Model of the Solar K Corona of the Minimum Type," *Annals of the Tokyo Astronomical Observatory*, University of Tokyo, Second Series, Vol. XII, No. 2, Mitaka, Tokyo, 1970.
24. Newkirk, G., "Structure of the Solar Corona," in *The Annual Review of Astronomy and Astrophysics*, Vol. 5, 1967.
25. Anderson, J. D., Esposito, P. B., Martin, W. L., Thornton, C. L., and Muhleman, D. O., "Experimental Test of General Relativity Using Time Delay Data from Mariner 6 and Mariner 7," in *The Astrophysical Journal*, Vol. 200, August 15, 1975.
26. Hundhausen, A. J., *Coronal Expansion and Solar Wind*, New York, Springer-Verlag, 1972.
27. Brandt, J. C., *Introduction to the Solar Wind*, W. H. Freeman and Company, San Francisco, 1970.
28. Ekers, R. D., and Little, L. T., "The Motion of the Solar Wind Close to the Sun," *Astronomy and Astrophysics*, Vol. 10, 1971.
29. Chang, H., *Analysis of Dual-Frequency Observations of Interplanetary Scintillations Taken by the Pioneer 9 Spacecraft*, Doctoral Dissertation, Department of Electrical Engineering, Stanford University, May 1976.

Table 1. Electron density measurements in the solar corona

Source	Reference	Year	ξ	Type of measurement
Edenhofer	18	1977	0.2	S-band range, Helios
Berman	2	1977	0.30	S-band doppler noise, Viking
Muhleman	5	1977	0.05	S-band range, Mariner 6
Muhleman	5	1977	0.08	S-band range, Mariner 7
Weisberg	19	1976	0.3 ^a	Pulsar time delay
Counselman	20	1972	0.4 ^b	Pulsar time delay
Muhleman	21	1971	0.41	S-band range, Mariner 6
Muhleman	22	1970	0.33	Radio interferometry
Saito	23	1970	0.5	Photometry
Newkirk	24	1967	0.34 ^c	Compilation of techniques
Blackwell	25	1967	0.33	Solar eclipse
Blackwell	25	1967	0.33	Solar eclipse
Blackwell	25	1966	0.3	Solar eclipse

^aOne of several solutions; this solution in best agreement with average in situ density values at 1 AU.

^bOne of several solutions; this solution included heliographic latitude.

^cComputed between $N_e(10r_\odot)$ and average in situ value (7.5 electrons/cm³) at 1 AU.

ORIGINAL PAGE IS
OF POOR QUALITY

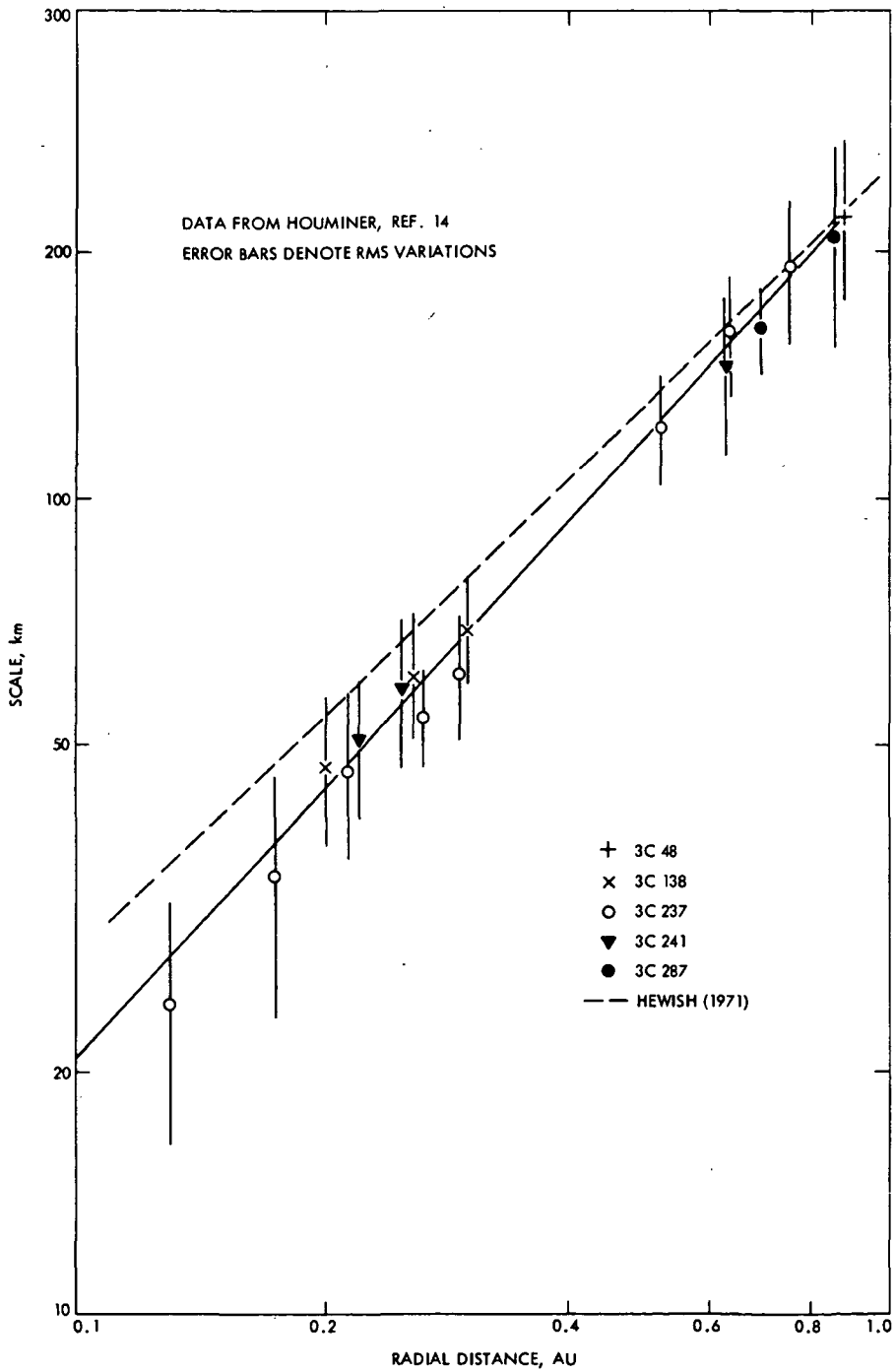


Fig. 1. Scale size of electron density irregularities at heliocentric distances between $0.1 \text{ AU} < r < 1.0 \text{ AU}$

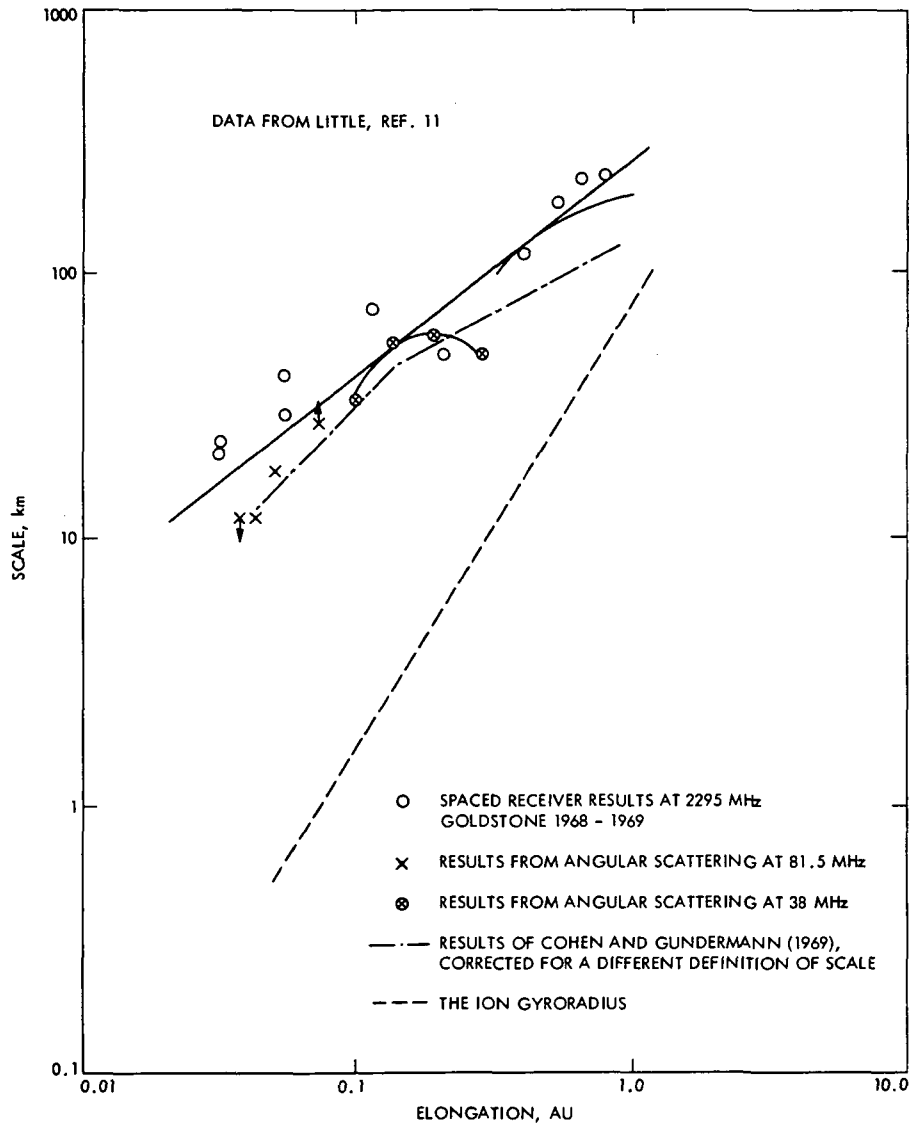


Fig. 2. Scale *l* of small-scale plasma irregularities as a function of solar elongation

ORIGINAL PAGE IS
OF POOR QUALITY

N78-11149

S-/X-Band Microwave Optics Design and Analysis for DSN 34-Meter-Diameter Antenna

D. L. Nixon and D. A. Bathker
Radio Frequency and Microwave Subsystems Section

This article discusses the feed system configuration, RF optics design and predicted gain and noise temperature performance of the DSN 26 meter antenna S-/X-Band upgrade project. A previously developed and demonstrated dual band feed system employing separately located feeds, combined in a reflex-dichroic reflector arrangement, has been adapted to a particularly compact overall design. Special design techniques to minimize aperture blockage and then evaluate resultant performance are discussed in detail. Final predicted antenna gain and noise temperature performance values are given.

I. Introduction

A DSN 26-m antenna S-X Conversion Project has been undertaken to extend the capability of an existing 26-m subnet. This is being done to support typical long-lived outer planet missions and to offload the heavily scheduled 64-m network during mission cruise phases. The 26-m meter S/X Conversion Project has been described from an overall project view in Ref. 1. This report will cover the RF optics design and performance estimates. The modifications to the basic antenna system primarily include increasing the aperture diameter from 26 to 34-m and adding a simultaneous X-band capability. To achieve the X-band system, it was decided to continue use of the DSN reflex-dichroic feed (Ref. 2) very similar to that being used by the 64-m subnet. In the 34-m case, a single new Cassegrain feedcone housing large enough to enclose both the S-band and X-band feed systems was designed. In order to minimize aperture blockage, the feedhorn focal point location was substantially moved toward the paraboloid vertex and necessarily offset from the paraboloid axis. As a result of this, a new subreflector and quadripod are required, and are a part of the conversion.

This article will emphasize the microwave optics design considerations and the series of very detailed computations needed to predict overall microwave gain. Final computed estimates of overall antenna system gain and noise temperature at both S- and X-bands are given. Conclusions and recommendations to the project are included.

II. Feed System Configuration

Figure 1 shows the outline configuration of the 34-m upgraded antenna. The original 26-m-diam paraboloid focal length of 1097.28 cm (432.000 in.) is maintained; new outer panels according to the required contour are added. The initial 26-m system f/d ratio of 0.4235 is therefore substantially reduced to a new value of 0.3227; in effect a much "deeper" dish is formed. In order to minimize central blockage caused by the substantial size of the single large feedcone and associated reflex-dichroic reflectors,¹ the original 26-m-diam Cas-

¹A description of the layout of the interior of the 34-m system feedcone is available in Ref. 3.

segrain design value for the separation of primary and secondary focal points of 579.12 cm (228.000 in.) is significantly increased to 701.04 cm (276.000 in.). This results in a feedhorn position above the paraboloid vertex fully 121.92 (48.000 in.) lower than the previous 26-m designs.

In the previous design, the feedhorn was visible, over the rim of the paraboloid, to an observer at a right angle to the antenna system boresight axis. In this new design, by a combination of the deeper dish and the increase in focal point separation distance, all feed components including the large ellipsoidal reflector are now within the envelope of the rim of the dish, as seen in Fig. 1. Also seen in this figure are the new quadripod feed support members, the necessarily larger apex structure, and the asymmetric subreflector. In contrast to the fully implemented 64-m system, wherein the asymmetric subreflector is rotatable about the paraboloid axis (to obtain switching among the three separate feedcones used in that system), the 34-m design described here uses a fixed subreflector, permanently focussed upon the X-band feedhorn focal point, located below the dichroic plate.

A cross-sectional detailed view of the reflex feed system geometry is shown in Fig. 2. Four of the elements, the S- and X-band feedhorns, the ellipsoidal reflector and the planar dichroic reflector as well as the critical relative positions of these four elements, are borrowed directly from the previously demonstrated 64-m design (Refs. 2 and 4). With these critical relative positions held constant, the optics design proceeded by rocking the subassembly about the primary focal point at the radius value equal to 701.04 cm (276.000 in.), until an approximately equal central blockage condition was achieved. This occurred, as shown in both Figs. 1 and 2, with an offset angle of 7.6677 deg. This offset angle, once determined, then allows design of the subreflector to proceed.

The subreflector consists of three parts; the hyperboloidal surface contour, the non-optic peripheral flange, and the non-optic vertex plate. The peripheral flange was iteratively designed by use of computed scattered patterns, noting the decrease in power spilled past the rim of the paraboloid, at both S- and X-bands. This is a key parameter in determining the noise performance of the finished design. The vertex plate was added to decrease the RF illumination into the central zone containing the feedcone and reflex reflectors.

III. Blockage

A rather novel approach in determining the blockage was used for this project. A precision scale shadow drawing (Fig. 3) was generated by using a computer. As can be seen, this includes direct optical planewave blockage as well as spherical wave shadows caused by any structure in the RF paths. The

hyperboloid outline, the apex portion of the quadripod, and the inner part of the quadripod shadows were projected on the planewave basis. All other shadows were projected from the paraboloidal focal point, on a spherical wave basis. The total blockage was then determined by mechanically integrating the complex outline to get the shadow area for all structures and comparing that to the overall paraboloid projected area. This results in a total of 4.24% for quadripod blockage and 1.82% central area blockage giving a total of 6.06% area blockage. It should be noted that this is the physical (optical) blockage; very short wavelength RF blockage will be essentially equal while longer wavelengths will typically be affected somewhat more than the optical value. At both S- and X-bands we estimate the true value to be close to the optical value.

IV. Vertex Plate

As mentioned, a vertex plate for the subreflector is considered an especially important part of the overall RF optics design for the 34-m S/X system. In many systems, vertex plates are employed totally or primarily to reduce the hyperboloid axial backscatter into the feedhorn, which results in feed system VSWR which, in turn, alternates with small changes in operating frequency because of the long (in wavelengths) path involved. As the size of a given Cassegrain system increases relative to wavelength, the behavior is such that the VSWR decreases. For a perfectly axial system of the 34-m S/X size, the estimated hyperboloid VSWR is 1.08:1 at S-band and substantially less at X-band. One concludes from this that a vertex plate for the 34-m S/X system is desirable at S-band, but not particularly useful at X-band, from a VSWR viewpoint. In this design, the large central blockage area (1.82%) caused by the substantial shadows from the feed system reflectors (ellipsoid-dichroic), leads one to question the amount of power contained within that area, assuming both feeds in transmission.² The amounts of power in the subject area were obtained from scattered patterns (without vertex plate) as 4.7 and 4.0% at S- and X-bands, respectively. If uncontrolled, this power would be scattered into generally unpredictable wide-angle pattern regions, by way of radiation off of the backside of the ellipsoidal reflector, feedcone sides, etc. One concludes from this that a vertex plate for the 34-m S/X system is highly desirable at both frequency bands.

A vertex plate was designed on an iterative basis using primarily the scattering program (FVSCATT) and the radiation pattern efficiency program (EFFIC). Also, other programs

²Reciprocity can alternately be considered to show the feeds in reception will be sensitive, at that power percentage, to whatever radiation (noise, perhaps) might emanate from that area (via reflections, for example).

used were (CASSDESIGN), (AZEXP), and (ANTPLOT).³ Vertex plates tend to be narrowband devices (in the sense of a frequency span from S- to X-bands), by virtue of the principle of displaced phase centers (displaced from the hyperboloid center of radiation). Nonetheless, a design was obtained and evaluated, resulting in the following performance. At S-band, the 4.7% central blockage zone power was reduced to 1.7%. The corresponding central blockage efficiency factor, (a function of blocked area and power) increased from 0.9197 to 0.9651. At X-band, the 4.0% central blockage zone power was reduced to 0.5%. The corresponding central blockage efficiency increased from 0.9430 to 0.9771.

On an overall efficiency basis, however, the above increases in the central blockage efficiency term are not necessarily realized as a final benefit. This is because the aperture illumination phase efficiency is typically degraded somewhat by including a vertex plate in a given design. In the 34-m S/X design discussed here, the phase efficiency degradations incurred as a result of using a vertex plate are 0.9 and 4.5% at S- and X-bands respectively. Other factors such as aperture illumination amplitude efficiency are impacted in minor ways; the overall efficiency at S-band was improved 0.14 dB by use of the vertex plate while the X-band was degraded 0.18 dB. In view of the overall operating system gain to noise temperature ratio, the tradeoff of 0.18 dB of X-band gain for the benefits of controlling most of the 4% stray radiation at X-band is considered highly desirable.

V. Azimuthal Mode Analysis

As seen in Figs. 1 and 2, the 34-m design includes the large offset feed angle of 7.7 deg as a key unavoidable feature of the design. This is a necessarily large angle due to the relatively shorter focal length of the 34-m antenna. One result of the large offset feed angle inherent in this size antenna is that the subreflector is substantially more asymmetric than in our previous 64-m designs. For example, the 64-m offset feed angle is approximately 4.5 deg; a gain degradation of 0.009 dB results from $m \neq 1$ mode conversion due to this amount of asymmetry (Ref. 4). One of the major design complexities of the 34-m RF optics design was to understand the impact of the expected additional mode conversion and associated gain degradation. As will be shown in the section on computed antenna gain, the gain degradation increased to 0.042 dB due to the greater relative asymmetry of the 34-m subreflector.

The azimuthal analysis of feed radiation patterns is well documented in previous literature (Refs. 2, 5, 6, 7), showing that only $m = 1$ components of the pattern contribute to the

antenna gain. Energy contained in the $m \neq 1$ components represents a degradation of antenna gain and, in a receive mode, a possible source of system noise temperature degradation.

A computer program (AZEXP) was used to individually compute the azimuthal mode component patterns. The input to this program is a group of radiation patterns obtained from the (FVSCATT) program. The computed output from (AZEXP) is a series of polar patterns for each m index from 0 to 3, together with the total power in each of the m -component patterns. The resulting modal power for S- and X-band as scattered from the subreflector is presented in Fig. 4. As can be seen in Fig. 4 there is significant power in the $m \neq 1$ modes. Not all of the reductions seen in the $m \neq 1$ terms in Fig. 4 result in overall antenna gain loss, however. Figure 4 addresses the relative amounts of mode power while the previously mentioned gain degradation of 0.042 dB (at X-band), which will be discussed further in Table 2, is representative of that portion that is destructive to the desired polarization. When using circular polarization, there is enough energy in these $m \neq 1$ modes to cause shift of the secondary beams of the overall antenna system. A preliminary analysis of this beam shift indicates a 0.004-deg shift at X-band and a 0.04-deg shift at S-band. (The corresponding gain loss at S-band, assuming perfect pointing is achieved at X-band, is 0.34 dB).

VI. Total System Noise Temperature

The total system noise temperature as a function of elevation angle is expected to be very similar to the 64-m antennas. The values shown in Fig. 5 are estimates of what to expect from the new 34-m design. The tolerance is ± 3.0 K for X-band and ± 2.5 K for S-band. The graphs do not extend below 20 deg elevation because of the difficulty of accurate prediction at such low angles. Final performance values will be determined from field measurements.

VII. Computed Performance

The design and evaluation of the 34-m S/X upgrade involved many computer programs. The major programs are listed in Table 1.

A brief narrative of the computational process is as follows, starting with the S-band system. A standard JPL corrugated conical feedhorn (Ref. 8) is used. First, the horn E- and H-plane amplitude and phase data from radiation patterns shown in Fig. 6 are numerically scattered off the ellipsoidal reflector using the computer program (FSCAT). The output from this program is shown in Fig. 7 and is then scattered off of the dichroic flat plate, using (FSCATT) a second time. That

³A summary of these and other JPL programs used is given in Table 1.

output, shown in Fig. 8 is then scattered off of the subreflector including a vertex plate and flange using (FVSCATT). This final output shown in Fig. 9 can then be evaluated using (AZEXP) and (EFFIC) to determine the overall illumination efficiency of the main paraboloid. It should be noted that a complex set of antenna patterns are used throughout this process with m modes of $m = 0$, $m = 1$, $m = 2$, and $m = 3$ included. This preserves and sums the beam asymmetry accumulation as it develops from each surface.

The X-band system evaluation is much simpler, because no asymmetric ellipsoid is present; of course, the asymmetric subreflector introduces the need for azimuthal modal decomposition analysis. The feedhorn radiation patterns (Fig. 6) (the S- and X-band feedhorn are the same design simply scaled for frequency) were scattered off of the same subreflector as the S-band system using (FVSCATT). The resultant illumination patterns are shown in (Fig. 10). This was the final scattering output and was evaluated using (AZEXP) and (EFFIC) similar to the S-band process.

VIII. Calculated Antenna Gain

Using the antenna patterns that were generated as described earlier, and at each stage preserving the azimuthal modes $m = 0$, 1, 2 and 3, the various antenna efficiency components are computed as shown in Table 2.

The predicted paraboloid/subreflector surface tolerance ranges from a low of 0.58 mm at reasonable elevation angles with no wind or thermals to a maximum of 1.0 mm at 60 deg elevation with wind and thermal distortion included. This maximum gives a gain loss at X-band of 0.51 dB and at S-band 0.04 dB. The higher RMS of 1.0 mm was used for the antenna gain in Table 2.

The gain values are defined at the input to the maser preamplifiers. The S-band gain as indicated in Table 2 is for the S-band low noise system. The expected gain for the S-band diplex mode is 0.09 dB less. The expected tolerance for X-band is +0.3 dB to -0.9 dB and the S-band gain tolerance is

± 0.6 dB. These tolerances include the major error sources which will be incurred during field measurements using natural radio sources. As stated in the section on azimuthal mode analysis, beam squint at S-band, due to higher order mode generation as a result of the asymmetric design, will degrade the S-band gain by 0.34 dB below the peak value indicated in Table 2. Thus, the operational system, assuming X-band beam peak pointing, is predicted to operate with +67.0 dB X-band gain and +56.1 dB S-band gain (low noise path) or +56.0 dB (diplex path).

IX. Summary and Recommendations

This report discusses the feed system configuration selected for the DSN 34-m S/X project (26-m antenna upgrade). A previously developed and demonstrated dual-band feed system employing separately located conventional feeds, combined in a reflex-dichroic reflector arrangement, has been adapted to a particularly compact overall design. It is an inescapable characteristic of this design that moderate to high central aperture blockage exists, as well as a higher degree of offset feed asymmetry than before. For these reasons, special analyses were done emphasizing blockage minimization and evaluation, including vertex plate design, as well as azimuthal mode analysis of the resultant moded illumination patterns on the paraboloid surface. Secondary beam squint was obtained at both frequency bands. Nevertheless, quite acceptable overall antenna gain and noise temperature performance at both frequency bands is indicated.

It is recommended that the project budget sufficient radio frequency (RF) system performance test time, following initial installation, and plan for a full evaluation of RF system parameters, especially overall system noise temperature at low elevation angles. Because the CONSCAN secondary beam pointing scheme as well as the subreflector focusing as a function of HA-Dec angles are very important to the RF performance achievement goals of the project, sufficient additional test time for a complete evaluation of those parameters is necessary. Finally, secondary beam squint between the two frequency bands requires evaluation in the field.

ORIGINAL PAGE IS
OF POOR QUALITY

References

1. Lobb, V. B. , "26 Meter S-X Conversion Project," *The Deep Space Network Progress Report No. 39*, Jet Propulsion Laboratory, Pasadena, Calif., June 15, 1977.
2. Potter, P. D., "S- and X-Band Feed System," Technical Report 32-1526, Vol. VIII, pp. 53-60, Jet Propulsion Laboratory, Pasadena, Calif., April 15, 1972.
3. Hartop, R. W., "Dual Frequency Feed System for 26-Meter Antenna Conversion," *DSN Progress Report 42-40*, pp. 146-149. Jet Propulsion Laboratory, Pasadena, Calif., Aug. 15, 1977.
4. Bathker, D. A., "Dual Frequency Dichroic Feed Performance," presented at the XXVI Meeting of the Arronus Panel, NATO-AGARD, Munich, Germany, Nov. 26-30, 1973.
5. Ludwig, A. C., "Calculation of Scattered Patterns from Asymmetrical Reflectors," Technical Report 32-1430, Jet Propulsion Laboratory, Pasadena, Calif., Feb. 15, 1970.
6. Potter, P. D., "Application of Spherical Wave Theory to Cassegrainian-Fed Paraboloids," *IEEE Trans. Ant. Prop.*, Vol. AP-15, No. 6, pp. 727-736, November 1967.
7. Ludwig, A. C., "Near-Field Far-Field Transformations Using Spherical-Wave Expansions," *IEEE Trans. Ant. Prop.*, Vol. AP-19, No. 2, pp. 214-220.
8. Brunstein, S. A., "A New Wideband Feedhorn with Equal E- and H-Plane Beamwidths and Suppressed Sidelobes," *Space Programs Summary 37-58*, Vol. II, pp. 61-64, Jet Propulsion Laboratory, Pasadena, Calif., July 1969.
9. Ludwig, A., "Antenna Feed Efficiency," *Space Programs Summary No. 37-26*, Vol. IV, pp. 200-208, Jet Propulsion Laboratory, Pasadena, Calif., April 30, 1964.
10. Potter, P. D., *A Computer Program for Machine Design of Cassegrain Feed Systems*, Technical Report 32-1202, Jet Propulsion Laboratory, Pasadena, Calif., Dec. 15, 1967.

Table 1. Antenna design computer programs

Name	Use
SPWEXP	Spherical wave expansion and fields calculation (Ref. 7)
FSCATT	Asymmetrical scattering (Ref. 5) (general conic surface)
AZEXP	Antenna pattern aximuthal fourier analysis (Ref. 5, Appendix A)
ANTPLOT	Antenna pattern plotting
EFFIC	General case efficiency (Ref. 9)
ANTDESIGN	Offset feed geometry
CASSDESIGN	Cassegrain feed system design (Ref. 10)
FVSCATT	Asymmetrical scattering (hyperbolic subreflector with flange and vertex plate)
NEWGRID	Grid spacing for FVSCATT
FLATGRID	Flat grid for FSCATT (dichroic plate)
FLATSURF	Flat surface input to FSCATT (dichroic plate)
FSURFELP	Ellipsoidal surface for FSCATT
FSURFGRID	Ellipsoidal grid for FSCATT

Table 2. 34-m Reflex feed system gain — S- and X-bands

Efficiency Component	X-Band 8.415 GHz		S-Band 2.295 GHz	
	Ratio	dB	Ratio	dB
Forward spillover	0.96005	-0.1771	0.98231	-0.0775
Rear spillover	0.99568	-0.0188	0.99637	-0.0158
Non-uniform amplitude illumination	0.84289	-0.7423	0.83582	-0.7789
Non-uniform phase illumination	0.91297	-0.3954	0.96398	-0.1593
Cross polarization	0.99942	-0.0025	0.99975	-0.0011
Mode conversion ($m \neq 1$)	0.99038	-0.0420	0.97503	-0.1098
Central blockage	0.9771	-0.1006	0.9651	-0.1543
Quadripod blockage (optical)	0.9170	-0.376	0.9170	-0.376
Paraboloid/subreflector surface tolerance loss (1.0 mm RMS)	0.88209	-0.5448	0.99096	-0.0394
Ellipsoid reflector tolerance loss (0.51 mm RMS)	Not applicable		0.99759	-0.0105
Transmission loss through planar reflector	0.9950	-0.0218	Not applicable	
Waveguide dissipative loss	0.9757	-0.107	0.9772	-0.100
Overall efficiency	0.5587	-2.529	0.6572	-1.823
Gain for 100% efficiency	8.9893×10^6	69.537	6.68623×10^5	58.252
Overall gain		67.01		56.43

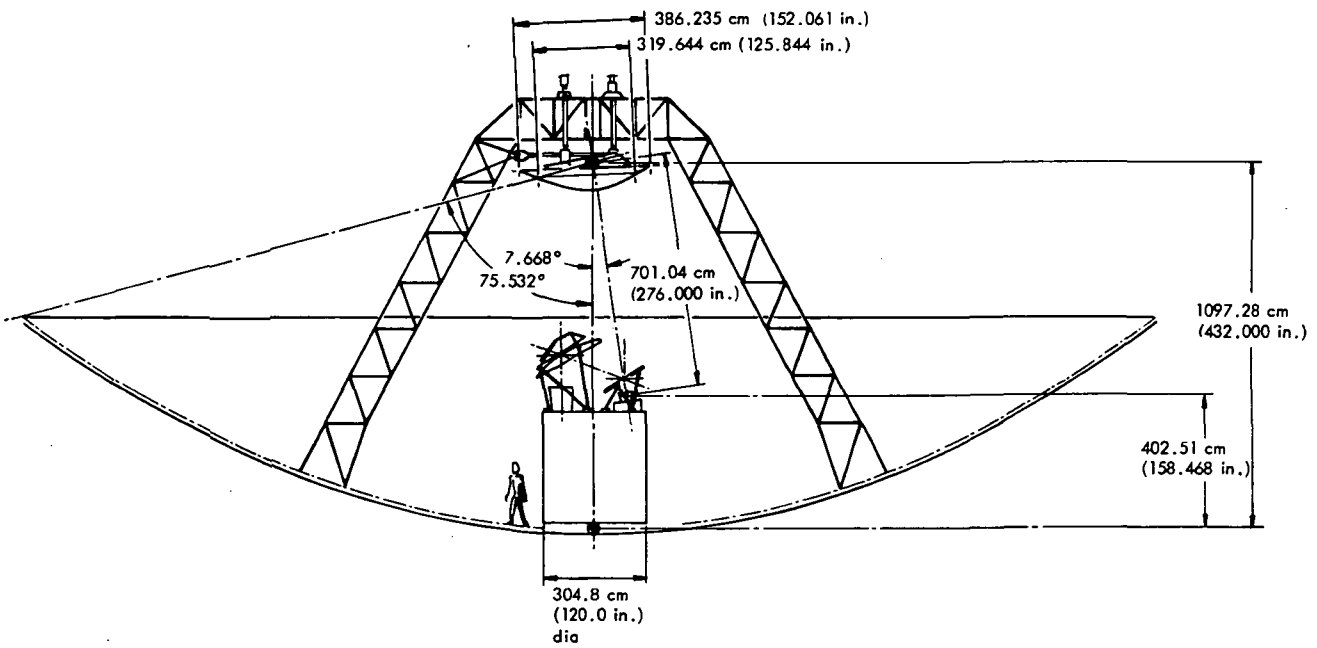


Fig. 1. 34-m S/X band antenna configuration

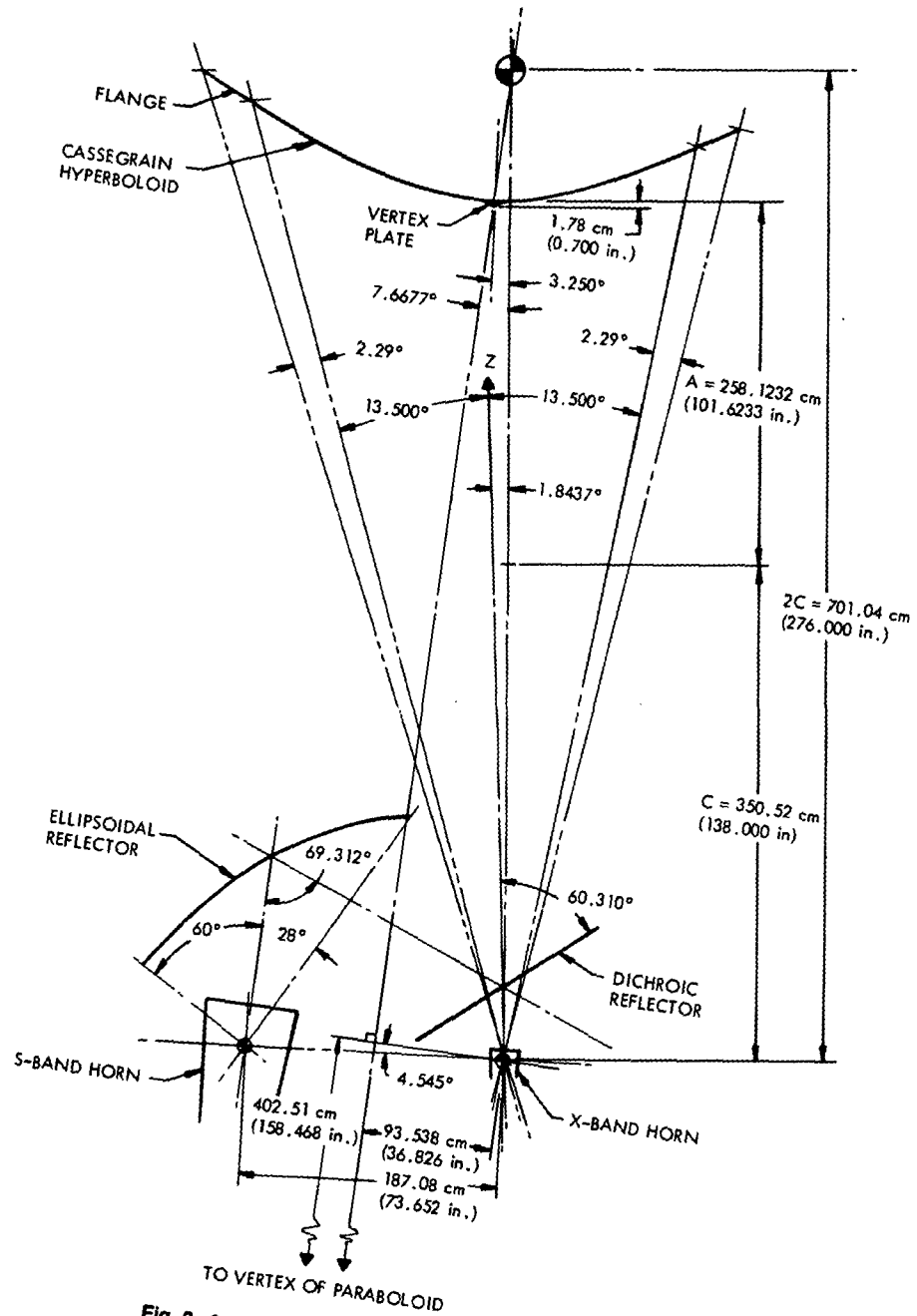


Fig. 2. 34-m S/X band reflex feed geometry

ORIGINAL PAGE IS
OF POOR QUALITY

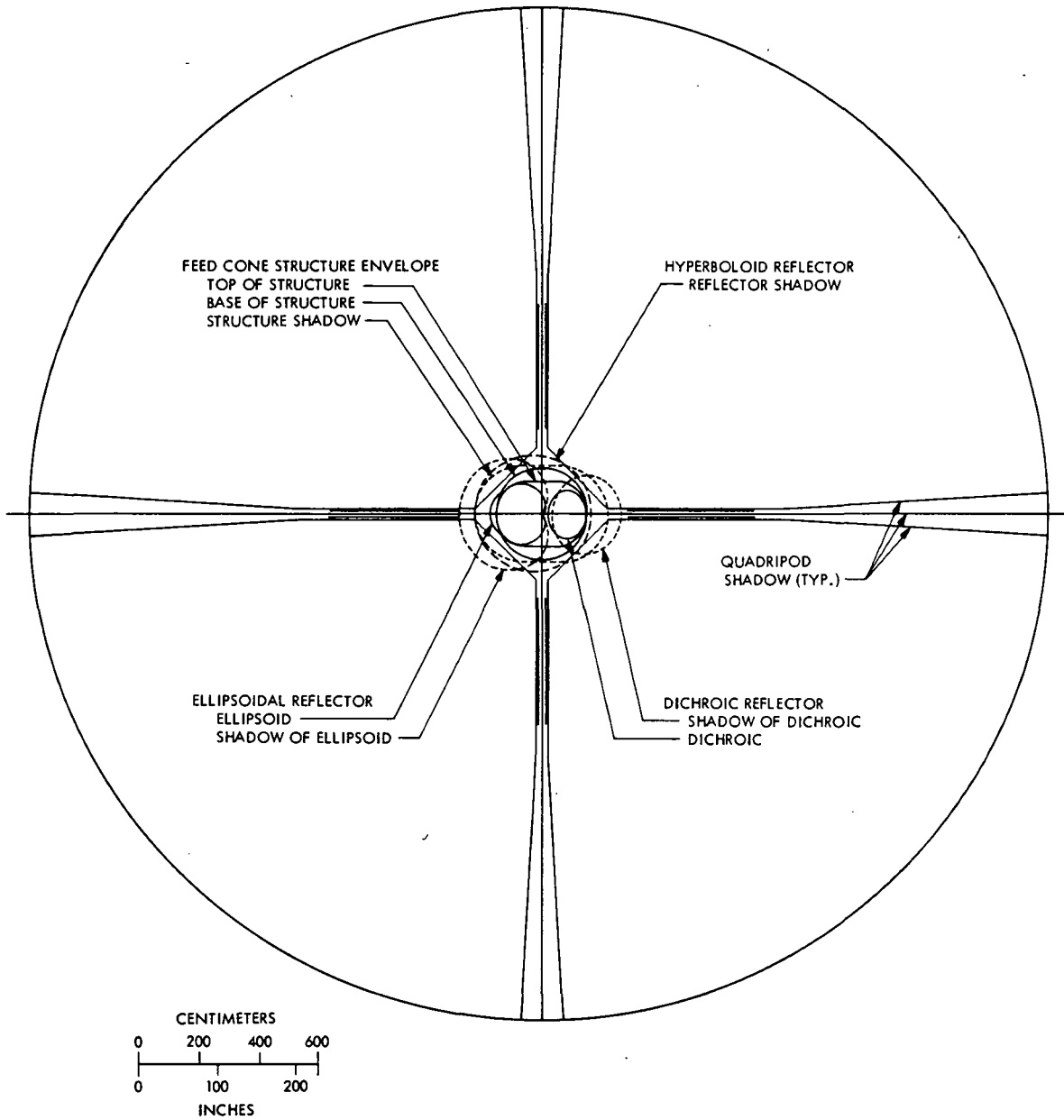


Fig. 3. Computed shadow drawing

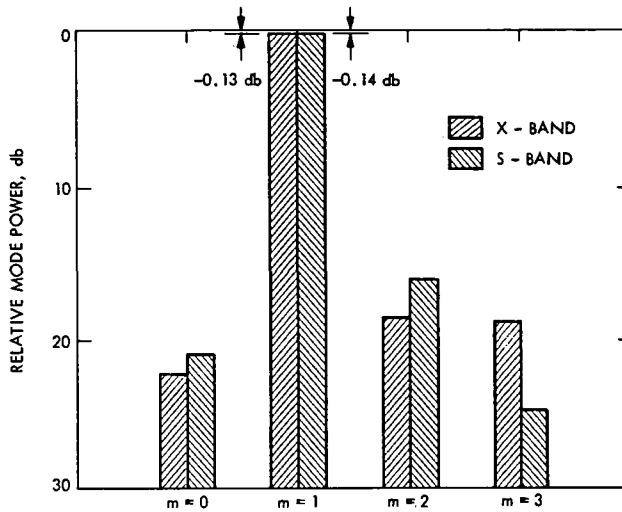


Fig. 4. Azimuthal mode power in the subreflector patterns, 34-m S/X system

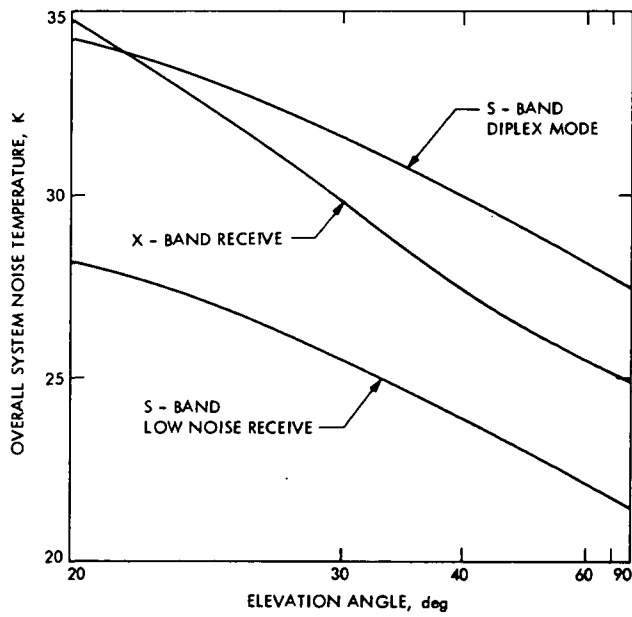


Fig. 5. 34-m S/X overall system noise temperature

ORIGINAL PAGE IS
OF POOR QUALITY

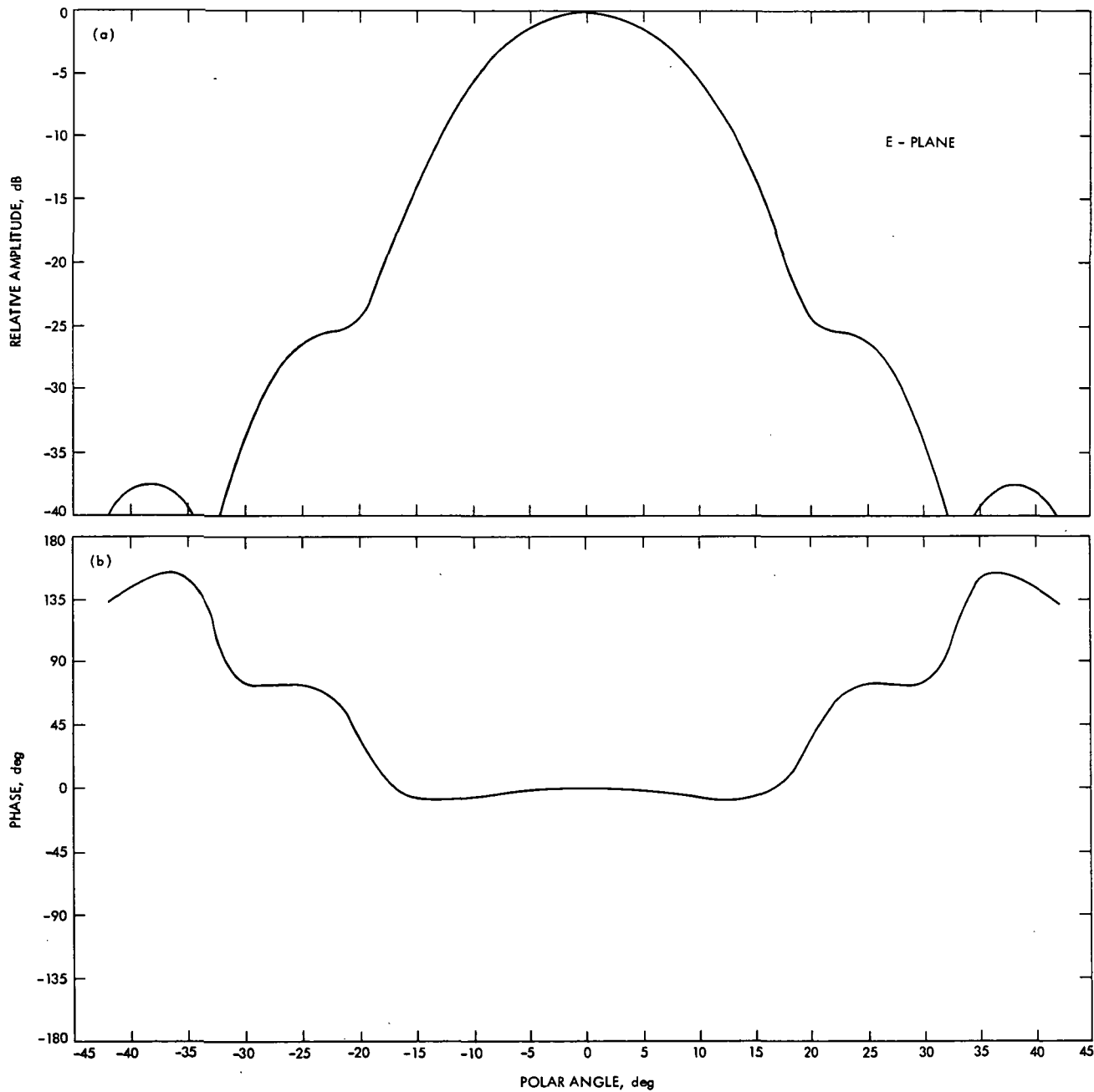


Fig. 6. S- and X-band 22-dB theoretical horn patterns: (a) E-plane amplitude and (b) E-plane phase

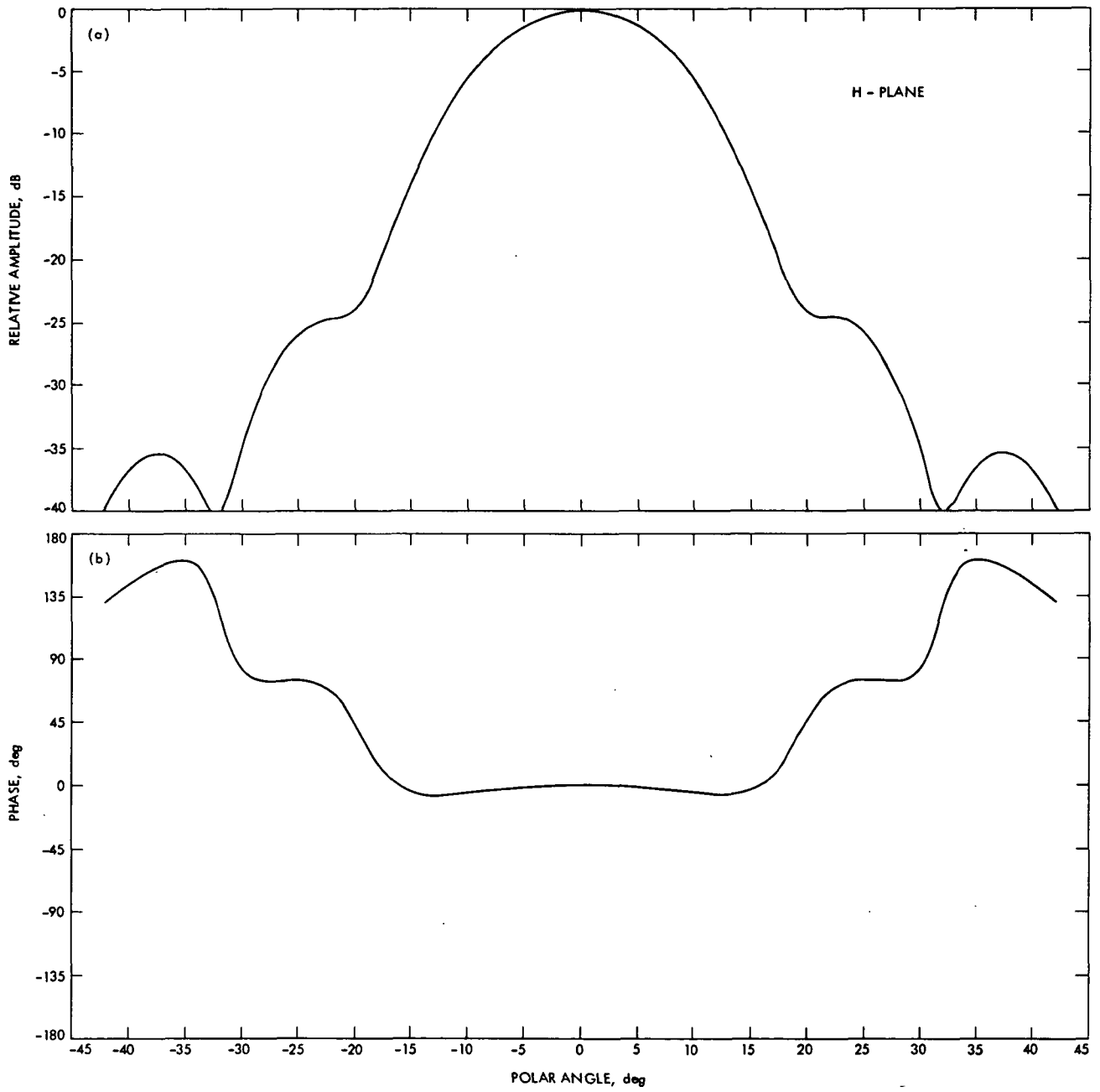


Fig. 7. S- and X-band 22 dB theoretical horn patterns: (a) H-plane amplitude and (b) H-plane phase

ORIGINAL PAGE IS
OF POOR QUALITY

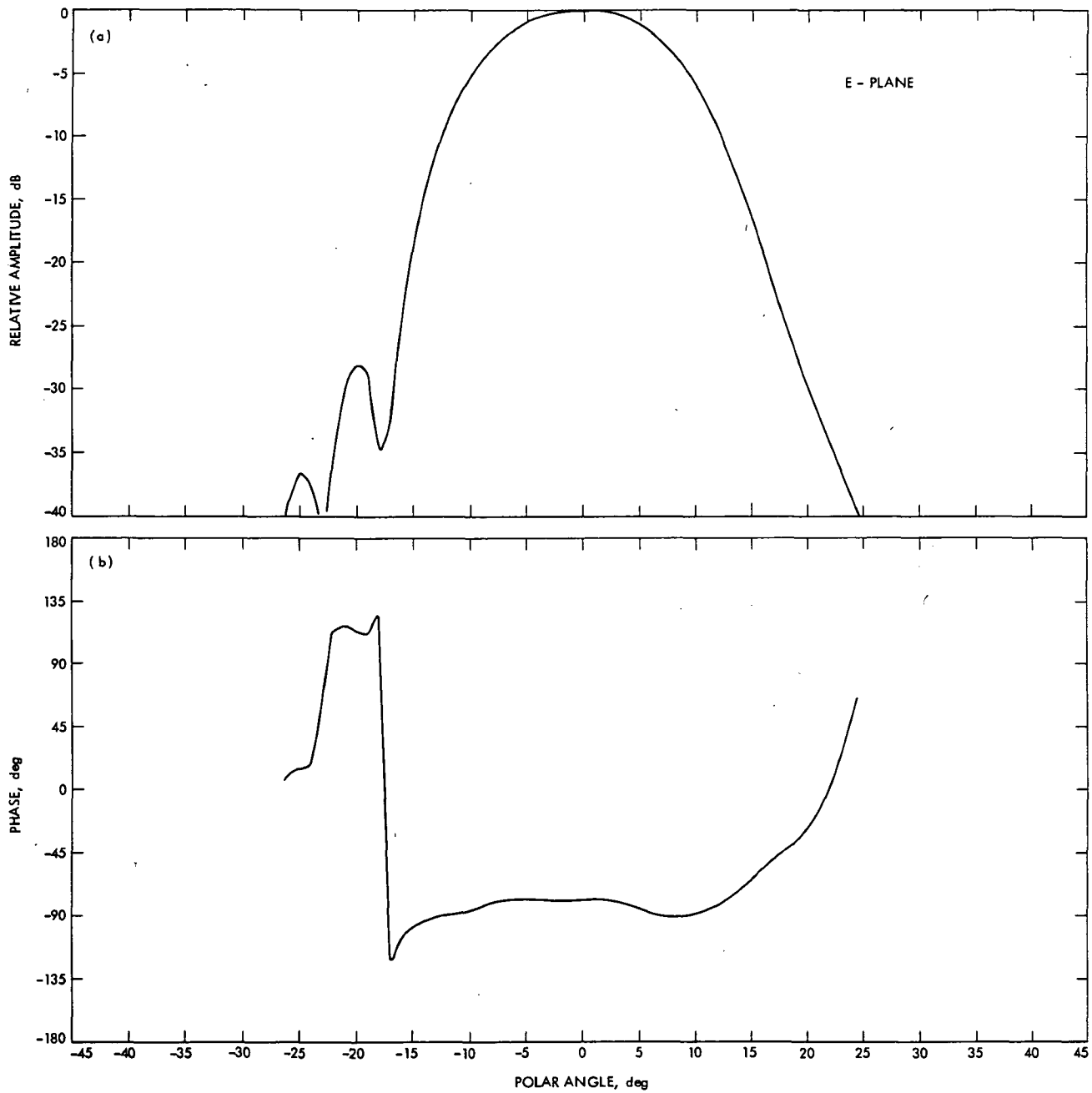


Fig. 8. S-band 34/64-m system: ellipsoid and dichroic plate (a) E-plane amplitude and (b) E-plane phase

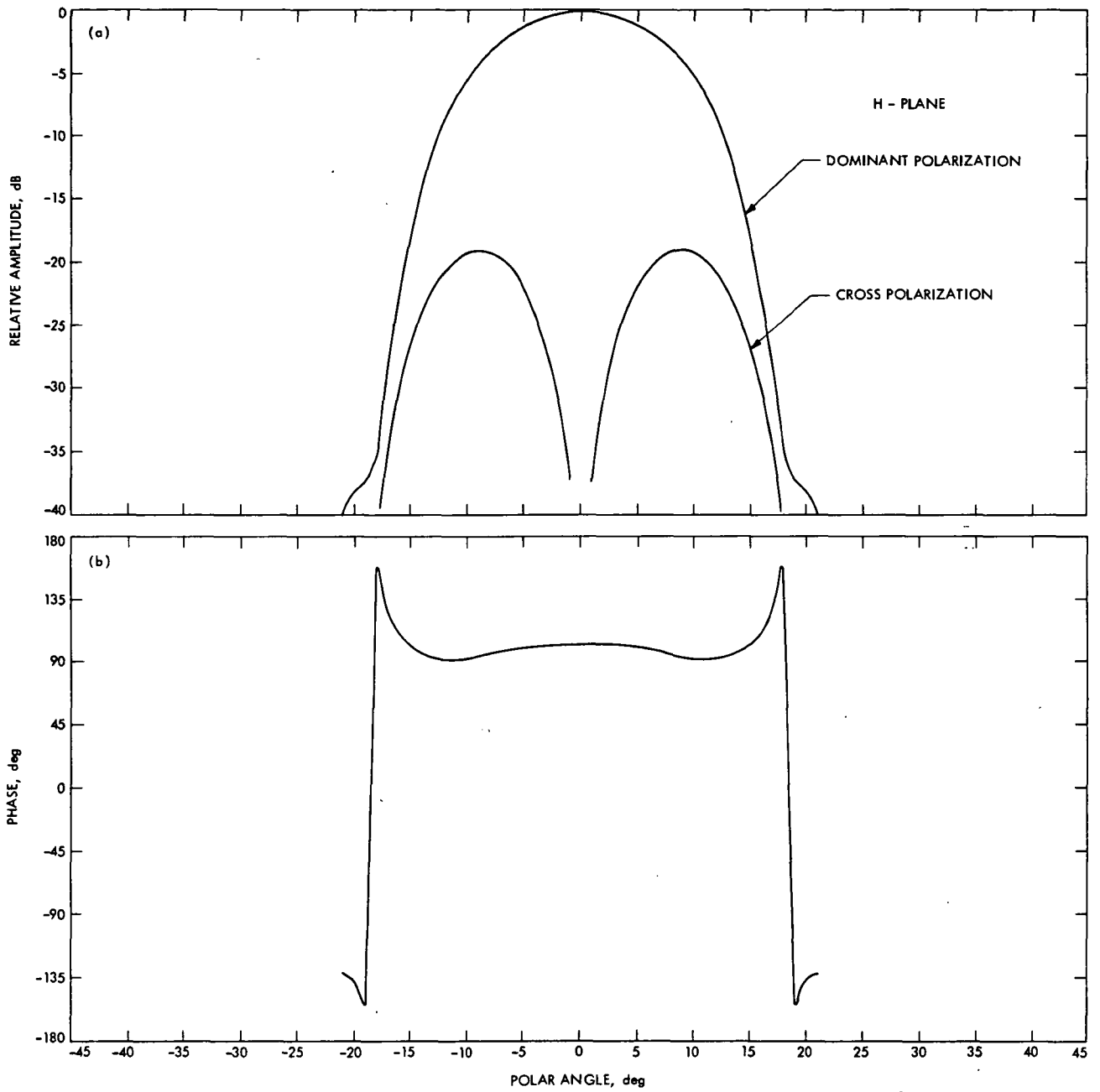


Fig. 9. S-band 34/64-m system: ellipsoid (a) H-plane amplitude and (b) H-plane phase

ORIGINAL PAGE IS
OF POOR QUALITY

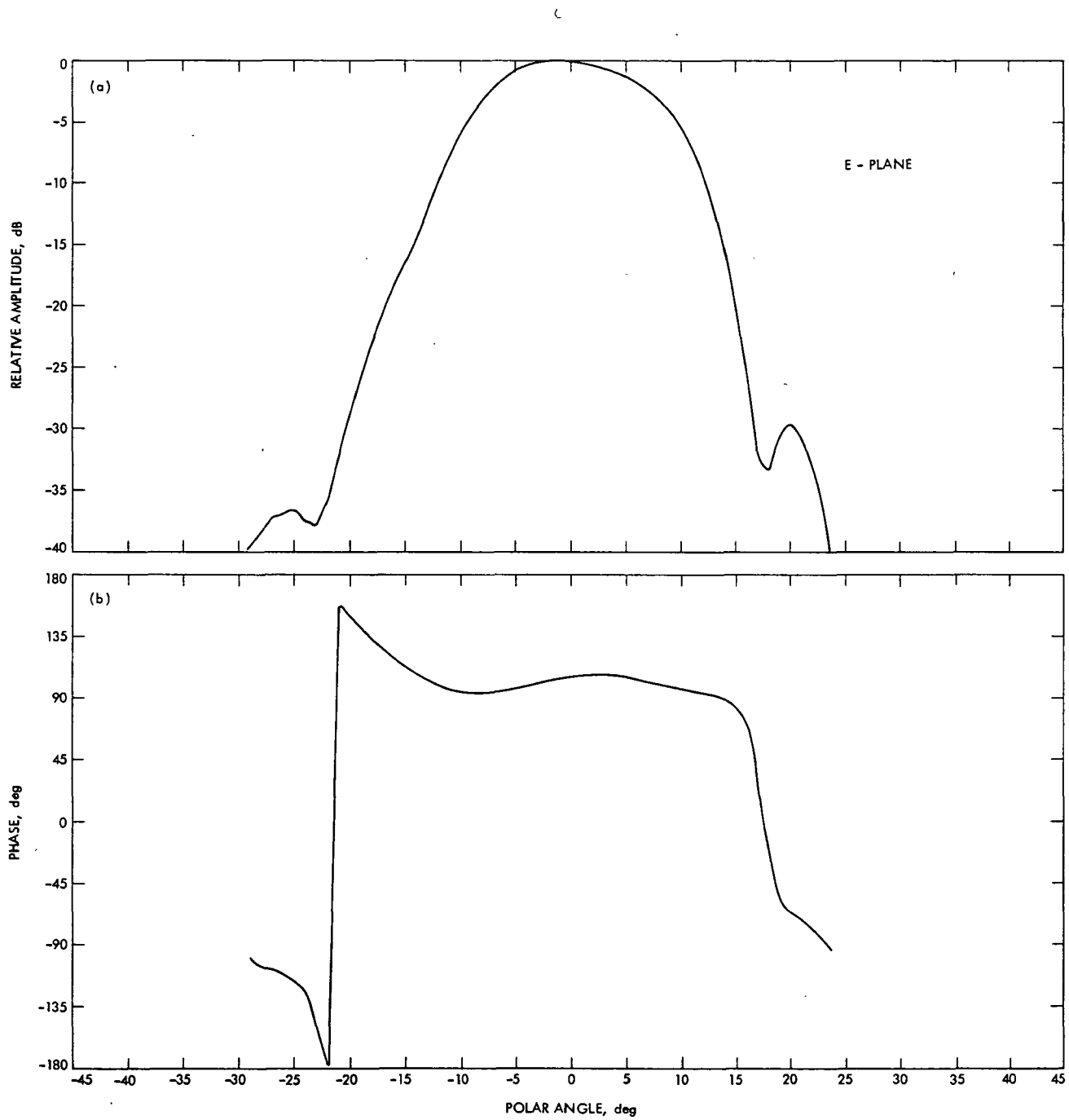


Fig. 10. S-band 34/64-m system: ellipsoid and dichroic plate (a) E-plane amplitude and (b) E-plane phase

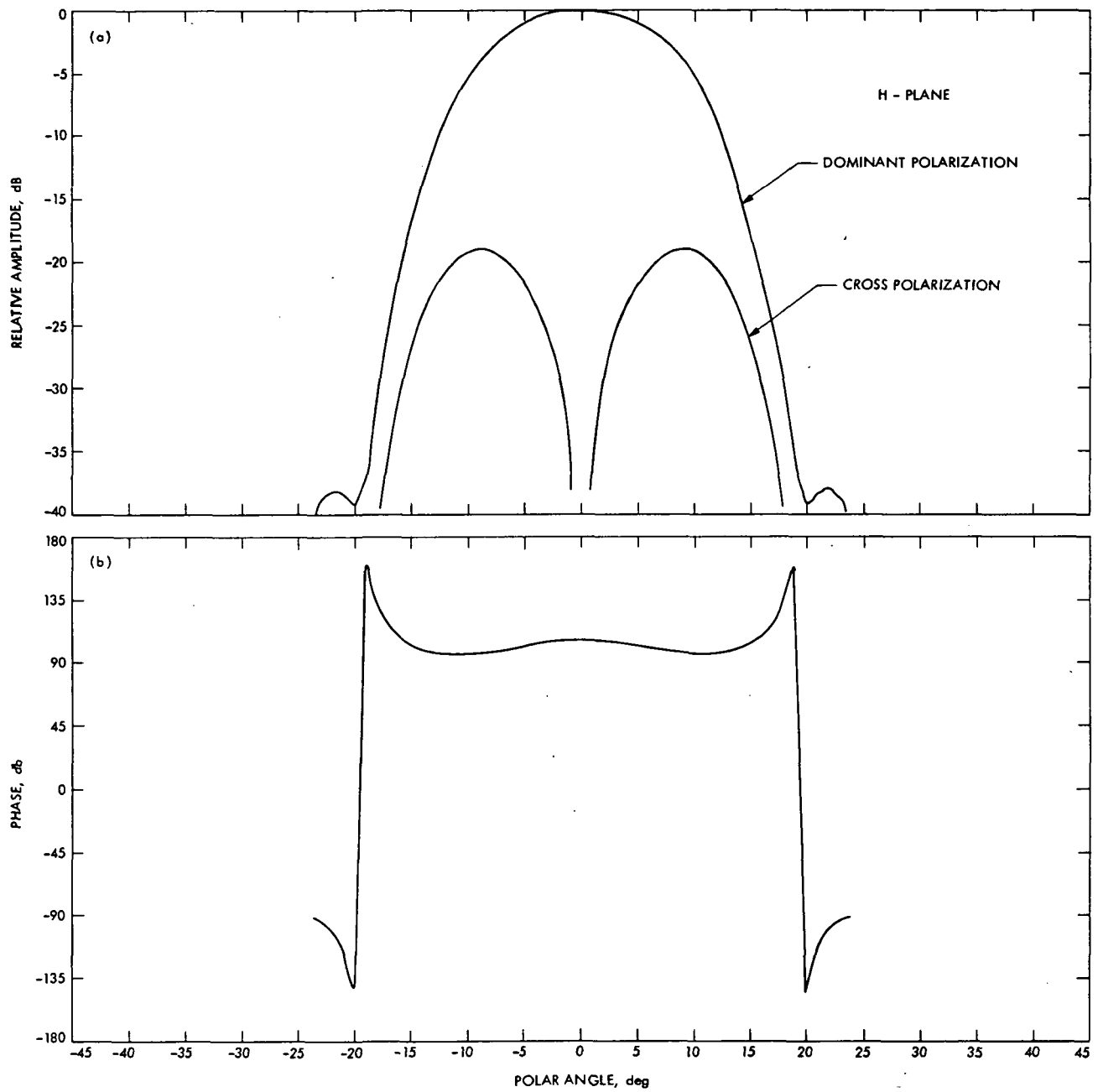


Fig. 11. S-band 34/64-m system: ellipsoid and dichroic plate (a) H-plane amplitude and (b) H-plane phase

ORIGINAL PAGE IS
OF POOR QUALITY

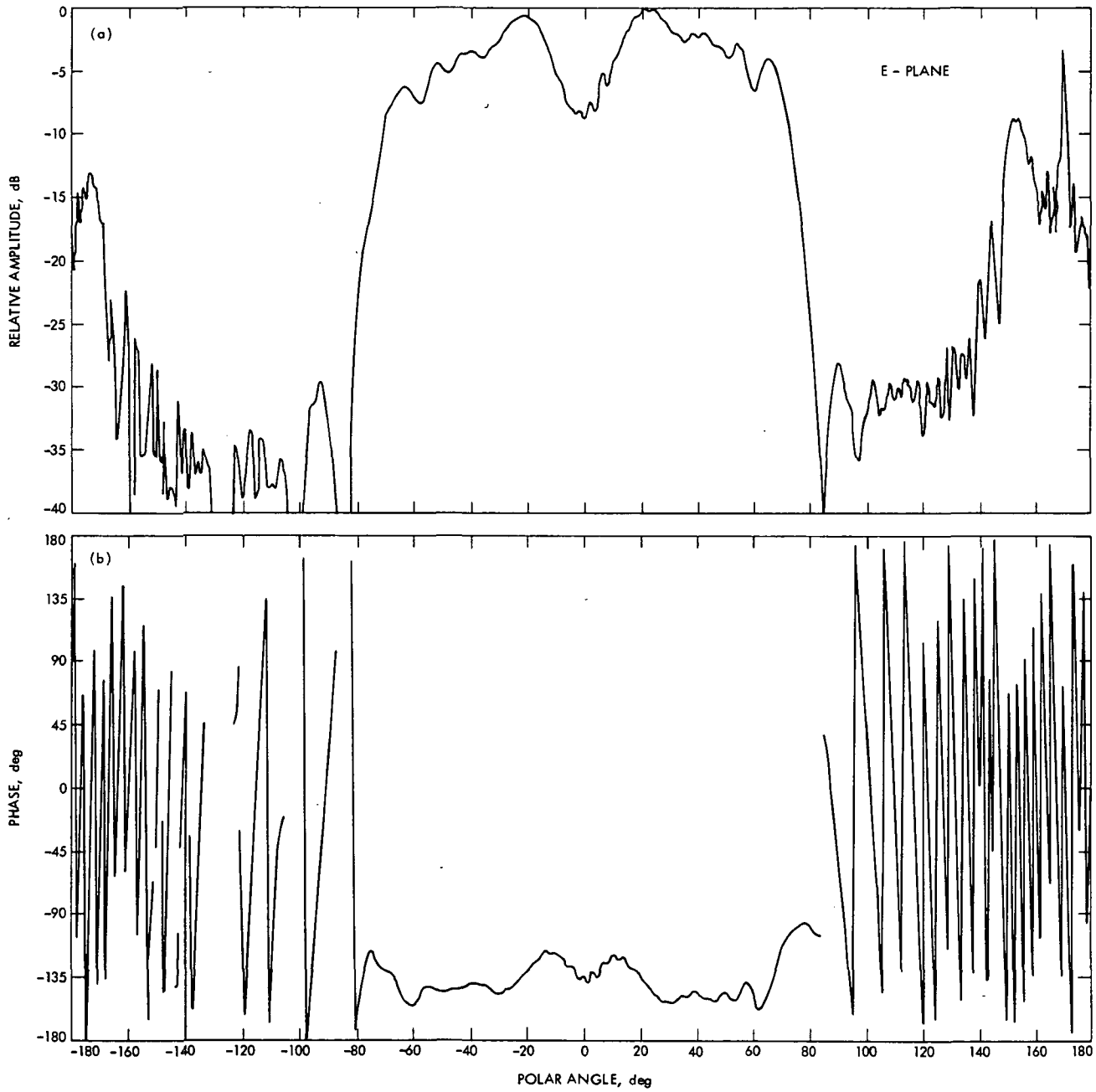


Fig. 12. S-band 34-m system: ellipsoid, dichroic plate, and subreflector (a) E-plane amplitude and (b) E-plane phase

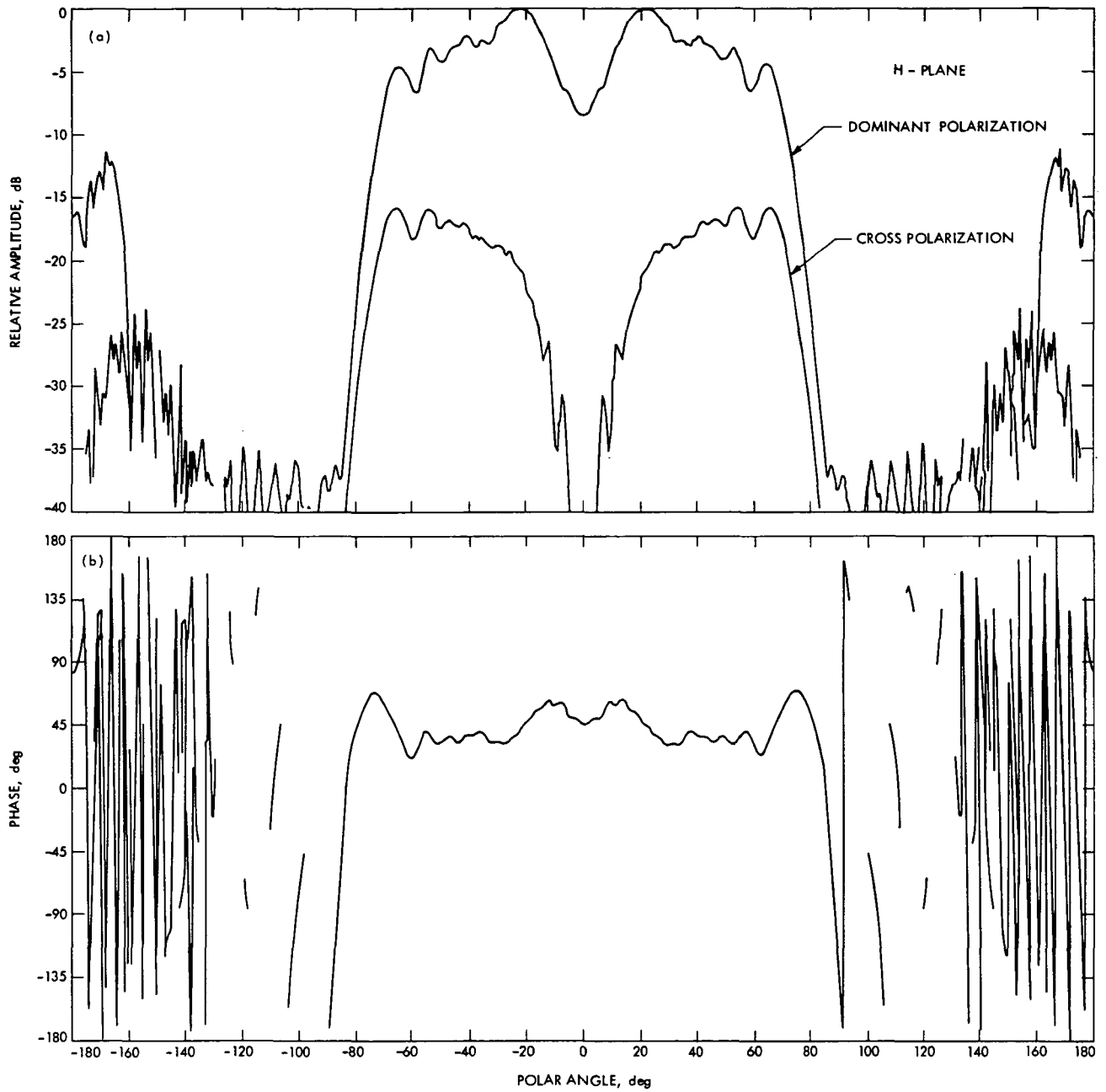


Fig. 13. S-band 34-m system: ellipsoid, dichroic plate, and subreflector (a) H-plane amplitude and (b) H-plane phase

ORIGINAL PAGE IS
OF POOR QUALITY

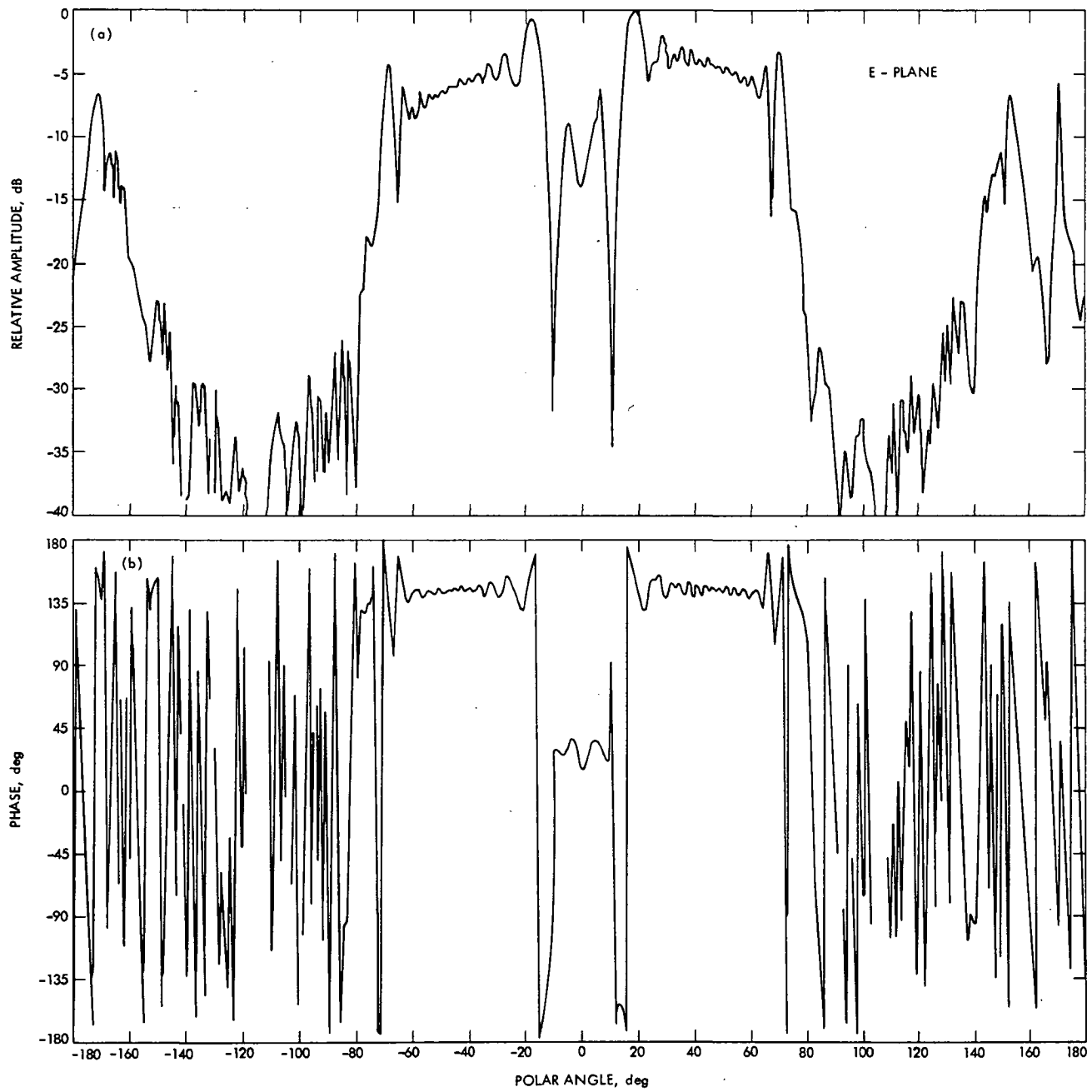


Fig. 14. X-band 34-m system: subreflector (a) E-plane amplitude and (b) E-plane phase

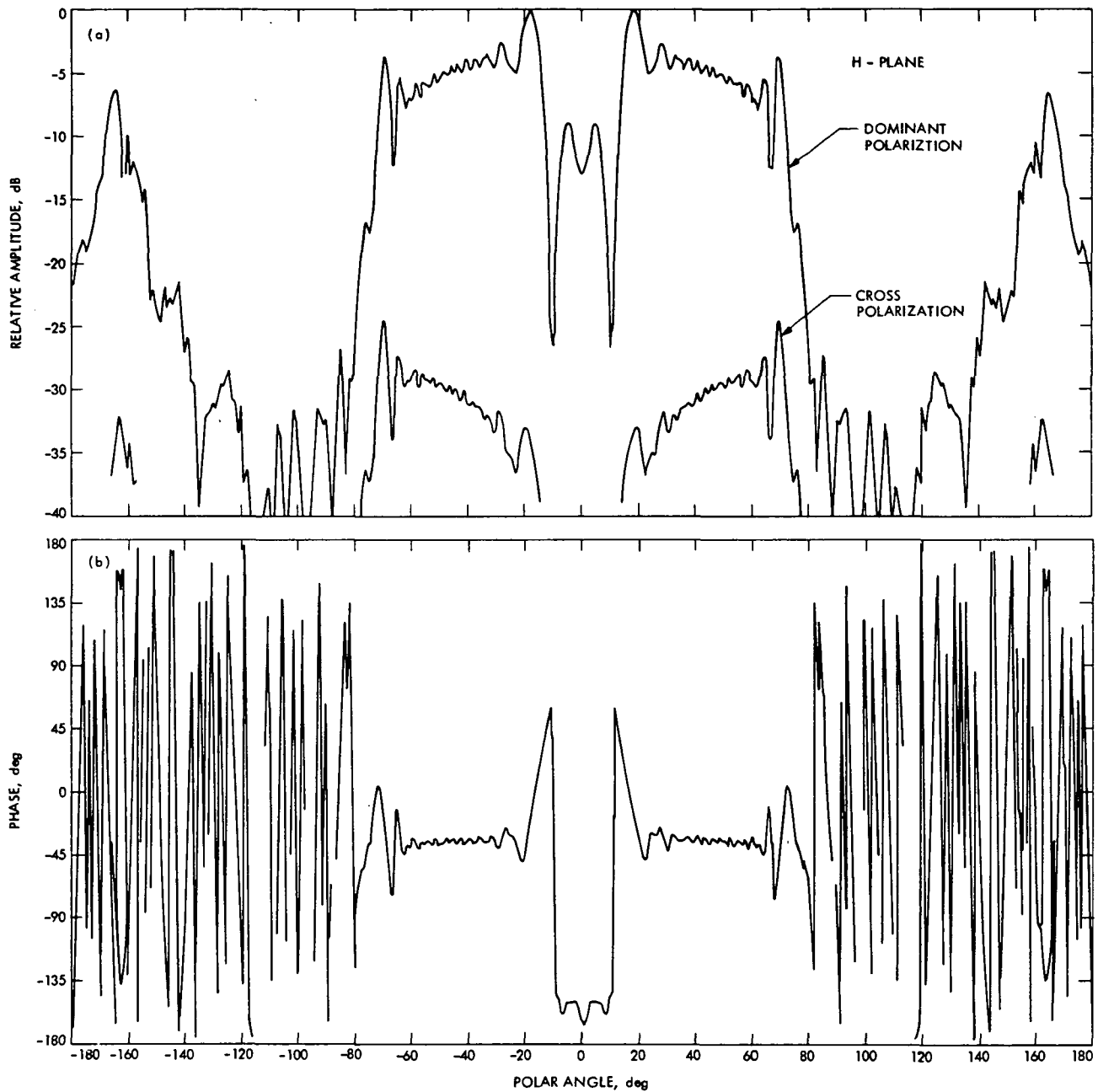


Fig. 15. X-band 34-m system: subreflector (a) H-plane amplitude and (b) H-plane phase

ORIGINAL PAGE IS
OF POOR QUALITY

N78-11150

Configuration Control and Audit Assembly

A. P. Irvine
DSN Data Systems Section

The Configuration Control and Audit Assembly (CCA) Project is responsible for the implementation of a computer-based system for acquiring, managing and distributing a subset of the Deep Space Network (DSN) Data Base of operational and management information. This system is presently scheduled for transfer to Operations July 1, 1981. Early in 1978 a demonstration will be conducted to provide data for a Life Cycle Cost (LCC) Analysis and verify design feasibility.

I. Introduction

This report presents an overview of the functions of the Configuration Control and Audit Assembly (CCA) and describes the subset of those functions that will be included in the CCA Demonstration to be performed at the Goldstone Deep Space Communications Complex (DSCC) early next year.

Its two primary objectives are:

- (1) Provide an easily accessible source of valid information to support DSN management activities.
- (2) Provide a more cost effective method of acquiring, maintaining and retrieving managerial data.

II. CCA Overview

The CCA is a unit in a network of four computers which is comprised of a CCA each at the Jet Propulsion Laboratory (JPL), Australia Deep Space Communications Complex (DSCC), Spain DSCC and Goldstone DSCC and linked together by high-speed data (HSD) lines (Fig. 1).

The CCA will perform the following major functions:

- (1) Data acquisition
- (2) Data management
- (3) Data distribution

III. Hardware

The CCA hardware configuration consists of the following: (Fig. 2)

- (1) One MODCOMP central processor, 64K word core memory
- (2) One 25.0 million byte (MB) disc
- (3) One 2.5 MB disc
- (4) One Magnetic Tape Unit
- (5) Remote Terminals
- (6) Operations Console

During the demonstration a Telemetry Process Assembly (TPA) at Compatibility Test Area (CTA) 21 will be used as a CCA. The demonstration configuration at the Goldstone DSCC will include additional peripherals (Figs. 3 and 4).

IV. Data Base

The DSN Data Base is a collection of stored operational data shared by the DSN personnel who periodically retrieve, update, add, and delete data. The DSN Data Base resides on various storage media and is maintained via various computerized and manual data management systems (Fig. 5). The CCA is concerned with those computerized portions of the data base applicable to station operations. The CCA demonstration will deal with a restricted subset of data comprising the physical equipment category (Fig. 6).

V. Demonstration Purpose

The primary purposes of the CCA Assembly demonstration are:

- (1) To provide valid cost data for a Life Cycle Cost (LCC) comparison analysis in order to establish economic feasibility.
- (2) To verify in an operational environment the functionality of a DSN distributed data base design.
- (3) To verify selected CCA Assembly requirements and the resulting technical design in an operational environment.

VI. Demonstration Software

The CCA system software supports two modes of operation:

- (1) On-line
- (2) Off-line (Fig. 3).

A. On-line

On-line operations fall into three major categories:

- (1) Transactions
- (2) Inquiries
- (3) MBASIC* applications programs

Transactions are operations performed to request changes to the data base. Those transactions that will be available to the user during the demonstration are:

- (1) Indicate equipment is to be shipped
- (2) Enter a new equipment Engineering Change Order (ECO)
- (3) Up-dating ECO status
- (4) Indicate equipment has failed
- (5) Assign a new control number
- (6) Receive shipment
- (7) Enter a new equipment design
- (8) Assign nomenclature to a drawing number

Inquiries are operations that retrieve information from the data base. Those transactions that will be available to the user during the demonstration are:

- (1) Request location of a possible spare
- (2) Request ECO status by drawing number
- (3) Request ECO status by DSN control number

An MBASIC* processor is being developed to allow application programs to access the data base in an information retrieval mode.

B. Off-line

Off-line operations fall into two major categories:

- (1) Update
- (2) High Speed Data (HSD)

The update portion of the off-line software takes care of all manipulations to the data base. Transactions from the transaction file are supplied to the data base after the "old" base is achieved. A new master data base is now formed as the result of transactions (Fig. 7).

The HSD portion of the off-line software handles transmissions between the CCA at the Goldstone DSCC and CTA-21. Any file selected by the operator may be transmitted in either direction.

*A trademark of the California Institute of Technology.

VII. Analysis

During the demonstration a series of monthly management reports will be generated by DSN Facility Support and Goldstone Operations. The monthly report will cover the following:

- (1) Overview of demonstration operations and progress
- (2) Significant problems
- (3) Data base comparison sampling results

The Life Cycle Cost (LCC) Analysis will be based primarily on this data.

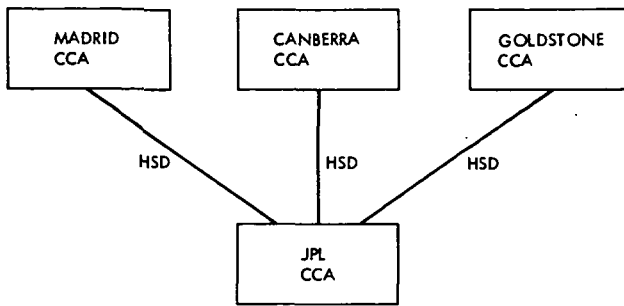


Fig. 1. Overview of the CCA configuration

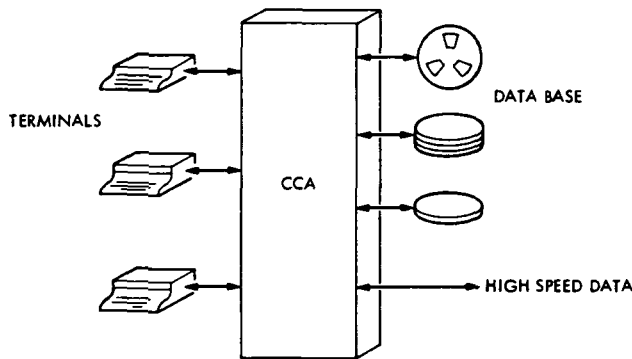


Fig. 2. CCA hardware configuration

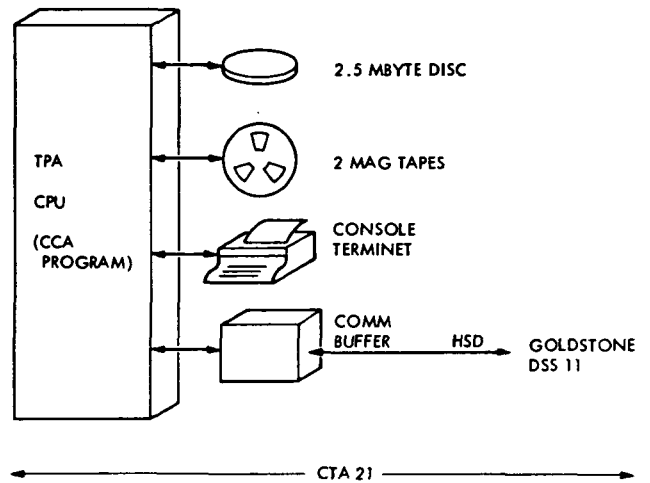


Fig. 3. CCA demonstration hardware at CTA 21

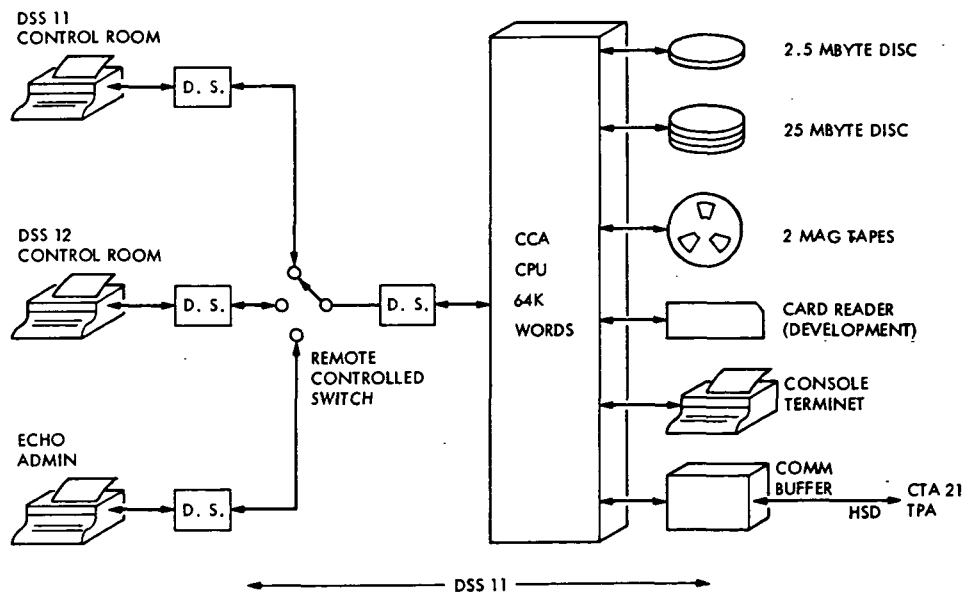


Fig. 4. CCA demonstration hardware at Goldstone

ORIGINAL PAGE IS
OF POOR QUALITY

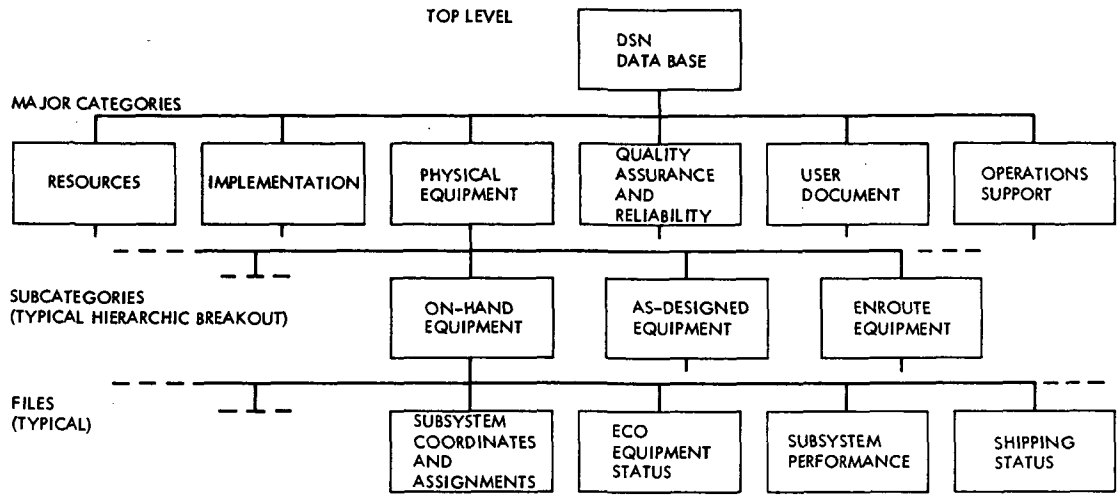


Fig. 5. The DSN data base

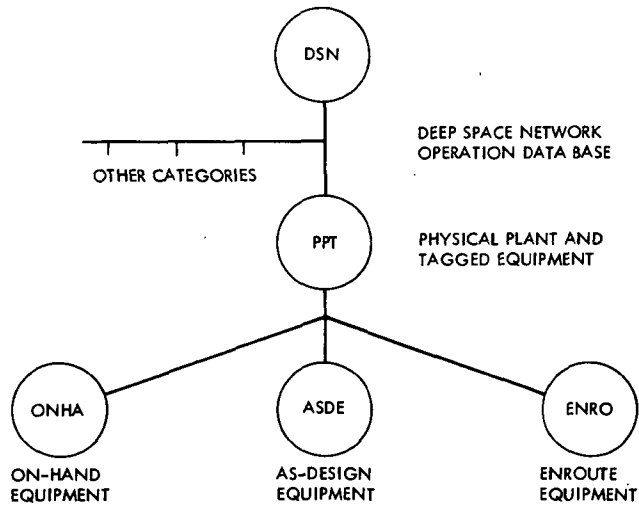


Fig. 6. The CCA demonstration data base

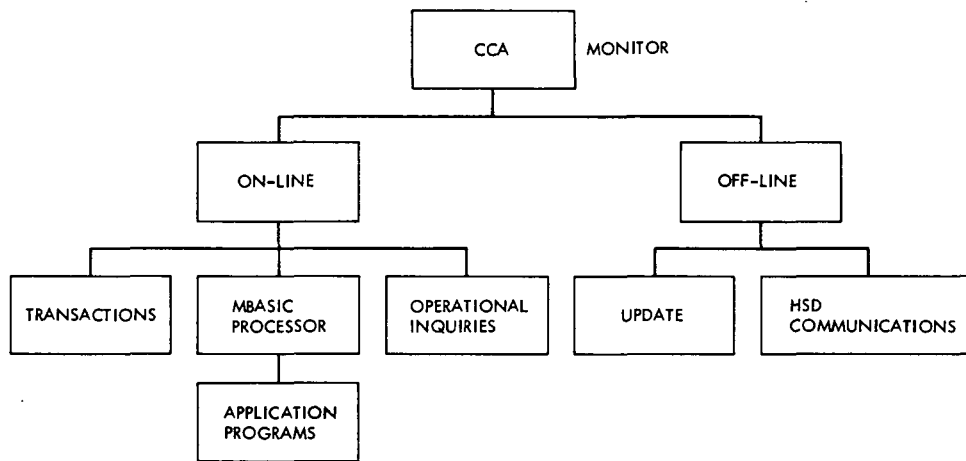


Fig. 7. Overview of CCA demonstration software

ORIGINAL PAGE IS
OF POOR QUALITY

0-3

N78-11151

Microcomputer Central Processing Unit Module

O. B. Parham

Radio Frequency and Microwave Subsystem Section

A new multipurpose microcomputer central processing unit (CPU) module is described and illustrated in this article. The module is designed to stand alone and also be upward expandable.

I. Introduction

A multipurpose stand-alone microcomputer in one standard DSN logic packaging subchassis (Ref. 1) has been designed for use in the Spectral Signal Indicator (SSI), the Noise Adding Radiometer (NAR), and the Precision Signal Power Measurement (PSPM) projects.

II. Configuration and Implementation

The three major parts of the microcomputer module are the central processing unit (CPU), the memory, and the input/output (I/O) as shown in the photo (Fig. 1). At the heart of the CPU is an 8080 microprocessor (Ref. 2), which with its support circuitry runs at a 2-MHz clock speed. Additional CPU implementations are a 4-channel direct memory access (DMA) controller for high-speed data transfer between the memory and I/O sections, and an 8-level priority interrupt controller for handling asynchronous events on an interrupt basis. (For a simplified block diagram of the module, see Fig. 2.)

The memory section is comprised of 12 K bytes of erasable programmable read-only memory (EPROM) for program storage and 4 K bytes of random access memory (RAM) for data

storage. To insure the multipurpose nature of the module, the EPROM section of memory incorporates the features necessary to allow the EPROMs to be programmed without being removed from the module. The RAM section of memory is implemented with parity error detection. Thus, should a memory error occur in the RAM area, an interrupt is generated to the CPU, notifying it that the memory has made an error. Additionally, when an error is detected, the upper half of the address bus is latched and can tell the processor which page of memory caused the error.

Components of the module's input/output structure are:

Two RS-232 ports with software controlled transmission rate, or baud rate

Nine parallel 8-bit ports for general-purpose usage

One IEEE-488 interface for interfacing to commercial equipment

One DSN interface for communication with station controllers

In addition to the data I/O, a three-section software controlled counter timer has been implemented. The first two sections of



the timer are cascaded and connected to the interrupt controller to generate real-time interrupts, with the interrupt periods programmable from 2 μ s to 2147 s. The third section is available for external use as a pulse generator or event counter.

The priority interrupt controller assigns the highest internal priority interrupt (Level 0) to memory parity error detection. The other internal priority interrupts are:

Level 1, to the real-time counter

Level 2, to the DSN interface, the IEEE-488 interface, and the DMA controller

Level 3, to the general-purpose I/O and the RS-232 ports

Four lower priority interrupts (levels 4, 5, 6, 7) are available externally. Internal priority interrupt levels shared between various ports on the same levels are individually maskable by software. The external interrupts are maskable in the interrupt controller.

III. System Expansion

Many applications require more memory and I/O than is available on the microcomputer module. To overcome this, the CPU bus has been buffered and brought out of the module. This allows simple, easy expansion of the system to the full capability of the 8080 CPU. A local memory disable feature has also been incorporated to allow for memory paging. This extends the memory capability beyond the normal 65 K limit of the 8080.

IV. Summary

A microcomputer module has been designed for use of the NAR, PSPM, and SSI projects. The multipurpose nature of the module will allow it alone, or with expansion modules, to meet the needs of many future DSN requirements.

References

1. Slaughter, D. W., "Standard High-Reliability Integrated Circuit Logic Packaging," in *The Deep Space Network Progress Report 42-39*, pp. 130-156, Jet Propulsion Laboratory Pasadena, Calif., March and April, 1977.
2. 8080 Microcomputer Systems User's Manual, Intel Corp., Santa Clara, Calif., July 1975.

ORIGINAL PAGE IS
OF POOR QUALITY

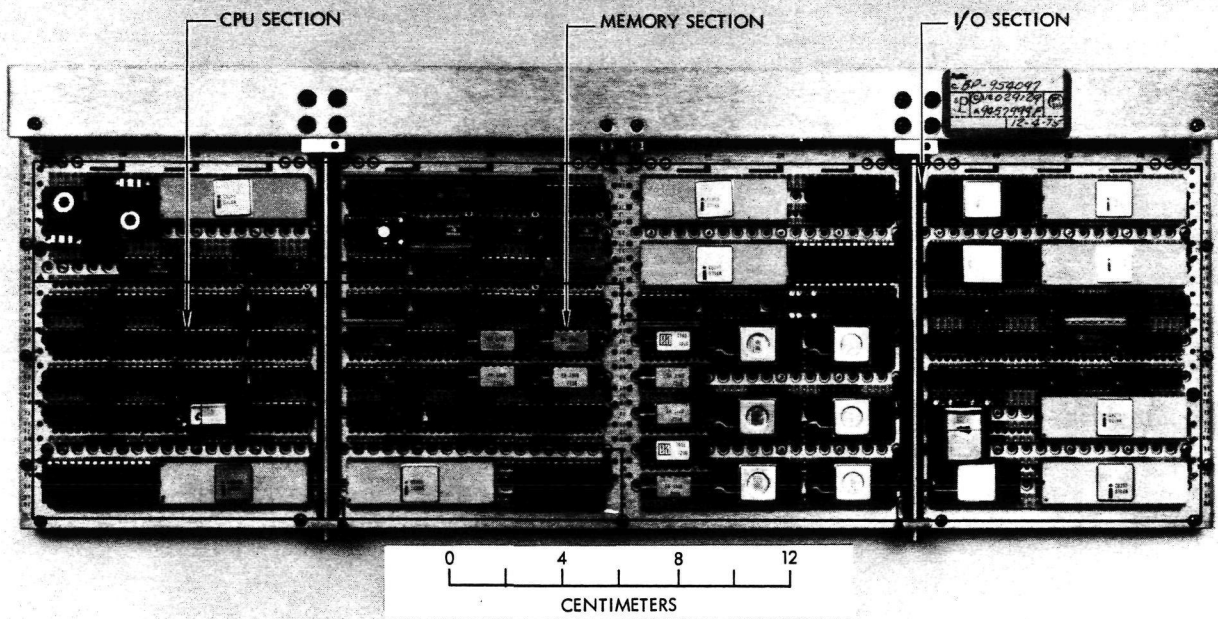


Fig. 1. Microcomputer CPU module

ORIGINAL PAGE IS
OF POOR QUALITY

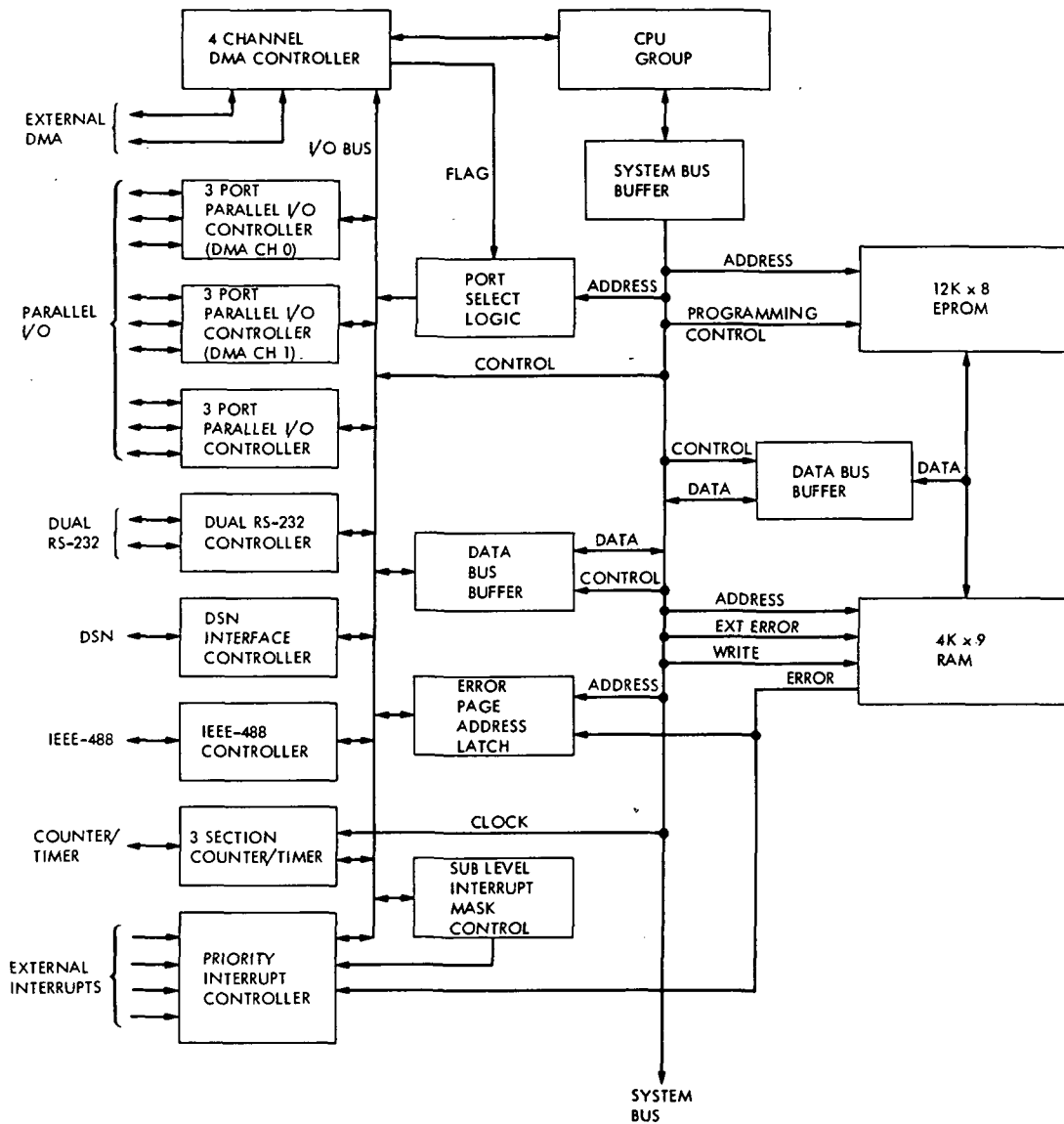


Fig. 2. Simplified block diagram of microcomputer CPU module

ORIGINAL PAGE IS
OF POOR QUALITY

N78-11152

Fast Algorithms for Computing Mersenne-Prime Number-Theoretic Transforms¹

I. S. Reed
University of Southern California

T. K. Truong
TDA Engineering Office

It is shown that Winograd's algorithm can be used to compute an integer transform over $GF(q)$, where q is a Mersenne prime. This new algorithm requires fewer multiplications than the conventional fast Fourier transform (FFT). This transform over $GF(q)$ can be implemented readily on a digital computer. This fact makes it possible to more easily encode and decode BCH and RS codes.

I. Introduction

Several authors (Ref. 1 through 12) have shown that transforms over finite fields or rings can be used to compute numerical convolutions without round-off error. Recently, Winograd (Ref. 13) developed a new class of algorithms for computing the conventional discrete Fourier transform (DFT). This new algorithm requires substantially fewer multiplications than the conventional FFT algorithm.

In this paper, a type of Winograd algorithm is employed to evaluate the transform over $GF(q)$, where $q = 2^p - 1$ is a Mersenne prime. This transform is comparable in speed with that given by Winograd (Ref. 13 and 14).

Recently, the authors (Ref. 15 and 16) proposed that transforms over $GF(F_n)$, where $F_n = 2^{2^n} + 1$ for $n = 1, 2, 3, 4$ is a Fermat prime, can be used to define RS codes and to improve the decoding efficiency of these codes. Therefore, an FFT over $GF(q)$ can be used to decode RS codes.

In order to use the methods of Winograd for computing the transform over $GF(q)$, a new method for computing every factorization of the polynomial $u^{p-1} - 1$ over $GF(q)$ is developed. Finally, it is shown that continued fractions can be used instead of the usual Euclid's algorithm to find the required inverse element of the polynomial over the finite field $GF(q)$.

¹This work was supported in part by the U.S. Air Force Office of Scientific Research under Grant AFOSR-75-2798.

II. Transforms Over $GF(q)$

Let $GF(q)$ be the finite field of residue classes modulo q , where $q = 2^p - 1$ is a Mersenne prime for $p = 2, 3, 5, 7, 13, 17, 19, 31, 61, \dots$. Also let d be an integer that divides $q - 1$. Finally let the element $\gamma \in GF(q)$ generate the cyclic subgroup of d elements, $G_d = (\gamma, \gamma^2, \dots, \gamma^{d-1}, \gamma^d = 1)$ in the multiplicative group of $GF(q)$.

The transform over G_d is

$$A_j = \sum_{i=0}^{d-1} a_i \gamma^{ij} \text{ for } 0 \leq j \leq d-1 \quad (1)$$

where $a_i \in GF(q)$ for $0 \leq i \leq d-1$

By Fermat's theorem, $2^p \equiv 2 \pmod{p}$. This implies $p|t$, where $t = q - 1 = 2^p - 2$. Also since $2^p - 2 \equiv 0 \pmod{3}$ or 2 , $t = 2^p - 2$ has the factors $2, 3$, and p . The factorizations of the different numbers $t = 2^p - 2$ for $p = 13, 17, 31, 61$ are shown in Table 1. A multidimensional technique will be developed herein to calculate the transform defined in (1).

To perform the transform over $GF(q)$ defined in (1), it is necessary to find primitive elements in the d -element cyclic subgroup G_d in $GF(q)$. To do this, by (Ref. 5), it is shown that 3 is quadratic nonresidue mod q . Thus $3(2^{p-2})/2 \equiv -1 \pmod{q}$. Hence by the same procedure used in the proof of Theorem 1 in Ref. 5, $\alpha = 3$ is a primitive element in $GF(q)$. Suppose $d = d_1 \cdot d_2 \cdot \dots \cdot d_{r-1} \cdot d_r$ where $d_1 = 2, d_2 = 3, d_r = p$, and $d_i = 5, 7, 9, 11$, and 13 for $i = 3, 4, \dots, r-1$. The generator of G_d , a multiplicative subgroup of order d , is evidently $\gamma = \alpha^{(q-1)/d}$ where α is a primitive element of $GF(q)$. Also, γ^j is a primitive element in G_d since $(j, d) = 1$.

Now γ satisfies,

$$\gamma^d = 1 \pmod{q} \quad (2)$$

But also

$$2^p = 1 \pmod{q} \quad (3)$$

Thus, combining (2) and (3), a generator γ of G_d must exist that satisfies

$$\gamma^{d/p} = 2 \pmod{q} \quad (4)$$

A computer program can be used to find a primitive element γ of G_d that satisfies (4). By Th. 1, Ref.5, γ satisfied

$$\gamma^{d/2} = -1 \quad (5)$$

From (4), (5), we observe that integer multiplications by γ^{d/d_i} or its powers for $d_i = 2$ or p can be accomplished simply by circular shifts instead of multiplications. Hence, a d_i -point DFT for $d_i = 2$ or p can be evaluated without integer multiplications. It will be shown next that a d_i -point DFT for $d_i = 3, 5, 9, 11$, and 13 can be computed by using the Winograd algorithm.

ORIGINAL PAGE IS
OF POOR QUALITY

III. Mathematical Preliminaries

In the next section the Chinese Remainder theorem for polynomials will be employed to compute fast transforms over $GF(q)$ of a small sequence. This well-known theorem is stated as follows without proof (see Ref. 17):

Theorem 1 (The Chinese Remainder theorem for polynomial): If $m_1(x), m_2(x), \dots, m_k(x)$ are polynomials which are relatively prime in pairs, then the system of congruences $x(u) \equiv g_i(u) \pmod{m_i(u)^{e_i}}$ for $i = 1, 2, \dots, k$ has a unique solution $x(u)$ given by

$$x(u) = \sum_{i=1}^k g_i(u) M_i(u) N_i(u) \quad (6a)$$

where

$$\begin{aligned} m(u) &= m_1(u)^{e_1} m_2(u)^{e_2} \dots m_k(u)^{e_k} \\ &= m_1(u)^{e_1} M_1(u) = m_2(u)^{e_2} M_2(u) = \dots = m_k(u)^{e_k} M_k(u) \end{aligned}$$

and $N_i(x)$ uniquely satisfies (modulo $m_i(u)^{e_i}$) the congruence

$$M_i(u) N_i(u) \equiv 1 \pmod{m_i(u)^{e_i}} \quad (6b)$$

To compute the inverse element of $M(u)$, i.e., $N(u)$ required in (6b), let $S(u) = M_i(u)/m_i(u)$. Then, using a procedure precisely similar to that used for a rational element $S = a/q$, where $a \in GF(q)$, described in Appendix A, it is possible to use continued fractions to develop a finite sequence of rational approximations to $S(u)$. The recursive formula for the convergents is given by

$$S_k(u) = \frac{a_k(u)p_{k-1}(u) - p_{k-2}(u)}{a_k(u)q_{k-1}(u) + q_{k-2}(u)} = \frac{p_k(u)}{q_k(u)} \quad (7)$$

where $p_{-1}(u) = 1, q_{-1}(u) = 0, p_0(u) = a_0(u)$, and $q_0(u) = 1$. The partial quotients $a_k(u)$ in (7) can be computed recursively by the following formula:

$$r_{k-2}(u) = a_k(u)r_{k-1}(u) + r_k(u), \deg r_k(u) < \deg r_{k-1}(u) \text{ for } k = 1, 2, \dots, n-1 \quad (8)$$

where the initial conditions are $r_{-1}(u) = m_i(u), r_{-2} = M_i(u)$, and $r_{n-2}(u) = r_{n-1}(u)a_n(u)$.

By applying Euclid's algorithm to the polynomial $S(u)$ over $GF(q)$, we observe that $S_k(u) = p_k(u)/q_k(u)$ will terminate with $S_n(u) = M_i(u)/m_i(u)$ when $r_n(u) = 0$. By the same procedure used in the derivation of Eq. (6A), we obtain

$$M_i(u)q_{n-1}(u) - m_i(u)p_{n-1}(u) = (-1)^{n+1} \quad (9)$$

There are two cases to consider:

Case I: If n is odd, then

$$M_i(u)q_{n-1}(u) + m_i(u)(-p_{n-1}(u)) = 1 \quad (10)$$

It follows that $N_i(u) = q_{n-1}(u)$ and $n_i(u) = -p_{n-1}(u)$ are solutions of $M(u)N_i(u) \equiv 1 \pmod{m_i(u)}$ and $m(u)n_i(u) \equiv 1 \pmod{M_i(u)}$, respectively.

Case II: If n is even, then

$$M_i(u)(-q_{n-1}(u)) + m_i(u)(p_{n-1}(u)) = 1 \quad (11)$$

Thus, $N_i(u) = -q_{n-1}(u)$ and $n_i(u) = p_{n-1}(u)$ are solutions of $N_i(u)M_i(u) \equiv 1 \pmod{m_i(u)}$ and $m_i(u)n_i(u) \equiv 1 \pmod{M_i(u)}$ respectively.

From (9), we see that it is necessary to compute the inverse element of a in $GF(q)$. To do this, let $S = a/q$. This inverse element is given by (A-7) in Appendix A.

The remainder of this section is based on ideas due to Winograd (Ref. 13). Let $X(u) = X_0 + X_1u + X_2u^2 + \dots + X_nu^n$ and $Y(u) = Y_0 + Y_1u + Y_2u^2 + \dots + Y_nu^n$ be two polynomials where $X_i, Y_i \in GF(q)$. It is well known that the linear convolution of $X(u)$ and $Y(u)$ is the set of coefficients of the product of $X(u)$ and $Y(u)$, i.e., $T(u) = X(u) \cdot Y(u)$. By (Ref. 13), the number of multiplications required to compute the coefficients of $T(u)$ can be obtained by using the Chinese Remainder theorem for polynomials over $GF(q)$.

To show this, let us choose $m + n + 1$ distinct scalars, i.e., $\alpha_0, \alpha_1, \dots, \alpha_{m+n}$. Then $T(u)$ with degree $n + m$ is equal to

$$T(u) = X(u) \cdot Y(u) \pmod{(u - \alpha_0) \cdot (u - \alpha_1) \cdot \dots \cdot (u - \alpha_{m+n})} \quad (12a)$$

or

$$T(u) = X(u) \cdot Y(u) \pmod{\prod_{i=0}^{m+n-1} (u - \alpha_i) + X_n Y_n \prod_{i=0}^{m+n-1} (u - \alpha_i)} \quad (12b)$$

Since $(u - \alpha_i)$ for $i = 0, 1, \dots, m + n$ are relatively prime in pairs, then by theorem 1, the system of congruences,

$$T_k(u) = X(\alpha_k) \cdot Y(\alpha_k) \equiv T(u) \pmod{(u - \alpha_k)} \text{ for } k = 0, 1, 2, \dots, n + m$$

has a unique solution $T(u)$ given by

$$T(u) = \sum_{k=0}^{m+n} T_k(u)M_k(u)N_k(u)$$

ORIGINAL PAGE IS
OF POOR QUALITY

where

$$\begin{aligned}
 m(u) &= m_1(u)m_2(u), \dots, m_k(u) \\
 &= (u - \alpha_0)(u - \alpha_1) \dots (u - \alpha_{n+m}) \\
 &= m_0(u)M_0(u) = m_1(u)M_1(u) = \dots = m_{m+n}(u)M_{m+n}(u)
 \end{aligned}$$

and $N_k(u)$ uniquely satisfies (module $m_k(u)$) the congruences

$$M_k(u)N_k(u) \equiv 1 \pmod{m_k(u)} \text{ for } k = 0, 1, 2, \dots, m+n$$

It can be shown that $T(u)$ given by (12a) can be reconstructed by

$$T(u) = \sum_{k=0}^{m+n} \left[\frac{\prod_{\substack{j=0 \\ j \neq k}}^{m+n} (u - \alpha_j)}{\prod_{\substack{j=0 \\ j \neq k}}^{m+n} (\alpha_k - \alpha_j)} \right] X(\alpha_k) \cdot Y(\alpha_k) \quad (13a)$$

If for example, one chooses $\alpha_k = \pm 2^n$ for $n \geq 0$, then each $T_k(u) = X(\alpha_k) \cdot Y(\alpha_k)$ can be computed with one multiply. Similarly, $T(u)$ given by (12b) is given by

$$T(u) = \sum_{k=0}^{m+n-1} \left[\frac{\prod_{\substack{j=0 \\ j \neq k}}^{m+n-1} (u - \alpha_j)}{\prod_{\substack{j=0 \\ j \neq k}}^{m+n-1} (\alpha_k - \alpha_j)} \right] X(\alpha_k) \cdot Y(\alpha_k) + X_n Y_n \sum_{i=0}^{m+n-1} (u - \alpha_i) \quad (13b)$$

The cyclic convolution of $(X_0, X_1, \dots, X_{n-1})$ and $(Y_0, Y_1, \dots, Y_{n-1})$ can be expressed as

$$\begin{pmatrix} \Psi_0 \\ \Psi_1 \\ \vdots \\ \Psi_{n-1} \end{pmatrix} = \begin{pmatrix} X_0 & X_1 & \dots & X_{n-2} & X_{n-1} \\ X_1 & X_2 & \dots & X_{n-1} & X_0 \\ \vdots & \vdots & & \vdots & \vdots \\ X_{n-1} & X_0 & \dots & X_{n-3} & X_{n-2} \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_{n-1} \end{pmatrix} \quad (13c)$$

As noted by Winograd, the cyclic $m \times n$ matrix in (13c) can be regarded as a "multiplication table" for the groups Z_n of the integers module n . If $n = n_1 \cdot n_2$, where n_1 and n_2 are relatively prime, then by the Chinese Remainder theorem (Theorem 1 above for the integers) Z_n is isomorphic to the direct product $Z_{n_1} \times Z_{n_2}$ of groups Z_{n_1} and Z_{n_2} . This fact is used to prove that a permutation of the rows and columns of the cyclic matrix in (13c) exists so that the resulting matrix is partitioned into blocks of $n_2 \times n_2$ cyclic matrices and so that the blocks form a $n_1 \times n_1$ cyclic matrix. This is proved in detail in Theorem 1 of Appendix B.

To illustrate the above, let n be $6 = 2 \cdot 3$. Since 2 and 3 are relatively prime, by the Chinese Remainder theorem an isomorphism

$$k \leftrightarrow (k_1, k_2)$$

exists between an integer k module 6 and the pair of integers k_1 and k_2 module 2 and 3, respectively, from the relationship,

$$k \equiv k_1 3 + k_2 4 \pmod{6}$$

Now suppose one has the cyclic convolution,

$$\begin{pmatrix} \Psi_0 \\ \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \\ \Psi_5 \end{pmatrix} = \begin{pmatrix} X_0 & X_1 & X_2 & X_3 & X_4 & X_5 \\ X_1 & X_2 & X_3 & X_4 & X_5 & X_0 \\ X_2 & X_3 & X_4 & X_5 & X_0 & X_1 \\ X_3 & X_4 & X_5 & X_0 & X_1 & X_2 \\ X_4 & X_5 & X_0 & X_1 & X_2 & X_3 \\ X_5 & X_0 & X_1 & X_2 & X_3 & X_4 \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{pmatrix} \quad (13d)$$

By Theorem 1 in Appendix B, there exists a permutation defined by

$$\begin{aligned} \pi = \alpha^{-1} \beta &= \begin{pmatrix} (0,0) & (0,1) & (0,2) & (1,0) & (1,1) & (1,2) \\ 0 & 4 & 2 & 3 & 1 & 5 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ (0,0) & (0,1) & (0,2) & (1,0) & (1,1) & (1,2) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 4 & 2 & 3 & 1 & 5 \end{pmatrix} \end{aligned}$$

of the rows and columns so that the above cyclic matrix can be partitioned into blocks of 3×3 cyclic matrices, such that the blocks form a 2×2 cyclic matrix. In other words, let the variable $Y_k = Y_{(k_1, k_2)}$, $X_k \equiv X_{(k_1, k_2)}$ and $\Psi_k \equiv \Psi_{(k_1, k_2)}$ be rearranged in such a manner that the first component k_1 of the index pair (k_1, k_2) is set to 0 and component k_2 is in ascending order, and that secondly component k_1 is set to 1 and component k_2 is in ascending order. The variable $X_{(k_1, k_2)}$ for (13c) are then rearranged in the order:

$$X_{(0,0)}, X_{(0,1)}, X_{(0,2)}, X_{(1,0)}, X_{(1,1)}, X_{(1,2)}$$

or in the variables X_k are in the order:

$$X_0, X_4, X_2, X_3, X_1, X_5$$

If such a rearrangement is made on variables Y_k , X_k , and Ψ_k , respectively, the cyclic convolution (13d) has the form,

$$\begin{pmatrix} \Psi_0 \\ \Psi_4 \\ \Psi_2 \\ \Psi_3 \\ \Psi_1 \\ \Psi_5 \end{pmatrix} = \begin{pmatrix} X_0 & X_4 & X_2 & X_3 & X_1 & X_5 \\ X_4 & X_2 & X_0 & X_1 & X_5 & X_3 \\ X_2 & X_0 & X_4 & X_5 & X_3 & X_1 \\ X_3 & X_1 & X_5 & X_0 & X_4 & X_2 \\ X_1 & X_5 & X_3 & X_4 & X_2 & X_0 \\ X_5 & X_3 & X_1 & X_2 & X_0 & X_4 \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_4 \\ Y_2 \\ Y_3 \\ Y_1 \\ Y_5 \end{pmatrix} \quad (13e)$$

This cyclic matrix has been partitioned into 3×3 blocks of 2×2 matrices. This technique, due to Winograd, of partitioning cyclic matrices, will be used repeatedly in the next section.

It is also readily established that Ψ_k of the cyclic convolution (13c) is the k -th coefficient of the polynomial

$$T(u) = X(u)Y(u) \text{ mod } (u^n - 1) \quad (13f)$$

where

$$X(u) = X_0 + X_1 u + \dots + X_{n-1} u^{n-1}$$

and

$$Y(u) = Y_0 + Y_1 u + \dots + Y_{n-1} u^{n-1}$$

Since $u^n - 1$ can be factored into polynomials over $GF(q)$, i.e.,

$$u^n - 1 = \prod_{i=1}^k g_i(u)$$

such that

$$(g_i(u), g_j(u)) = 1$$

for $j \neq i$, then, by the Chinese Remainder theorem, the coefficients of $T(u) \text{ mod } (u^n - 1)$ can be computed from the system of congruences defined by

$$T_i(u) = T(u) \bmod g_i(u) \text{ for } i = 1, 2, \dots, k \quad (14)$$

To get the factors, $g_i(u)$ defined in (14), first factor $u^n - 1$ into a product of irreducible polynomials. Each $g_i(u)$ is a product of one or more of these factors. Assume that α is an element of order n possible in some extension field of $GF(q)$. Then $\alpha^0, \alpha^1, \alpha^2, \dots, \alpha^{n-1}$ are all roots of $u^n - 1$ over $GF(q)$. By Th. 4.32, Ref. 17, it follows that

$$u^n - 1 = \prod_{\substack{d \\ d|n}} Q^{(d)}(u) \quad (15)$$

where $Q^{(d)}(u)$ is a polynomial whose roots are all elements of order d . $Q^{(d)}(u)$ is called the cyclotomic polynomial. By Th. 4.33, Ref. 17, the cyclotomic polynomial is given by

$$Q^{(d)}(u) = \prod_{k|d} (u^{d/k} - 1)^{\mu(d)} \quad (16)$$

where $\mu(d)$ is the Moebius function defined by

$$\mu(d) = \begin{cases} 1 & \text{if } d = 1 \\ (-1)^k & \text{if } d \text{ is the product of } k \text{ distinct primes} \\ 0 & \text{if } d \text{ contains any repeated prime factors.} \end{cases}$$

If n is a factor of $q - 1 = 2^p - 2$, then one can find an element α of order n in $GF(q)$ such that $\alpha^0, \alpha^1, \dots, \alpha^{n-1}$ are elements in $GF(q)$ and roots of $u^n - 1$. That is,

$$u^n - 1 = \sum_{i=0}^{n-1} (u - \alpha^i), \text{ where } \alpha^i \in GF(q)$$

Otherwise, one needs to compute the factorization of the $Q^{(d)}(u)$ into irreducible factors. To achieve this, if α is a root of $Q^{(d)}(u)$ with degree ℓ , then α is an element of order d in some field of characteristic q . By Th. 4.408, Ref. 17, α^{p^i} for $i = 0, 1, 2, \dots, \ell - 1$ are all roots of $Q^{(d)}(u)$. Suppose β is one of these roots. There are two cases to consider:

Case I: If $q \equiv 1 \pmod{n}$, then $\beta^q \equiv \beta \pmod{q}$. By Th. 4.407, Ref. 17, β is an element in $GF(q)$, thus, the factorization of $Q^{(d)}(u)$ is

$$Q^{(d)}(u) = \sum_{i=0}^{\ell-1} (u - \alpha_i), \text{ where } \alpha_i \in GF(q)$$

Case II: If $q \not\equiv 1 \pmod{n}$, then $\beta^q \not\equiv \beta \pmod{q}$. By Th. 4.407, Ref. 17, this implies that $\beta \notin GF(q)$. Thus, $Q^{(d)}(u)$ is an irreducible polynomial over $GF(q)$.

IV. Winograd's Algorithm for Computing the Transform over GF(q)

The discrete Fourier transform can be defined by

$$A_0 = \sum_{i=0}^{d-1} a_i \quad (17a)$$

and

$$A_j = a_0 + \sum_{i=1}^{d-1} a_i \gamma^{ij} \text{ for } j = 1, 2, \dots, d-1$$

or

$$A = B + Ia_0 = Wa + Ia_0 \quad (17b)$$

where

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_{d-1} \end{pmatrix}, \quad W = (\gamma^{ij})_{i,j \neq 0}, \quad a = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{d-1} \end{pmatrix}$$

and I is a unit matrix. If d is a prime, i.e., $d = p$, one can find an element $\alpha \in GF(p)$ which generates the cyclic subgroup of $d - 1$ elements, so that a permutation or substitution σ can be defined by

$$\sigma = \begin{pmatrix} 1, 2, \dots, p-2, p-1 \\ \alpha, \alpha^2, \dots, \alpha^{p-2}, \alpha^{p-1} = 1 \end{pmatrix}$$

Using the above permutation, one can permute the indices of B , defined in (17b), so that the matrix $\bar{W} = (\gamma^{\sigma(i)\sigma(j)})_{i,j \neq 0}$ is cyclic. That is,

$$B_{\sigma(j)} = \sum_{i=1}^{p-1} a_{\sigma(i)} \gamma^{\sigma(i)\sigma(j)} \text{ for } j = 1, 2, \dots, p-1$$

or

$$\bar{B} = \bar{W} \bar{a} \quad (18)$$

where

$$\bar{B} = \begin{pmatrix} B_{\sigma(1)} \\ B_{\sigma(2)} \\ \cdot \\ \cdot \\ \cdot \\ B_{\sigma(p-1)} \end{pmatrix}, \quad \bar{W} = (\gamma^{\sigma(i)\sigma(j)})_{i,j \neq 0}$$

and

$$\bar{a} = \begin{pmatrix} a_{\sigma(1)} \\ a_{\sigma(2)} \\ \cdot \\ \cdot \\ \cdot \\ a_{\sigma(p-1)} \end{pmatrix}$$

From (18), $B_{\sigma(j)}$ is a cyclic convolution of $a_{\sigma(p-j)}$ and $\gamma^{\sigma(j)}$ for $j = 1, 2, \dots, p-1$. Thus by the last section Eq. (18) is the set of coefficients of

$$T(u) = \left(\sum_{i=1}^{p-1} a_{\sigma(p-i)} u^{i-1} \right) \left(\sum_{i=1}^{p-1} \gamma^{\sigma(i)} u^{i-1} \right) \text{ mod } u^{p-1} - 1$$

Using the algorithm for factoring the polynomial $u^n - 1$ over $GF(q)$, described at the end of the last section, one can factor $u^{p-1} - 1$ over $GF(q)$ into irreducible relatively prime factors. That is,

$$u^{p-1} - 1 = \prod_{i=1}^k g_i(u), \text{ where } (g_i(u), g_j(u)) = 1 \text{ for } i \neq j$$

After computing the residues of $T(u) \text{ mod } g_i(u)$ for $i = 1, 2, \dots, k$, the Chinese Remainder theorem can be used to evaluate $T(u)$ with these residues. If $d = p^r$ is a factor of q , where $p \neq 2$, the number of integers relatively prime to p^r is $(p-1)p^{r-1}$. In this case, by Ref. 18, a set

$$\left\{ \gamma, \gamma^2, \dots, \gamma^{(p-1)p^{r-1}} = 1 \right\}$$

in $GF(q)$ can be found which is a cyclic group. Thus, by a procedure similar to that used to compute the above case for $d = p$, a p' point DFT can be obtained.

Now consider a d_i -point DFT over $GF(q)$ for $d_i = 3$ or 5 or 7 or 9 or 11 or 13. For a 3-point DFT over $GF(q)$, it is straightforward to show that the number of multiplications used to perform this transform is 2. Consider $d_i = 5$. Since 2 is a primitive element in $GF(5)$, the permutation σ is given by

$$\sigma = \begin{pmatrix} 1, 2, 3, 4 \\ 2, 4, 3, 1 \end{pmatrix}$$

using the above permutation, the matrix \bar{B} in (17b) is

$$\bar{B} = \bar{w} \bar{a}$$

where

$$\bar{B} = \begin{pmatrix} b_{\sigma(1)} \\ b_{\sigma(2)} \\ b_{\sigma(3)} \\ b_{\sigma(4)} \end{pmatrix}, \bar{w} = (\gamma^{\sigma(i)\sigma(j)})_{i,j \neq 0} \text{ and } \bar{a} = \begin{pmatrix} a_{\sigma(1)} \\ a_{\sigma(2)} \\ a_{\sigma(3)} \\ a_{\sigma(4)} \end{pmatrix}$$

More explicitly \bar{B} is

$$\begin{pmatrix} b_2 \\ b_4 \\ b_3 \\ b_1 \end{pmatrix} = \begin{pmatrix} \gamma^4 & \gamma^3 & \gamma & \gamma^2 \\ \gamma^3 & \gamma & \gamma^2 & \gamma^4 \\ \gamma & \gamma^2 & \gamma^4 & \gamma^3 \\ \gamma^2 & \gamma^4 & \gamma^3 & \gamma \end{pmatrix} \begin{pmatrix} a_2 \\ a_4 \\ a_3 \\ a_1 \end{pmatrix} \quad (19)$$

where γ is a 5th root of unity in $GF(q)$. $T(u)$ in (19) is obtained by computing the set of coefficients of

$$\begin{aligned} T(u) &= X(u) \cdot Y(u) \\ &= (\gamma^2 + \gamma^4 u + \gamma^3 u^2 + \gamma u^3) \cdot (a_1 + a_3 u + a_4 u^2 + a_2 u^3) \text{ mod } u^4 - 1 = (u - 1)(u + 1)(u^2 + 1) \end{aligned}$$

Let

$$\begin{aligned} m(u) &= (u - 1)(u + 1)(u^2 + 1) = m_1(u)m_2(u)m_3(u) \\ &= m_1(u)M_1(u) = m_2(u)M_2(u) = m_3(u)M_3(u) \end{aligned}$$

The system of congruences $T(u) \equiv T_i(u) \pmod{m_i(u)}$ for $i = 1, 2, 3$ is given by

$$\begin{aligned} T_1(u) &= X(1) \cdot Y(1) \equiv (\gamma^2 + \gamma^4 + \gamma^3 + \gamma) \cdot (a_1 + a_3 + a_4 + a_2) \\ &\equiv -(a_1 + a_3 + a_4 + a_2) \equiv C_1 \pmod{u-1}, \end{aligned}$$

$$\begin{aligned} T_2(u) &= X(-1) \cdot Y(-1) \equiv (\gamma^2 - \gamma^4 + \gamma^3 - \gamma) \cdot (a_1 - a_3 + a_4 - a_2) \\ &\equiv C_2 \pmod{u+1}, \end{aligned}$$

and

$$\begin{aligned} T_3(u) &= X(u) \cdot Y(u) \\ &\equiv [(\gamma^4 - \gamma)u + (\gamma^2 - \gamma^3)] \cdot [(a_3 - a_2)u + (a_1 - a_4)] \\ &\equiv [(au + b) \cdot (cu + d)] \pmod{u^2 + 1} \end{aligned} \tag{20}$$

where

$$C_1, C_2, a, b, c, d \in GF(q)$$

In order to compute (20), by Eq. (12b),

$$\begin{aligned} B(u) &= (au + b)(cu + d) \\ &= (au + b)(cu + d) \pmod{u(u+1)} + a \cdot c u(u+1) \end{aligned}$$

Let

$$R(u) = (au + b)(cu + d) \pmod{u(u+1)}$$

Then,

$$R_1(u) \equiv b \cdot d \equiv K_1 \pmod{u},$$

$$R_2(u) \equiv (b - a) \cdot (d - c) \equiv K_2 \pmod{u+1},$$

where $K_1, K_2 \in GF(q)$. Using Eq. (6a) this yields

$$R(u) = K_1 + (K_1 - K_2)u$$

ORIGINAL PAGE IS
OF POOR QUALITY

Thus,

$$B(u) = K_1 + (K_1 - K_2 + K_3)u + K_3u^2, \text{ where } K_3 = a \cdot c$$

Hence,

$$\begin{aligned} T_3(u) &\equiv (K_1 - K_3) + (K_1 - K_2 + K_3)u \\ &\equiv C_3 + C_4u \pmod{u^2 + 1} \end{aligned}$$

where

$$C_3 = K_1 - K_3 \text{ and } C_4 = (K_1 - K_2 + K_3)$$

Using Eq. (6a), $T(u)$ is

$$\begin{aligned} T(u) &= 2^{p-2}(C_1 - C_2 - 2C_4)u^3 + (C_1 + C_2 - C_3)u^2 + (C_1 - C_2 + 2C_4)u + (C_1 + C_2 + 2C_3) \\ &\equiv b_1u^3 + b_3u^2 + b_4u + b_2 \end{aligned}$$

It follows from this example that the number of integer multiplications used to perform a 5-point transform is 4.

For $d_i = 7$, the permutation σ is given by

$$\sigma = \begin{pmatrix} 1, 2, 3, 4, 5, 6, \\ 3, 2, 6, 4, 5, 1 \end{pmatrix}$$

Applying the above permutation to (17b), one obtains $\bar{B} = \bar{W}\bar{a}$ as

$$\begin{pmatrix} b_3 \\ b_2 \\ b_6 \\ b_4 \\ b_5 \\ b_1 \end{pmatrix} = \begin{pmatrix} \gamma^2, \gamma^6, \gamma^4, \gamma^5, \gamma^1, \gamma^3 \\ \gamma^6, \gamma^4, \gamma^5, \gamma^1, \gamma^3, \gamma^2 \\ \gamma^4, \gamma^5, \gamma^1, \gamma^3, \gamma^2, \gamma^6 \\ \gamma^5, \gamma^1, \gamma^3, \gamma^2, \gamma^6, \gamma^4 \\ \gamma^1, \gamma^3, \gamma^2, \gamma^6, \gamma^4, \gamma^5 \\ \gamma^3, \gamma^2, \gamma^6, \gamma^4, \gamma^5, \gamma^1 \end{pmatrix} \begin{pmatrix} a_3 \\ a_2 \\ a_6 \\ a_4 \\ a_5 \\ a_1 \end{pmatrix}$$

where γ is a 7th root of unity in $GF(q)$.

In the last section it was mentioned that there exists a permutation π of rows and columns so that the above cyclic matrix can be partitioned into a 2×2 block matrix of 3×3 cyclic blocks in the manner given in (13e). This permutation of the rows and columns is

$$\begin{pmatrix} b_3 \\ b_5 \\ b_6 \\ b_4 \\ b_2 \\ b_7 \end{pmatrix} = \begin{pmatrix} \gamma^2 \gamma^1 \gamma^4 \gamma^5 \gamma^6 \gamma^3 \\ \gamma^1 \gamma^4 \gamma^2 \gamma^6 \gamma^3 \gamma^5 \\ \gamma^4 \gamma^2 \gamma^1 \gamma^3 \gamma^5 \gamma^6 \\ \gamma^5 \gamma^6 \gamma^3 \gamma^2 \gamma^1 \gamma^4 \\ \gamma^6 \gamma^3 \gamma^5 \gamma^1 \gamma^4 \gamma^2 \\ \gamma^3 \gamma^5 \gamma^6 \gamma^4 \gamma^2 \gamma^1 \end{pmatrix} \begin{pmatrix} a_3 \\ a_5 \\ a_6 \\ a_4 \\ a_2 \\ a_7 \end{pmatrix} \quad (21a)$$

This has the form

$$\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = \begin{pmatrix} A & B \\ B & A \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = 2^{-1} \begin{pmatrix} (A+B)(Y_1+Y_2) + (A-B)(Y_1-Y_2) \\ (A+B)(Y_1+Y_2) - (A-B)(Y_1-Y_2) \end{pmatrix} \\ = 2^{p-1} \begin{pmatrix} D+E \\ D-E \end{pmatrix} \quad (21b)$$

Since A and B are cyclic matrices, it is evident that the matrices $A+B$ and $A-B$ are also cyclic matrices. In (21b), D is defined as

$$D = \begin{pmatrix} d_0 \\ d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} X_0 & X_1 & X_2 \\ X_1 & X_2 & X_0 \\ X_2 & X_0 & X_1 \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_1 \\ Y_2 \end{pmatrix} \quad (22)$$

This matrix can be obtained by computing the set of coefficients of

$$T(u) = (X_2 + X_0 u + X_1 u^2) (Y_2 + Y_1 u + Y_0 u^2) \bmod u^3 - 1 = (u-1)(u^2 + u + 1) \quad (23)$$

The system of congruences, given in (23), is

$$T_1(u) = (X_0 + X_1 + X_2) \cdot (Y_0 + Y_1 + Y_2) \equiv C_1 \bmod (u-1)$$

and

$$T_2(u) \equiv [(X_2 - X_1) + (X_0 - X_1)u] \cdot [(Y_2 - Y_0) + (Y_1 - Y_0)u] \\ \equiv C_2 + C_3 u \bmod u^2 + u + 1$$

ORIGINAL PAGE IS
OF POOR QUALITY

where $C_1, C_2, C_3 \in GF(q)$. By the same procedure for computing Eq. (20), $T_2(u)$ can be obtained, using only 3 multiplications. By (6a), $T(u)$ is given by

$$\begin{aligned} T(u) &= [3^{-1}(C_1 + C_2 + C_3) - C_2] + [3^{-1}(C_1 + C_2 + C_3) - C_3]u + [3^{-1}(C_1 + C_2 + C_3)]u^2 \\ &= d_0 + d_1u + d_2u^2 \end{aligned} \quad (24)$$

In a similar fashion matrix E , given in (21b), can also be obtained. Thus, the number of multiplications used to perform a 7-point transform is 8.

Consider $d_i = 3^2$. Since the integers 1, 2, 4, 5, 7, 8 are relatively prime to 9, the permutation σ is defined by

$$\sigma = \begin{pmatrix} 1, 2, 4, 5, 7, 8 \\ 2, 4, 8, 7, 5, 1 \end{pmatrix} \quad (25)$$

Rearranging the rows and columns of B in such a manner that the elements of the matrix with indices relatively prime to 9 form a block, one has,

$$\begin{pmatrix} b_1 \\ b_2 \\ b_4 \\ b_5 \\ b_7 \\ b_8 \\ b_3 \\ b_6 \end{pmatrix} = \begin{pmatrix} \gamma^{1 \cdot 1}, \gamma^{1 \cdot 2}, \gamma^{1 \cdot 4}, \gamma^{1 \cdot 5}, \gamma^{1 \cdot 7}, \gamma^{1 \cdot 8}, \gamma^{1 \cdot 3}, \gamma^{1 \cdot 6} \\ \gamma^{2 \cdot 1}, \gamma^{2 \cdot 2}, \gamma^{2 \cdot 4}, \gamma^{2 \cdot 5}, \gamma^{2 \cdot 7}, \gamma^{2 \cdot 8}, \gamma^{2 \cdot 3}, \gamma^{2 \cdot 6} \\ \gamma^{4 \cdot 1}, \gamma^{4 \cdot 2}, \gamma^{4 \cdot 4}, \gamma^{4 \cdot 5}, \gamma^{4 \cdot 7}, \gamma^{4 \cdot 8}, \gamma^{4 \cdot 3}, \gamma^{4 \cdot 6} \\ \gamma^{5 \cdot 1}, \gamma^{5 \cdot 2}, \gamma^{5 \cdot 4}, \gamma^{5 \cdot 5}, \gamma^{5 \cdot 7}, \gamma^{5 \cdot 8}, \gamma^{5 \cdot 3}, \gamma^{5 \cdot 6} \\ \gamma^{7 \cdot 1}, \gamma^{7 \cdot 2}, \gamma^{7 \cdot 4}, \gamma^{7 \cdot 5}, \gamma^{7 \cdot 7}, \gamma^{7 \cdot 8}, \gamma^{7 \cdot 3}, \gamma^{7 \cdot 6} \\ \gamma^{8 \cdot 1}, \gamma^{8 \cdot 2}, \gamma^{8 \cdot 4}, \gamma^{8 \cdot 5}, \gamma^{8 \cdot 7}, \gamma^{8 \cdot 8}, \gamma^{8 \cdot 3}, \gamma^{8 \cdot 6} \\ \gamma^{3 \cdot 1}, \gamma^{3 \cdot 2}, \gamma^{3 \cdot 4}, \gamma^{3 \cdot 5}, \gamma^{3 \cdot 7}, \gamma^{3 \cdot 8}, \gamma^{3 \cdot 3}, \gamma^{3 \cdot 6} \\ \gamma^{6 \cdot 1}, \gamma^{6 \cdot 2}, \gamma^{6 \cdot 4}, \gamma^{6 \cdot 5}, \gamma^{6 \cdot 7}, \gamma^{6 \cdot 8}, \gamma^{6 \cdot 3}, \gamma^{6 \cdot 6} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_4 \\ a_5 \\ a_7 \\ a_8 \\ a_3 \\ a_6 \end{pmatrix} \quad (26)$$

Define the upper left 6×6 matrix of (26) as

$$\begin{pmatrix} y_1 \\ y_2 \\ y_4 \\ y_5 \\ y_7 \\ y_8 \end{pmatrix} = \begin{pmatrix} \gamma^{1 \cdot 1}, \gamma^{1 \cdot 2}, \gamma^{1 \cdot 4}, \gamma^{1 \cdot 5}, \gamma^{1 \cdot 7}, \gamma^{1 \cdot 8} \\ \gamma^{2 \cdot 1}, \gamma^{2 \cdot 2}, \gamma^{2 \cdot 4}, \gamma^{2 \cdot 5}, \gamma^{2 \cdot 7}, \gamma^{2 \cdot 8} \\ \gamma^{4 \cdot 1}, \gamma^{4 \cdot 2}, \gamma^{4 \cdot 4}, \gamma^{4 \cdot 5}, \gamma^{4 \cdot 7}, \gamma^{4 \cdot 8} \\ \gamma^{5 \cdot 1}, \gamma^{5 \cdot 2}, \gamma^{5 \cdot 4}, \gamma^{5 \cdot 5}, \gamma^{5 \cdot 7}, \gamma^{5 \cdot 8} \\ \gamma^{7 \cdot 1}, \gamma^{7 \cdot 2}, \gamma^{7 \cdot 4}, \gamma^{7 \cdot 5}, \gamma^{7 \cdot 7}, \gamma^{7 \cdot 8} \\ \gamma^{8 \cdot 1}, \gamma^{8 \cdot 2}, \gamma^{8 \cdot 4}, \gamma^{8 \cdot 5}, \gamma^{8 \cdot 7}, \gamma^{8 \cdot 8} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_4 \\ a_5 \\ a_7 \\ a_8 \end{pmatrix} \quad (27)$$

Applying the permutations defined in (25) to the indices of (26), one obtains

$$\begin{pmatrix} y_2 \\ y_4 \\ y_8 \\ y_7 \\ y_5 \\ y_1 \end{pmatrix} = \begin{pmatrix} \gamma^4, \gamma^8, \gamma^7, \gamma^5, \gamma^1, \gamma^2 \\ \gamma^8, \gamma^7, \gamma^5, \gamma^1, \gamma^2, \gamma^4 \\ \gamma^7, \gamma^5, \gamma^1, \gamma^2, \gamma^4, \gamma^8 \\ \gamma^5, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^7 \\ \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^7, \gamma^5 \\ \gamma^2, \gamma^4, \gamma^8, \gamma^7, \gamma^5, \gamma^1 \end{pmatrix} \begin{pmatrix} a_2 \\ a_4 \\ a_8 \\ a_7 \\ a_5 \\ a_1 \end{pmatrix}$$

By a similar procedure used to partition the matrix (13e) the above matrix becomes a 2×2 block matrix of 3×3 cyclic blocks as follows:

$$\begin{pmatrix} y_2 \\ y_5 \\ y_8 \\ y_7 \\ y_4 \\ y_1 \end{pmatrix} = \begin{pmatrix} \gamma^4, \gamma^1, \gamma^7, \gamma^5, \gamma^8, \gamma^2 \\ \gamma^1, \gamma^7, \gamma^4, \gamma^8, \gamma^2, \gamma^5 \\ \gamma^7, \gamma^4, \gamma^1, \gamma^2, \gamma^5, \gamma^8 \\ \gamma^5, \gamma^8, \gamma^2, \gamma^4, \gamma^1, \gamma^7 \\ \gamma^8, \gamma^2, \gamma^5, \gamma^1, \gamma^7, \gamma^4 \\ \gamma^2, \gamma^5, \gamma^8, \gamma^7, \gamma^4, \gamma^1 \end{pmatrix} \begin{pmatrix} a_2 \\ a_5 \\ a_8 \\ a_7 \\ a_4 \\ a_1 \end{pmatrix} \quad (28)$$

Using the same procedure for computing the 6×6 cyclic matrix, described previously, we know that the number of multiplications required to perform (28) is 8. The last two columns of the matrix defined in (26) can be obtained by computing the following 2×2 cyclic matrix,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \gamma^3 & \gamma^6 \\ \gamma^6 & \gamma^3 \end{pmatrix} \begin{pmatrix} a_3 \\ a_6 \end{pmatrix} \quad (29)$$

The last two rows of the matrix defined in (26) can be obtained by computing the following cyclic matrix

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \gamma^3 & \gamma^6 \\ \gamma^6 & \gamma^3 \end{pmatrix} \begin{pmatrix} a_1 + a_4 + a_7 \\ a_2 + a_5 + a_8 \end{pmatrix} \quad (30)$$

Note that the computation of (29) and (30) are the same as computing the 2×2 cyclic matrix in a 3-point DFT. Hence, the total number of multiplications used to perform this transform is 12.

For $d_i = 11$, the permutation σ is given by

$$\sigma = \begin{pmatrix} 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 \\ 2, 4, 8, 5, 10, 9, 7, 3, 6, 1 \end{pmatrix}$$

Applying the above permutation to (17b) one obtains $\overline{B} = \overline{W}\overline{a}$, i.e.,

$$\begin{pmatrix} b_2 \\ b_4 \\ b_8 \\ b_5 \\ b_{10} \\ b_9 \\ b_7 \\ b_3 \\ b_6 \\ b_1 \end{pmatrix} = \begin{pmatrix} \gamma^4, \gamma^8, \gamma^5, \gamma^{10}, \gamma^9, \gamma^7, \gamma^3, \gamma^6, \gamma^1, \gamma^2 \\ \gamma^8, \gamma^5, \gamma^{10}, \gamma^9, \gamma^7, \gamma^3, \gamma^6, \gamma^1, \gamma^2, \gamma^4 \\ \gamma^5, \gamma^{10}, \gamma^9, \gamma^7, \gamma^3, \gamma^6, \gamma^1, \gamma^2, \gamma^4, \gamma^8 \\ \gamma^{10}, \gamma^9, \gamma^7, \gamma^3, \gamma^6, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^5 \\ \gamma^9, \gamma^7, \gamma^3, \gamma^6, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^5, \gamma^{10} \\ \gamma^7, \gamma^3, \gamma^6, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^5, \gamma^{10}, \gamma^9 \\ \gamma^3, \gamma^6, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^5, \gamma^{10}, \gamma^9, \gamma^7 \\ \gamma^6, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^5, \gamma^{10}, \gamma^9, \gamma^7, \gamma^3 \\ \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^5, \gamma^{10}, \gamma^9, \gamma^7, \gamma^3, \gamma^6 \\ \gamma^2, \gamma^4, \gamma^8, \gamma^5, \gamma^{10}, \gamma^9, \gamma^7, \gamma^3, \gamma^6, \gamma^1 \end{pmatrix} \begin{pmatrix} a_2 \\ a_4 \\ a_8 \\ a_5 \\ a_{10} \\ a_9 \\ a_7 \\ a_3 \\ a_6 \\ a_1 \end{pmatrix}$$

where γ is a 11th root of unity in $GF(q)$. By the same procedure used to partition the matrix given in (13e) and (21a) the above matrix can be partitioned into blocks of 5×5 cyclic matrices, such that the blocks form a 2×2 cyclic matrix. That is,

$$\begin{pmatrix} b_2 \\ b_7 \\ b_8 \\ b_6 \\ b_{10} \\ b_9 \\ b_4 \\ b_3 \\ b_5 \\ b_1 \end{pmatrix} = \begin{pmatrix} \gamma^4, \gamma^3, \gamma^5, \gamma^1, \gamma^9 & \gamma^7, \gamma^8, \gamma^6, \gamma^{10}, \gamma^2 \\ \gamma^3, \gamma^5, \gamma^1, \gamma^9, \gamma^4 & \gamma^8, \gamma^6, \gamma^{10}, \gamma^2, \gamma^7 \\ \gamma^5, \gamma^1, \gamma^9, \gamma^4, \gamma^3 & \gamma^6, \gamma^{10}, \gamma^2, \gamma^7, \gamma^8 \\ \gamma^1, \gamma^9, \gamma^4, \gamma^3, \gamma^5 & \gamma^{10}, \gamma^2, \gamma^7, \gamma^8, \gamma^6 \\ \gamma^9, \gamma^4, \gamma^3, \gamma^5, \gamma^1 & \gamma^2, \gamma^7, \gamma^8, \gamma^6, \gamma^{10} \\ \gamma^7, \gamma^8, \gamma^6, \gamma^{10}, \gamma^2 & \gamma^4, \gamma^3, \gamma^5, \gamma^1, \gamma^9 \\ \gamma^8, \gamma^6, \gamma^{10}, \gamma^2, \gamma^7 & \gamma^3, \gamma^5, \gamma^1, \gamma^9, \gamma^4 \\ \gamma^6, \gamma^{10}, \gamma^2, \gamma^7, \gamma^8 & \gamma^5, \gamma^1, \gamma^9, \gamma^4, \gamma^3 \\ \gamma^{10}, \gamma^2, \gamma^7, \gamma^8, \gamma^6 & \gamma^1, \gamma^9, \gamma^4, \gamma^3, \gamma^5 \\ \gamma^2, \gamma^7, \gamma^8, \gamma^6, \gamma^{10} & \gamma^9, \gamma^4, \gamma^3, \gamma^5, \gamma^1 \end{pmatrix} \begin{pmatrix} a_2 \\ a_7 \\ a_8 \\ a_6 \\ a_{10} \\ a_9 \\ a_4 \\ a_3 \\ a_5 \\ a_1 \end{pmatrix}$$

This has a form similar to the matrix (21b), where A and B are 5×5 cyclic matrices. The corresponding 5-point vectors D and E can be obtained by direct computations without using 5-point cyclic convolutions. Thus, the number of multiplications needed to perform a 11-point transform is 50.

For $d_i = 13$, the permutation σ is given by

$$\sigma = \begin{pmatrix} 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 \\ 2, 4, 8, 3, 6, 12, 11, 9, 5, 10, 7, 1 \end{pmatrix}$$

Applying the above permutation σ to (17b), this yields

$$\begin{pmatrix} b_2 \\ b_4 \\ b_8 \\ b_3 \\ b_6 \\ b_{12} \\ b_{11} \\ b_9 \\ b_5 \\ b_{10} \\ b_7 \\ b_1 \end{pmatrix} = \begin{pmatrix} \gamma^4, \gamma^8, \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10}, \gamma^7, \gamma^1, \gamma^2 \\ \gamma^8, \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10}, \gamma^7, \gamma^1, \gamma^2, \gamma^4 \\ \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10}, \gamma^7, \gamma^1, \gamma^2, \gamma^4, \gamma^8 \\ \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10}, \gamma^7, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^3 \\ \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10}, \gamma^7, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^3, \gamma^6 \\ \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10}, \gamma^7, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^3, \gamma^6, \gamma^{12} \\ \gamma^9, \gamma^5, \gamma^{10}, \gamma^7, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11} \\ \gamma^5, \gamma^{10}, \gamma^7, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9 \\ \gamma^{10}, \gamma^7, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5 \\ \gamma^7, \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10} \\ \gamma^1, \gamma^2, \gamma^4, \gamma^8, \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10}, \gamma^7 \\ \gamma^2, \gamma^4, \gamma^8, \gamma^3, \gamma^6, \gamma^{12}, \gamma^{11}, \gamma^9, \gamma^5, \gamma^{10}, \gamma^7, \gamma^1 \end{pmatrix} \begin{pmatrix} a_2 \\ a_4 \\ a_8 \\ a_3 \\ a_6 \\ a_{12} \\ a_{11} \\ a_9 \\ a_5 \\ a_{10} \\ a_7 \\ a_1 \end{pmatrix}$$

where γ is a 13th root of unity in $GF(q)$. Then by a similar procedure used to partition the matrix (13e) or (21a), the above matrix can be partitioned into blocks of 4×4 cyclic matrices, such that the blocks form a 3×3 cyclic matrix. That is,

$$\begin{pmatrix} E_0 \\ E_1 \\ E_2 \end{pmatrix} = \begin{pmatrix} A & B & C \\ B & C & A \\ C & A & B \end{pmatrix} \begin{pmatrix} Y_0 \\ Y_1 \\ Y_2 \end{pmatrix} \quad (31)$$

where

$$E_0 = \begin{pmatrix} b_2 \\ b_{10} \\ b_{11} \\ b_3 \end{pmatrix}, \quad E_1 = \begin{pmatrix} b_6 \\ b_4 \\ b_7 \\ b_9 \end{pmatrix}, \quad E_2 = \begin{pmatrix} b_5 \\ b_{12} \\ b_8 \\ b_{11} \end{pmatrix}$$

$$A = \begin{pmatrix} \gamma^4 \gamma^7 \gamma^9 \gamma^6 \\ \gamma^7 \gamma^9 \gamma^6 \gamma^4 \\ \gamma^9 \gamma^6 \gamma^4 \gamma^7 \\ \gamma^6 \gamma^4 \gamma^7 \gamma^9 \end{pmatrix} \quad B = \begin{pmatrix} \gamma^{12} \gamma^8 \gamma^1 \gamma^5 \\ \gamma^8 \gamma^1 \gamma^5 \gamma^{12} \\ \gamma^1 \gamma^5 \gamma^{12} \gamma^8 \\ \gamma^5 \gamma^{12} \gamma^8 \gamma^1 \end{pmatrix}$$

$$C = \begin{pmatrix} \gamma^{10} \gamma^{11} \gamma^3 \gamma^2 \\ \gamma^{11} \gamma^3 \gamma^2 \gamma^{10} \\ \gamma^3 \gamma^2 \gamma^{10} \gamma^{11} \\ \gamma^2 \gamma^{10} \gamma^{11} \gamma^3 \end{pmatrix} \quad Y_0 = \begin{pmatrix} a_2 \\ a_{10} \\ a_{11} \\ a_3 \end{pmatrix}$$

$$Y_1 = \begin{pmatrix} a_6 \\ a_4 \\ a_7 \\ a_9 \end{pmatrix} \quad Y_2 = \begin{pmatrix} a_5 \\ a_{12} \\ a_8 \\ a_{11} \end{pmatrix}$$

Note that A , B , and C are 4×4 cyclic matrices. Using the same procedure for computing the system given by (23), the above system can be obtained as

$$\begin{aligned} E_0 &= 3^{-1}(C_1 + C_2 + C_3) - C_2 \\ E_1 &= 3^{-1}(C_1 + C_2 + C_3) - C_3 \\ E_2 &= 3^{-1}(C_1 + C_2 + C_3) \end{aligned} \quad (32)$$

where

$$\begin{aligned} C_1 &= (C + A + B)(Y_0 + Y_1 + Y_2) \\ C_2 &= M_1 - M_2 = (C - B)(Y_2 - Y_0) - (A - B)(Y_1 - Y_0) \\ C_3 &= M_1 - M_3 = (C - B)(Y_2 - Y_0) - (C - A)(Y_2 - Y_0) \end{aligned}$$

In (32), it is evident that the computation of C_1 or M_1 or M_2 or M_3 requires 5 multiplies. Hence the total number of multiplications needed to perform the transform defined by (31) is 16.

The total number of integer multiplications of d_i -point transform over $GF(q)$ for $d_i = 2, 3, 5, 7, 9, 11, 13$, P is shown in Table 2.

V. The DFT Over $GF(q)$ by Multidimensional Techniques

To compute the transform of longer sequences, let $d = d_1 \cdot d_2 \dots d_r$, where $(d_i, d_j) = 1$, for $i \neq j$, and let R_{d_i} be the ring of integers modulo d_i . Then, by using the Chinese Remainder theorem (Theorem 1 for integers), it can be shown that the direct sum of rings

$$R_d = R_{d_1} \dot{+} R_{d_2} \dot{+} \dots \dot{+} R_{d_r}$$

$$= \left\{ (\alpha_1, \alpha_2, \dots, \alpha_r) \mid \alpha_k \in R_{d_k} \text{ for } k = 1, \dots, r \right\}$$

where

$$(\alpha_1, \alpha_2, \dots, \alpha_r) + (\beta_1, \beta_2, \dots, \beta_r) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_r + \beta_r)$$

and

$$(\alpha_1, \alpha_2, \dots, \alpha_r) \cdot (\beta_1, \beta_2, \dots, \beta_r) = (\alpha_1 \cdot \beta_1, \alpha_2 \cdot \beta_2, \dots, \alpha_r \cdot \beta_r)$$

is a ring of d elements which is isomorphic to the ring R_d . If $d = d_1 \cdot d_2 \dots d_r$, by using the direct sum of finite rings, it is shown in Ref. 13 that the d -point DFT over $GF(q)$ defined in (1) can be decomposed into multidimensional DFT as follows:

$$A_{(j_1, j_2, \dots, j_r)} = \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \dots \sum_{i_r=0}^{d_r-1} a_{(i_1, i_2, \dots, i_r)} \gamma^{(i_1 j_1, 0, \dots, 0)}$$

$$\gamma^{(0, i_2 j_2, \dots, 0)} \dots \gamma^{(0, \dots, i_r j_r)}$$

$$= \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \dots \sum_{i_r=0}^{d_r-1} a_{(i_1, i_2, \dots, i_r)} \gamma^{(1, 0, \dots, 0)} i_1^{j_1}$$

$$\gamma^{(0, 1, 0, \dots, 0)} i_2^{j_2} \dots \gamma^{(0, 0, \dots, 1)} i_r^{j_r}$$

$$\equiv \sum_{i_1=0}^{d_1-1} \sum_{i_2=0}^{d_2-1} \dots \sum_{i_r=0}^{d_r-1} a_{(i_1, i_2, \dots, i_r)} \gamma_1^{i_1^{j_1}} \gamma_2^{i_2^{j_2}} \dots \gamma_r^{i_r^{j_r}} \quad (33)$$

where $\gamma_k = \gamma^{(0, \dots, 0, 1, 0, \dots, 0)}$ with the 1 in the k -th position as a d_k -th root of unity in R_{d_k} . Assume the number of multiplications used to perform d_i -point DFT for $i = 1, 2, \dots, r$ is m_i . Then, by (33), it is evident that the number of multiplications for computing d -point DFT over $GF(q)$ is equal to

$$N = d_1 \cdot d_2 \dots d_{r-1} m_r + d_1 \cdot d_2 \dots d_{r-2} d_r m_{r-1} + \dots + d_2 d_3 \dots d_r m_1$$

A simple example for computing the DFT over $GF(q)$ by using multidimensional techniques is now presented.

Example: Let $q = 2^5 - 1$ and $d = 2^5 - 2 = 2 \cdot 3 \cdot 5 = d_1 d_2 d_3$. Compute the 30-point DFT over $GF(q)$.

Since $\alpha = 3$ is a primitive element in $GF(31)$, then 3^j is also a primitive element in $GF(31)$ for $j = 7, 11, 13, 17, 19, 21, 23, 27, 29$. The choice of $r = 3^{19}$ yields

$$(\gamma)^6 = (3^{19})^6 \equiv 3^{24} \equiv 2 \pmod{31}$$

Thus, one can find the primitive element $\gamma = 3^{19}$ such that $\gamma_1 = \gamma^{(1,0,0)} = \gamma^{15} = -1, \gamma_2 = \gamma^{(0,1,0)} = \gamma^{10} = -6, \gamma_3 = \gamma^{(0,0,1)} = \gamma^6 = 2$ are the elements of order 2, 3, 5, respectively. By (33), this FFT algorithm consists of the following 3 stages:

Stage 1:

$$\begin{aligned} A^1_{(i_1, i_2, j_3)} &= \sum_{i_3=0}^4 a_{(i_1, i_2, i_3)} \gamma^{(0,0,1)} i_3^{j_3} \\ &= \sum_{i_3=0}^4 a_{(i_1, i_2, i_3)} 2^{i_3 j_3} \text{ for } j_3 = 0, 1, 2, 3, 4 \end{aligned} \quad (34)$$

Stage 2:

$$\begin{aligned} A^2_{(i_1, j_2, j_3)} &= \sum_{i_2=0}^2 A^1_{(i_1, i_2, j_3)} \gamma^{(0,1,0)} i_2^{j_2} \\ &= \sum_{i_2=0}^2 A^1_{(i_1, i_2, j_3)} (-6)^{i_2 j_2} \text{ for } j_2 = 0, 1, 2 \end{aligned}$$

Stage 3:

$$\begin{aligned} A^3_{(j_1, j_2, j_3)} &= \sum_{i_1=0}^1 A^2_{(i_1, j_2, j_3)} \gamma^{(1,0,0)} i_1^{j_1} \\ &= \sum_{i_1=0}^1 A^2_{(i_1, j_2, j_3)} (-1)^{i_1 j_1} \text{ for } j_1 = 0, 1 \end{aligned}$$

In (34), we observe that both $A^1_{(i_1, i_2, j_3)}$ and $A^3_{(j_1, j_2, j_3)}$ can be evaluated without multiplications and that $A^2_{(i_1, j_2, j_3)}$ is a 3-point DFT over $GF(q)$. By Table 1, the number of multiplications used to perform $A^2_{(i_1, j_2, j_3)}$ for $j_2 = 0, 1, 2$ is 1. This requires a total of $N = 26 \cdot 0 + 10 \cdot 1 + 15 \cdot 0 = 10$ integer multiplications modulo 31 for evaluating (30).

For most applications to digital filters, the two most important Mersenne primes are $2^{31} - 1$ and $2^{61} - 1$. The number of real integer multiplications used to perform a DFT over $GF(q)$ of $d = 2 \cdot p \cdot r_3$, where $r_3 = 5, 7, 9, 11, \text{ or } 13$ for $q = 2^{31} - 1$ and $q = 2^{61} - 1$ is given in Table 3. The present algorithm, and Winograd's new algorithm (Ref. 13) are compared in Table 3 by giving the number of real multiplications needed to perform these algorithms. These results for Winograd's algorithm come from Table 2, in Reference 13.

In Table 3, one can see that the transform over $GF(q)$ appears comparable in speed with that given by Winograd (Ref. 13).

Acknowledgement

The authors wish to thank Dr. N. A. Renzetti, Manager of Tracking and Data Acquisition Engineering, and the members of the Advanced Engineering Group in that organization at the Jet Propulsion Laboratory for their early support, suggestions, and encouragement of the research which led to this paper. We also thank Dr. C. A. Greenhall for his mathematical suggestions.

References

1. Pollar, J. M.: "The Fast Fourier Transform in a Finite Field," *Math. Comput.*, 1971, 25, pp. 365-374.
2. Schonhage, A., and Strassen, V.: "Schnelle Multiplication Grosser Zahlen," *Computing*, 1971, 7, pp. 281-292.
3. Rader, C. M.: "Discrete Convolution via Mersenne Transforms," *IEEE Trans. Comp.* 1972, C-21, pp. 1269-1273.
4. Agarwal, R. C., and Burrus, C. S.: "Number Theoretic Transforms to Implement Fast Digital Convolution," *Proc. IEEE*, 1975, 63, pp. 550-560.
5. Reed, I. S., and Truong, T. K.: "The Use of Finite Fields to Compute Convolutions," *IEEE Trans. Inform. Theory*, 1975, It-21, pp. 208-213.
6. Reed, I. S., and Truong, T. K.: "Complex Integer Convolution Over a Direct Sum of Galois Fields," *IEEE Trans. Inform. Theory*, 1975, IT-21, pp. 657-661.
7. Vegh, E., and Leibowitz, L. M.: "Fast Complex Convolution in Finite Rings," *IEEE Trans.*, 1976, ASSP-24, pp. 343-344.
8. Golomb, S. W., and Reed, I. S., and Truong, T. K.: "Integer Convolutions Over the Finite Field $GF(3 \cdot 2^n + 1)$," *SIAM J. on Applied Math.*, Vol. 32, No. 2, March 1977.
9. Golomb, S. W.: "Properties of the Sequences $3 \cdot 2^n + 1$," *Math. Comput.* 1976, 30, pp. 657-663.

10. Pollar, J. M.: "Implementation of Number-Theoretic Transforms," *Electro. Lett.*, 1976, 12, pp. 378-379.
11. Liu, K. Y., Reed, I. S., and Truong, T. K.: "Fast Number-Theoretic Transforms for Digital Filtering," *Electron. Lett.*, 1976, 12, pp. 644-646.
12. Reed, I. S., Truong, T. K., and Liu, K. Y., "A New Fast Algorithm for Computing Complex Number-Theoretic Transforms", *Electron Lett.*, 1977, pp. 278-82.
13. Winograd, S.: "On Computing the Discrete Fourier Transform," *Proc Nat. Acad. Sci. USA*, 1976, 73, pp. 1005-1006.
14. S. Winograd, "On Computing the Discrete Fourier Transform," Research Report, Math. Science Dept., IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10592.
15. J. Justesen, "On the Complexity of Decoding of Reed-Solomon Codes," *IEEE Trans. Inform Theory*, Vol. IT-22, March 1976, pp. 237-238.
16. I. S. Reed, R. A. Scholtz, T. K. Truong, L. R. Welch, "The Fast Decoding of Reed-Solomon Codes Using Fermat Theoretic Transforms and Continued Fractions," to be published in *IEEE Trans. Inform theory*.
17. E. R. Berlekamp, *Algebraic Coding Theory*, New York, McGraw Hill, 1968; Chapter 7.
18. I. M. Vinogradov, *Elements of Number Theory*, Dover Publications, New York, 1954.

Table 1. Factorization of the integers $2^p - 2$

p	$2^p - 2$
13	$3^2 \cdot 2 \cdot 5 \cdot 7 \cdot 13$
17	$2 \cdot 3 \cdot 5 \cdot 17 \cdot 257$
31	$2 \cdot 3^2 \cdot 7 \cdot 11 \cdot 31 \cdot 331 \cdot 151$
61	$2 \cdot 3^2 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 31 \cdot 41 \cdot 61 \cdot 151 \cdot 331 \cdot 1321$

Table 2. The complexity of transforms over $GF(q)$ of small sequences where $q = 2^p - 1$

d_i	No. of Integer Multiplications
2	0
3	1
5	4
7	8
9	12
11	50
13	21
p	0

Table 3. The complexity of the transform over $GF(q)$ where $q = 2^{31} - 1$ and $q = 2^{61} - 1$

d	No. real integer multiplications of transform over $GF(2^{31} - 1)$	No. real integer multiplications of transform over $GF(2^{61} - 1)$	No. real integer multiplications of Winograd's new algorithm for real data
60			72
62	0		
120			144
122		0	
168			216
186	62		
240			324
310	248		
366		122	
420			648
434	496		
504			936
558	744		
610		488	
806	1302		
840			1296
854		976	
1008			2106
1098		1464	
1586		1952	
1674	2790		
2520			5616
3348		5490	

ORIGINAL PAGE IS
OF POOR QUALITY

Appendix A

Let $a \in GF(q)$, where $(a, q) = 1$ and also let $S = a/q$. In this Appendix, it will be shown that the inverse element of a can be obtained by using continued fractions.

If $S = a/q$, where $a \in GF(q)$, using Euclid's algorithm, i.e.,

$$r_{k-2} = a_k r_{k-1} + r_k, \quad 0 < r_k < r_{k-1} \quad \text{for } k = 1, 2, \dots, n-1 \quad (\text{A-1})$$

with initial conditions $r_{-1} = q$, $r_{-2} = a$ and $r_{n-2} = r_{n-1} a_n$, one generates the sequence of the partial quotients, a_0, a_1, \dots, a_n . By (A-1), S can be developed into a continued fraction.

$$S = a_0 + (a_1 + (\dots (a_k + \alpha_k)^{-1} \dots)^{-1} \dots)^{-1}, \quad k \leq n \quad (\text{A-2})$$

By setting $\alpha_k = 0$ in (A-2), one can determine a k^{th} order approximation to S , which is called a *convergent*,

$$S_k = a_0 + (a_1 + (\dots (\dots (a_k)^{-1} \dots)^{-1} \dots)^{-1})^{-1}$$

From (A-1), $S_0, S_1, \dots, S_k, \dots$ will terminate with S_n since $r_n = 0$. Thus, $S_n = a/q$, where n is a finite number.

A recursive formula for convergents is generated as follows:

$$S_0 = \frac{a_0}{1} = \frac{p_0}{q_0}$$

$$S_1 = a_0 + \frac{1}{a_1} = \frac{a_1 a_0 + 1}{a_1 \cdot 1 + 0} = \frac{a_1 p_0 + p_{-1}}{a_1 q_0 + q_{-1}} = \frac{p_1}{q_1}$$

$$S_2 = a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = \frac{(a_1 + 1/a_2)p_0 + p_{-1}}{(a_1 + 1/a_2)q_0 + q_{-1}} = \frac{a_2 p_1 + p_0}{a_2 q_1 + q_0} = \frac{p_2}{q_2}$$

The recursive convergents are defined as

$$S_k = \frac{a_k p_{k-1} + p_{k-2}}{a_k q_{k-1} + q_{k-2}} = \frac{p_k}{q_k} \quad (\text{A-3})$$

where $p_{-1} = 1$, $q_{-1} = 0$, $p_0 = a_0$, and $q_0 = 1$ for $k = 1, 2, \dots, n$.

Let

$$I_k = p_k q_{k-1} - q_k p_{k-1}$$

By (A-3),

$$I_k = p_k q_{k-1} - q_k p_{k-1} = -I_{k-1} \quad (\text{A-4})$$

Since $I_0 = p_0 q_{-1} - q_0 p_{-1} = a_0 \cdot 0 - 1 \cdot 1 = -1$, one has by (A-4), $I_1 = -I_0 = 1$. With the above result, one has $I_k = (-1)^{k+1}$. It follows that

$$p_k q_{k-1} - q_k p_{k-1} = (-1)^{k+1} \text{ for } k \geq 0 \quad (\text{A-5})$$

If $n = k$, then $S_n = p_n/q_n = a/q$. Thus, (A-5) becomes

$$a q_{n-1} - q p_{n-1} = (-1)^{n+1} \quad (\text{A-6})$$

It follows from (A-6) that the solutions of $a x \equiv 1 \pmod{q}$ are given by

$$x \equiv q_{n-1} \pmod{q} \text{ if } n \text{ is odd,}$$

$$x \equiv -q_{n-1} \pmod{q} \text{ if } n \text{ is even} \quad (\text{A-7})$$

In order to determine an upper bound on the number of partial quotients to form a continued fraction for a/q , $1 < a < q$, the following lemma and theorem are needed:

Lemma 1: Sequence $\{q_k\}$ defined in (A-3) as a function of a_1, a_2, \dots, a_k increases most slowly for $a_1 = a_2 = \dots = a_k = 1$ for all k .

Proof: By (A-3), $q_k = a_k q_{k-1} + q_{k-2}$, where $q_k \geq 1$ for all k and $q_1 = a_1$. q_1 is a minimum for $q_1 = a_1 = 1$. For purposes of induction, assume the theorem is true for all $a \leq k \leq n$, i.e., a_1, \dots, a_k $^{\text{Min}}$ q_k is achieved for $a_1 = a_2 = \dots = a_k = 1$ for $k \leq n$. Now

$$q_{n+1} = a_{n+1} q_n + q_{n-1}$$

so that

$$\begin{aligned} \text{Min}_{a_1, \dots, a_k} q_{n+1} &= \left(\text{Min}_{a_{n+1}} a_{n+1} \right) \left(\text{Min}_{a_1, \dots, a_n} q_n \right) + \left(\text{Min}_{a_1, \dots, a_{n-1}} q_{n-1} \right) \\ &= 1(q_n | a_1 = a_2 \dots a_n = 1) + (q_{n-1} | a_1 = a_2 \dots a_{n-1} = 1) \\ &= q_{n+1} | a_1 = a_2 \dots a_{n+1} = 1 \end{aligned}$$

and the induction is complete.

Definition: The number b_n is called a *Fibonacci number* if $b_n = b_{n-1} + b_{n-2}$ with $b_0 = 1, b_1 = 1$.

Theorem 2: Let b_n be a Fibonacci number. If $b_{n-1} \leq p \leq b_n$ then, the upper bound on the number of partial quotients needed to form a continued fractions for $a/q, 1 < a < q$, is n .

Proof: If $1 = a_1 = a_2 = a_3 \dots$, then $q_0 = 1, q_1 = 1, q_2 = 1 + 1 = 2, q_3 = q_2 + q_1 = 2 + 1 = 3, q_4 = 3 + 2 = 5$, etc. These are the *Fibonacci numbers*, b_n for $(n = 0, 1, 2, \dots)$. Thus, if $q_n = q$ is a Fibonacci number, then $a_1 = a_2 = \dots a_n = 1$. Thus, by lemma, the largest value of n is achieved.

If $q_k = q$ for $b_{n-1} < q \leq b_n$, then, by the lemma, $b_k < q_k = q \leq b_n$. Hence $k < n$, and n is the upper bound of partial quotients to form continued fractions for $a/q, 1 < a < q$.

A simple example is now presented for finding the inverse element in $GF(q)$.

Example: Let $GF(q)$ be the field of integers modulo the Mersenne prime $q = 2^7 - 1$. Find an inverse element of 19 in $GF(127)$.

Let $a = 19$ and let $S = 19/127$. From the tabular form (Table A-1) when $k = n = 4$, one observes $r_4 = 0$. Thus, $S = S_4 = 19/127$. Hence, $q_3 = 20$. By (A-7), $a^{-1} \equiv -20 \equiv 107 \pmod{127}$ is the inverse element of 19. Since $b_4 = 5 < q_4 = 127 < b_{11} = 144$, by Theorem 2, the upper bound on the numbers of partial quotients is 11.

Table A-1. A computation of convergents to a continued fraction

k	$r_{k-2} = q_k r_{k-1} + r_k$	a_k	r_k	$p_k = a_k p_{k-1} + p_{k-2}$	$q_k = a_k q_{k-1} + q_{k-2}$	$S_k = p_k/q_k$
-2						
-1						
0	$19 = 0 \cdot 127 + 19$	0	19	$p_0 = 0 \cdot 1 + 0 = 0 = a_0$	$q_0 = 0 \cdot 0 + 1 = 1$	$S_0 = 0$
1	$127 = 6 \cdot 19 + 13$	6	13	$p_1 = 6 \cdot 0 + 1 = 1$	$q_1 = 6 \cdot 1 + 0 = a_1$	$S_1 = 1/6$
2	$19 = 1 \cdot 13 + 6$	1	6	$p_2 = 1 \cdot 1 + 0 = 1$	$q_2 = 1 \cdot 6 + 1 = 7$	$S_2 = 1/7$
3	$13 = 2 \cdot 6 + 1$	2	1	$p_3 = 2 \cdot 1 + 1 = 3$	$q_3 = 2 \cdot 7 + 6 = 20$	$S_3 = 3/20$
4	$6 = 6 \cdot 1$	6	0	$p_4 = 6 \cdot 3 + 1 = 19$	$q_4 = 6 \cdot 20 + 7 = 127$	$S_4 = 19/127$

Appendix B

Let a and b be relatively prime and let A be a cyclic $ab \times ab$ matrix. In this Appendix, it will be shown in the following theorem that there exists a permutation π of the rows and columns so that A can be partitioned into blocks of $b \times b$ cyclic matrices, such that the blocks form a $a \times a$ cyclic matrix.

Theorem 1: Let a and b be relatively prime. Let A be the cyclic $ab \times ab$ matrix given by

$$A_{(x,y)} = f_{(x+y \bmod ab)}, \quad 0 \leq x, y < ab$$

If π is a permutation of $\{0, 1, \dots, ab - 1\}$, let

$$B_{(x,y)} = A_{(\pi(x), \pi(y))}$$

Then there exists a permutation π such that, if B is partitioned into $b \times b$ submatrices, then each submatrix is cyclic and the submatrices form a $a \times a$ cyclic matrix.

Proof: Let $Z_n = \{0, 1, \dots, n - 1\}$ be the additive group of integers modulo n . By the Chinese Remainder Theorem, the mapping $\alpha: Z_{ab} \rightarrow Z_a \times Z_b$ given by

$$\alpha(x) = (x \bmod a, x \bmod b), \quad x \in Z_{ab}$$

is an isomorphism. Define also the mapping $\beta: Z_{ab} \rightarrow Z_a \times Z_b$ given by

$$\beta(x) = \left(u = \frac{x - v}{b}, v \equiv x \bmod b \right)$$

for $x \in Z_{ab}$. Then β is a one-to-one and onto mapping, and $\beta^{-1}(u, v) = bu + v, u \in Z_a, v \in Z_b$.

Let $\pi = \alpha^{-1}\beta$. Then π is a permutation of Z_{ab} . Let $\beta_{ij}, i, j \in Z_a$, be the $(i, j)^{\text{th}}$ $b \times b$ submatrix of B . Then for $v, w \in Z_b$,

$$\begin{aligned} B_{ij}(v, w) &= B_{(bi+v, bj+w)} \\ &= B_{(\beta^{-1}(i, v), \beta^{-1}(j, w))} \\ &= A_{(\pi\beta^{-1}(i, v), \pi\beta^{-1}(j, w))} \\ &= A_{(\alpha^{-1}(i, v), \alpha^{-1}(j, w))} \\ &= f_{((\alpha^{-1}(i, v) + \alpha^{-1}(j, w)) \bmod a \cdot b)} \end{aligned}$$

Since α^{-1} is an isomorphism

$$\beta_{ij}(v,w) = f_{(\alpha^{-1}((i+j) \bmod a, (v+w) \bmod b))}$$

If we fix i and j in the above eq., it is evident that the $(i,j)^{\text{th}}$ $b \times b$ submatrix of B is cyclic matrix. Similarly, by fixing v and w , the submatrices B_{ij} form a $a \times a$ cyclic matrix.

Example: Let $a = 2, b = 3$. Then $\alpha(x) = (x \bmod 2, x \bmod 3)$ and $\alpha^{-1}(u,v) = 3u + 4v \bmod 6$. Also $\beta(x) = (u = (x - v)/3, v \equiv x \bmod 3)$ and $\beta^{-1}(u,v) = 3u + v$. Finally $\pi = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 4 & 2 & 3 & 1 & 5 \end{pmatrix}$ or the 2-cycle (14).

ORIGINAL PAGE IS
OF POOR QUALITY

N78-11153

GCF Central Performance Monitor — Functional Description

N. C. S. Lau
DSN Data Systems Section

This article describes the Ground Communications Facilities (GCF) Central Performance Monitor (CPM) which gathers information on the operational status of the Central High Speed and Wideband Assemblies. This monitor information is formatted and forwarded to the Central Communication Monitor Processor. The CCM Processor, in turn, provides near real time displays and printouts of the status information and any resulting alarm indications. The ability of the GCF to quickly detect and correct circuit failures will be enhanced by the addition of the CPM.

I. Introduction

The Central Performance Monitor (CPM) is an integrated part of the Central Communication Monitor Assembly (CCMA) within the Ground Communication Facilities (GCF) Monitor and Control Subsystem located in the Central Communication Terminal (CCT) at JPL. The GCF provides for interconnecting the Deep Space Stations with the Space Flight Mission Operations Center and the Network Operations Control Center for the DSN. The Mission Operations Center (MOC) processes the data acquired by the Deep Space Network Stations and issues commands to be sent to the spacecraft through the Stations. The Network Operations Control Center processes data to determine network status and issues messages for network control purposes. These data interchanges are conducted in real-time and must be accomplished with high reliability.

The CCMA is provided with the status and parameters by the CPM. The CCMA, in turn, provides Operations with cen-

tralized real-time display on the performance of all GCF circuits. Prompt correction can then be accomplished for faulty data links. The CPM, therefore, plays a key role in accumulating real-time performance data of the communication channels.

II. Hardware Functional Description

The Central Performance Monitor (CPM) provides the hardware function of monitoring circuit performance signals for up to four wideband channels and up to eighteen High Speed Data Channels (see Fig. 1). The performance signals monitored can be grouped into three categories: Signals to be sensed, signals to be sensed and held, and signals to be accumulated.

The CPM has two identical CCM computer ports. Each port provides for a 14 line interface. The 14 lines of each port include a unidirectional stimulus from device (STD), a unidirectional stimulus from computer (STC), a bidirectional

response line (RSP), a bidirectional ready line (RDY), eight bidirectional data lines (D1 through D8), and two bidirectional function lines (F0 and F1). All bidirectional lines are used to both transmit and receive. A special power sense (PS) line is routed to the computer to inform the computer whether the CPM power is on or off.

The CPM is switch selectable to be brought on-line with either of the two CCMA computers. Upon request from the on-line CCMA computer, the CPM will arrange and forward all the gathered information in a format acceptable to the computer through the 14-line interface.

The CPM reporting period can be varied, and is defined by software. However, the reporting period should not extend beyond the time that it takes to accumulate a maximum count of 255 of either Data Block Detected signals or Block Error Detected signals from any one of the four wideband channels. Figure 2 shows the functional block diagram for the CPM.

III. Information Monitored

The specific signals monitored for each of the four (4) wideband channels are carrier ON/OFF, SEARCH, NO MATCH, SYNC FAIL, REQUEST-TO-SEND, REQUEST-TO-SEND FAIL, ON-SITE-COMPUTER RECEIVING, DATA BLOCK DETECTED AND BLOCK ERROR DETECTED.

The carrier ON/OFF and SEARCH are sense only signals. The status of these signals are reported as "ON" only if they are "ON" at the instant of computer interrogation.

NO MATCH, SYNC FAIL, REQUEST-TO-SEND, REQUEST-TO-SEND FAIL, and ON-SITE COMPUTER-RECEIVING are sense and hold signals. An occurrence of an event during the interrogation interval is stored and reported when interrogated.

The Data Block Detected and Block Error Detected signals are accumulated respectively during the interrogation interval and the count results reported to the computer.

For the High Speed Subsystem, the only signal that is monitored by the CPM is the Data Set Carrier ON/OFF information which is a sense only signal. However, other monitor information for the High Speed, similar to that gathered for the Wideband, is obtained through the Error Detection Correction (EDC) and the High Speed Switch (HSW) Assemblies of the GCF.

IV. Conclusion

The CCM computer accepts the monitored data as provided by the CPM, EDC, and HSW Assemblies and calculates each circuit throughput and displays the status and alarm, if any, on the display and other data recording devices for the GCF technical controller use.

The combining of the High Speed and Wideband circuit performance information into one centralized display eliminated display switching and promotes speedier operator responses to alarmed circuits. Thus, any channel anomalies and faulty circuits can be rapidly located and prompt substitution or service can be performed.

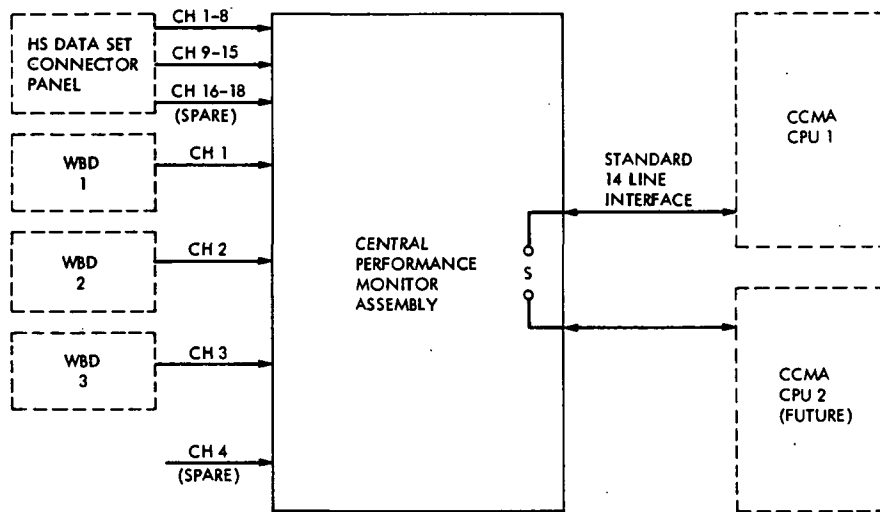


Fig. 1. Central performance monitor interfaces

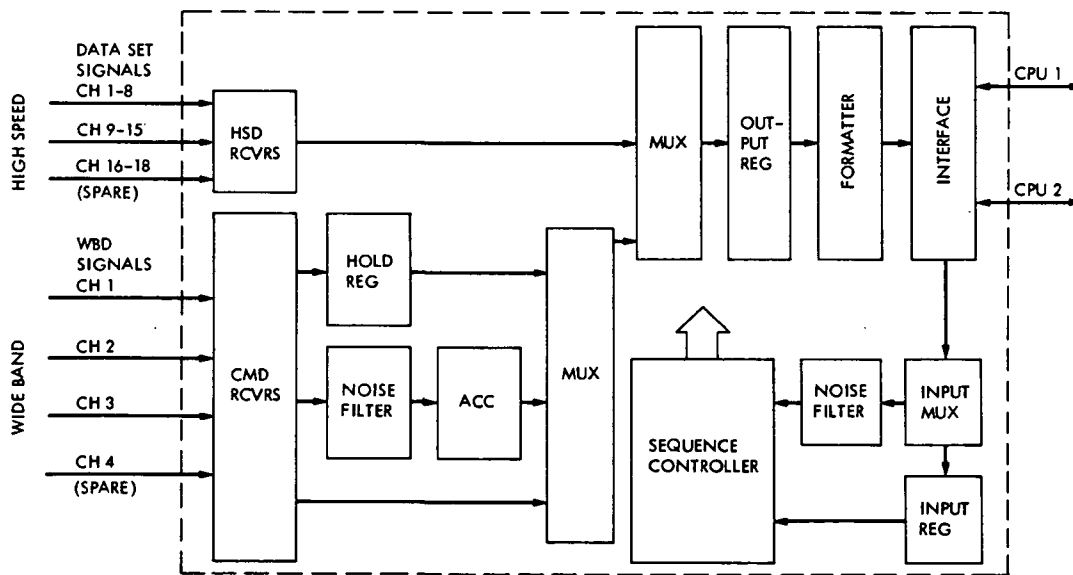


Fig. 2. Central performance monitor functional block diagram

222
N78-11154

A Thermodynamic Analysis of a Solar-Powered Jet Refrigeration System

F. L. Lansing and V. W. Chai
DSN Engineering Section

The article describes and analyzes a method of using solar energy to drive a jet refrigeration system. A new technique is presented in the form of a performance nomogram combining the energy and momentum equations to determine the performance characteristics. A numerical example, using water as the working fluid, is given to illustrate the nomogram procedure. The resulting coefficient of performance was found comparable with other refrigeration systems such as the solar-absorption system or the solar-Rankine turbocompressor system.

I. Introduction

As part of the Deep Space Network energy conservation project made for the Deep Space Communication Complex (DSCC) at Goldstone, California, many energy-conservation money-saving ideas are being studied to reduce the complex dependence on fossil fuel or electric-powered devices. The existing heating, ventilation, and air-conditioning equipment (HVAC) presents a sizable share of the total energy consumption of the complex, and the concept of utilizing the abundant solar energy falling on the site is under investigation.

The present state-of-the-art of solar-powered cooling devices is still limited to several options. Examples are: (1) a solar-powered Rankine engine driving a vapor compression refrigeration machine, (2) solar-powered photovoltaic array (concentrated on nonconcentrated) driving a vapor compression refrigeration machine, and (3) a solar-powered absorption refrigeration. The first option employs two rotating parts such

as turbines (or engines), and compressors (excluding the pump) that have to be specially designed to adapt the low temperature nature of solar energy. The second option, though directly generates electricity with less mechanical parts, is neither economically competitive nor has the same conversion efficiency with the first. The third option possesses the highest coefficient of performance, but has some size, maintenance, and operation difficulties.

One of the least commonly used refrigeration techniques today is the jet refrigeration system (Refs. 1, 2 and 3). This system has not enjoyed wide application despite its advantages of simplicity, ruggedness of design, vibration-free, reliability, and its low operation and maintenance cost. Coupling this system with solar-energy as a driving source instead of conventional gas-fired boilers adds a new option to solar-powered cooling devices. This work intends to introduce the concept, describe the relevant parameters, and present the new analytic methodology used in performance determination.

**ORIGINAL PAGE IS
OF POOR QUALITY**

II. System Description

The basic concept of a solar-powered vapor compression refrigeration system, as illustrated in Fig. 1, is to use a solar-Rankine power cycle to drive a vapor compression refrigeration cycle. Solar energy is used to vaporize the working fluid of the power cycle to saturated vapor at the boiler. Nonconcentrating flat-plate solar collectors or medium-concentration types can be used for this task to provide fluid temperatures that range from 90 to 160°C. For most working fluids, this range of temperatures is sufficient for vaporization. The vapor expands in the turbine and produces the mechanical work that is needed to drive the refrigeration compressor. The combination of the turbine and the compressor on a single shaft is sometimes called a turbo-compressor unit (Ref. 4).

The working fluids and the temperatures used in both cycles so far can be independent of each other. In each cycle, the choice of the working fluid depends on performance, stability, economics, and safety factors.

Figure 2 shows a modified solar-powered vapor compression refrigeration system where both power and refrigeration cycles are employing the same working fluid and have a common condenser. The common condenser acts for both cycles as a heat sink at a uniform temperature. This configuration with a single working fluid has been recommended (Ref. 5) for its compactness and low maintenance cost benefits, though it does not have the choice of many working fluids due to chemical and thermal stability in the combined cycle.

The only role of the turbo-compressor unit (Figs. 1 and 2) is to expand in the turbine the vapor from the boiler pressure down to the condenser pressure and to compress in the compressor the vapor from the evaporator pressure up to the condenser pressure, thus exchanging two equal but opposite mechanical operations. This energy exchange, from thermodynamics viewpoint has a null effect on the overall cycle performance. The same function of the turbocompressor can be replaced and made simply by using the ejector scheme as shown in Figs. 3 and 4. The convergent-divergent nozzle replaces the turbine function in converting the enthalpy drop between the boiler pressure and the evaporation pressure to kinetic energy. In practice, the pressure at the nozzle exit (state 4) is made slightly lower than the evaporator pressure (state 5) to overcome the pressure drop in the line. In the mixing chamber, entrainment with momentum exchange occurs between the vapor with supersonic speed leaving the nozzle and the vapor with subsonic speed leaving the evaporator. The mixture (state 6), still with supersonic velocity, will emerge into the diffuser section where it converts its kinetic

energy back into an enthalpy increase, thus raising its pressure to the condenser pressure (state 7). The diffuser replaces the compressor function. The condensate from the condenser (state 1) is divided to two branches; one branch goes to the power cycle loop through the pump (state 2) and the solar collector/boiler section to state (3) as vapor, and the other branch goes to the refrigeration cycle loop through an expansion valve (state 8) to the evaporator.

Figure 5 presents the sequence of events for the jet refrigeration scheme on the pressure-enthalpy diagram. Flow irreversibilities in the nozzle and the diffuser occur through the deviation of the isentropic states (state 4_{isen} and state 7_{isen}) from the actual ones (states 4 and 7).

The jet-refrigeration system has many advantages besides operating without gross moving parts (except for the small circulating pump). It is vibration- and noise-free, relatively simple and inexpensive to manufacture with a well-developed technology of nozzles, diffusers, and heat exchangers. The system is low in operation and maintenance cost and its performance is highly competitive with solar-powered turbo-compressor units and solar powered absorption refrigeration units, as will be shown later.

III. Assumptions and Governing Equations

The adiabatic control volume including the nozzle, the mixing chamber, and the diffuser is shown schematically in Fig. 4. Some assumptions are made in the analysis and are listed as follows:

- (1) The velocity of the fluid entering the nozzle (state 3) is negligible.
- (2) The velocity of the fluid leaving the evaporator (state 5) is negligible.
- (3) The velocity of the fluid leaving the diffuser (state 7) is negligible.
- (4) The pressure drop through the heat exchangers, and lines is negligible.
- (5) Saturated conditions for states 1, 3, and 5 are assumed.
- (6) Expansion and compression processes in the nozzle and the diffuser, respectively, are adiabatic.
- (7) Heat exchange processes in the evaporator and the condenser are ideal with no thermal losses by radiation.

According to assumptions 1, 2, and 3, the total and static enthalpy of states 3, 5 and 7, as depicted in Fig. 6, will coincide, i.e.,

$$\left. \begin{aligned} H_3 &= H_{3t} \\ H_5 &= H_{5t} \\ H_7 &= H_{7t} \end{aligned} \right\} \quad (1)$$

Applying the first law of thermodynamics on the control volume of Fig. 4 gives:

$$M_1 H_3 + M_2 H_5 = (M_1 + M_2) H_7 \quad (2)$$

where M_1 is the mass flow rate for the power cycle and M_2 is the mass flow rate for the refrigeration cycle. Defining a mass ratio R as

$$R = \frac{M_1}{M_2} \quad (3)$$

Equation (2) can be rewritten in either one of the following forms:

$$\left. \begin{aligned} R &= \frac{H_7 - H_5}{H_3 - H_7} \\ \left(\frac{R}{R+1} \right) &= \frac{H_7 - H_5}{H_3 - H_5} \end{aligned} \right\} \quad (4)$$

Equation (4) indicates that the total enthalpy H_7 is located between the enthalpies H_3 and H_5 and divides their enthalpy difference ($H_3 - H_5$) inversely by the ratio R , as shown in Fig. 6. Furthermore, applying the first law of thermodynamics and the irreversibility relationships for the nozzle, gives

$$\frac{V_4^2}{2} = (H_3 - H_4) = (H_3 - H_{4, \text{isen}}) \times \eta_n \quad (5)$$

and for the diffuser

$$\frac{V_6^2}{2} = H_7 - H_6 = \frac{1}{\eta_d} (H_{7, \text{isen}} - H_6) \quad (6)$$

where η_n and η_d are the isentropic efficiencies for the nozzle and the diffuser, respectively.

The momentum equation in the constant pressure mixing chamber alone would give:

$$\left. \begin{aligned} M_1 \cdot V_4 &= (M_1 + M_2) \cdot V_6 \\ \frac{V_6^2}{2} &= \left(\frac{R}{1+R} \right)^2 \cdot \frac{V_4^2}{2} \end{aligned} \right\} \quad (7)$$

The thermodynamic properties of states 1, 3 and 5 are completely determined once the three independent temperatures of the boiler, evaporator, and condenser are given. The solution of the set of Eqs. (4), (5), (6), and (7) requires the use of one of the property diagrams together with the design values of η_n and η_d . Instead of the usual trial and error procedure, a performance nomogram will be introduced to match the operating conditions.

A first manipulation of Eqs. (5), (6) and (7) give the following relationship:

$$(H_{7, \text{isen}} - H_6) = \left(\frac{R}{R+1} \right)^2 \cdot \eta_n \cdot \eta_d (H_3 - H_{4, \text{isen}}) \quad (8)$$

The left-hand side of Eq. (8) represents the ideal compression work needed if the diffuser were replaced by an adiabatic compressor. On the other hand, the enthalpy difference ($H_3 - H_{4, \text{isen}}$) represents the ideal expansion work obtained if the nozzle were replaced by an adiabatic expander. The unique turbocompressor coupling effect made by the jet refrigeration scheme is clearly shown by the extra quantity $(R/R+1)^2$ in Eq. (8). The fact that the quantity $(R/R+1)^2$ appears is due to the entrainment phenomenon in the mixing chamber and the momentum exchange.

The overall coefficient of performance (COP) for the jet refrigeration system is defined as

$$COP = \frac{\text{refrigeration effect}}{\text{external energy input}} \quad (9)$$

The external energy added to the system (Fig. 3) is composed of both the heating part from state (2) to state (3) and the pump work from state (1) to state (3). Accordingly,

$$COP = \frac{(H_5 - H_1)}{R (H_3 - H_1)} \quad (10)$$

IV. Carnot's Cycle Idealization

The maximum coefficient of performance of the jet refrigeration scheme can be determined using Carnot's principles as shown in Fig. 7. The Carnot's power cycle shown works between the boiler temperature (T_B) and the condenser temperature (T_c) as its limits. The Carnot's refrigeration cycle works between the condenser temperature (T_c) and the evaporator temperature (T_E) as its limits. Both cycles exchange exactly the same amount of mechanical work, which can be written either as

$$W = Q_B \left(1 - \frac{T_c}{T_B} \right) \quad (11)$$

or

$$W = Q_E \left(\frac{T_c}{T_E} - 1 \right)$$

where Q refers to the amount of heat exchanged as shown in Fig. 7. The maximum coefficient of performance can then be written as

$$(COP)_{\max} = \frac{Q_E}{Q_B} = \frac{(T_B - T_c)}{(T_B)} \cdot \frac{(T_E)}{(T_c - T_E)} \quad (12)$$

Combining Eqs. (10) and (12) together, will give the minimum mass ratio R for a given set of enthalpies (H_1 , H_3 , and H_5) as

$$R_{\min} = \frac{(H_5 - H_1)}{(H_3 - H_1)(COP)_{\max}} \quad (13)$$

V. Performance Nomogram

This is a graphical presentation of the system equations presented in Section III including energy, momentum, and continuity equations. The nomogram is illustrated in Fig. 8 and is made up of four curves, one for each quadrant. Knowing the three independent temperatures of the system (for the boiler, the condenser and the evaporator) together with the fifth assumption of saturated conditions, the thermodynamic properties of states 1, 3 and 5 are completely determined. Curve 1 can be constructed using the ejector energy equation, Eq. (4), to plot the mass ratio R as an ordinate versus H_7 as an abscissa. Equation (4) shows that R ranges from zero (at $H_7 = H_5$) to infinity (at $H_7 = H_3$). The Carnot's principle further limits the mass ratio R to only above or at least equal to R_{\min} as given by Eq. (13). Accordingly, curve 1 will be divided to two regions: an upper region (above R_{\min}), where the performance does not violate the second law of

thermodynamics, and the lower region (below R_{\min}), which is forbidden. Combining the energy and momentum Eqs. (4), (5), (6), and (7) to eliminate the mass ratio R , the resulting equation will be in the form

$$H_6 = H_7 - \left(\frac{H_7 - H_5}{H_3 - H_5} \right)^2 \cdot \eta_n \cdot [H_3 - H_{4, \text{isen}}] \quad (14)$$

The construction of the set of curves (2) can then follow using curve 1 and Eq. (14), with the nozzle efficiency (η_n) as a parameter. Curve (2) is a plot of the enthalpy H_6 versus the mass ratio R as shown in Fig. 8. Starting with a suitable value of H_7 at a given η_n , the value of R is determined graphically from curve 1, or numerically by Eq. (4), and together with Eq. (14) a point can be located on curve 2. The nozzle efficiency can then be varied and the set of curves (2) will be completed.

The above method in constructing curve 2 is found more straightforward in computations than other methods. For example, if the enthalpy H_7 was eliminated from Eqs. (4), (5), (6), and (7) instead of R , the following quadratic equation would be solved for a given value of H_6 instead of Eq. (14),

$$R^2 (H_6 - H_3 + \eta_n (H_3 - H_{4, \text{isen}})) - R (H_3 + H_5 - 2H_6) - (H_5 - H_6) = 0 \quad (15)$$

The set of curves (3) in Fig. 8, is plotted with H_6 and H_7 as coordinates, using Eq. (6) and one of the property diagrams with the diffuser efficiency (η_d) as a parameter. The nomogram is then completed by drawing a straight line in the fourth quadrant that passes through the origin with a 45-deg slope.

The performance characteristics are now ready to be determined for the given boiler, condenser, and evaporator temperatures and the nozzle and diffuser isentropic efficiencies. The path may start, for example, from point A (Fig. 8) at an arbitrary value of H_7 , down to curve (4), horizontally to meet curve (3) at point C at the selected diffuser efficiency, up to meet curve (2) at point D at the selected nozzle efficiency, and horizontally to meet curve (1) at point E. At point E the enthalpy H_7 should match the first value chosen at point A, otherwise the loop is repeated until the solution converges.

VI. Nomogram Case Example

The following is a numerical example selected for a jet refrigeration system using water as the working fluid. The

following five independent input conditions are sufficient to describe the system. Numerically they are assigned as

- (1) Boiler temperature 93.33°C (200°F)
- (2) Condenser temperature 26.67°C (80°F)
- (3) Evaporator temperature 4.44°C (40°F)
- (4) Nozzle efficiency 0.80
- (5) Diffuser efficiency 0.80

The maximum (COP) is calculated from Eq. (12) as 2.27 and the minimum mass ratio (R_{\min}) as 0.414. Curves 1, 2, 3 and 4 are constructed as shown in Fig. 8. The coefficient of performance for the above example converges to 0.465 corresponding to a mass ratio of R of 2.02. Also, from the nomogram, the coefficient of performance drops to 0.289 corresponding to a mass ratio of 3.25 if the nozzle and diffuser efficiencies become 0.7 each. On the other hand, the coefficient of performance increases to 0.618 corresponding to a mass ratio of about 1.52 if the nozzle and diffuser efficiencies become 0.9 each. The above indicates how sensitive the performance is to the quality of manufacturing the nozzle and the diffuser. Also, it shows that the performance is comparable to other heat-powered refrigeration techniques.

VII. Summary and Remarks

The following points can be made from this first phase study of solar-powered jet refrigeration and the given example of performance:

(1) A new technique is presented in the form of a performance nomogram combining energy and momentum equations to describe the performance characteristics.

(2) The coefficient of performance obtained in the cited example with a nozzle efficiency of 0.80 and a diffuser efficiency of 0.80 was 0.465. This compares favorably with about 0.6 for the solar-powered absorption unit and with about 0.5 for solar-powered Rankine driving a vapor-compression refrigerator. The jet refrigeration performance has improved to about 0.62 when the nozzle and diffuser efficiencies increased to 0.90. This means that an improvement of the product efficiency from 0.64 to 0.81 (an increase of 27%) has produced a performance increase of about 50%.

(3) The numerical example given can also be very helpful in sizing a 3-ton refrigeration unit (36,000 Btu/hr) to fit a conventional residence house. Based on the mass flow ratio R of 2.02, the mass flow rates entering the evaporator and the nozzle are 15.84 kg/hr and 32.48 kg/hr, respectively, and the boiler heat rate would be about 23 kWh_t (78,610 Btu/hr). If the local solar radiancy is taken as 0.8 kw_t/m² (~254 Btu/hr-ft²), the collector area needed at 45% collector efficiency would be about 64 m² (~690 ft²); an area that is adequate for most south-facing roofs. The ejector unit itself for this type of flow conditions does not exceed 50 cm in length according to manufacturer's specifications.

(4) Performance optimization studies are underway to analyze the effects of different types of working fluids, or the effects of different operating temperatures for the solar collector/boiler, condenser and evaporator. These will be reported later as a second phase of the program.

ORIGINAL PAGE IS
OF POOR QUALITY

Definition of Terms

COP	Coefficient of performance
COP_{max}	Maximum coefficient of performance
H	Enthalpy
M_1	Mass flow rate for the power cycle
M_2	Mass flow rate for the refrigeration cycle
R	Mass flow ratio
R_{min}	Minimum flow ratio by Carnot principle
T	Temperature
V	Velocity
η	Efficiency
Subscripts	
B	Boiler
c	Condenser
d	Diffuser
E	Evaporator
n	Nozzle
t	Total (stagnation)
isen	Isentropic

References

1. Anderson, H., "Assessment of Solar Powered Vapor Jet Air-conditioning Systems," presented at 1975 International Solar Energy Congress and Exposition (ISES) held at UCLA, Los Angeles, California, p. 408.
2. Threlkeld, J. L., *Thermal Environmental Engineering*, Sect. 6. Prentice-Hall, Inc., New Jersey, 1962.
3. *ASHRAE Guide and Data Book*, "Steam Jet Refrigeration Equipment," Chap. 12, pp. 133-138, 1967.
4. Barber, R. E., "Solar-air Conditioning Systems Using Rankine Power Cycles – Design and Test Results of Prototype Three-ton Unit," in *Proceedings of the Institute of Environmental Science*, Vol. 1, pp. 170-179, 1975.
5. Biancardi, F. R., et al., "Design and Operation of a Solar-powered Turbocompressor Air-conditioning and Heating System," in *IECEC '75 Record*, pp. 186-194.

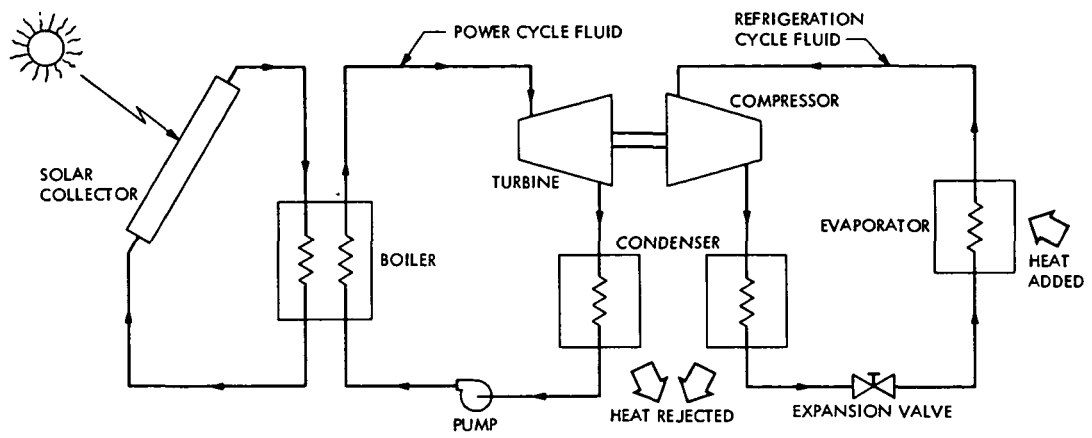


Fig. 1. Double fluid turbocompressor system

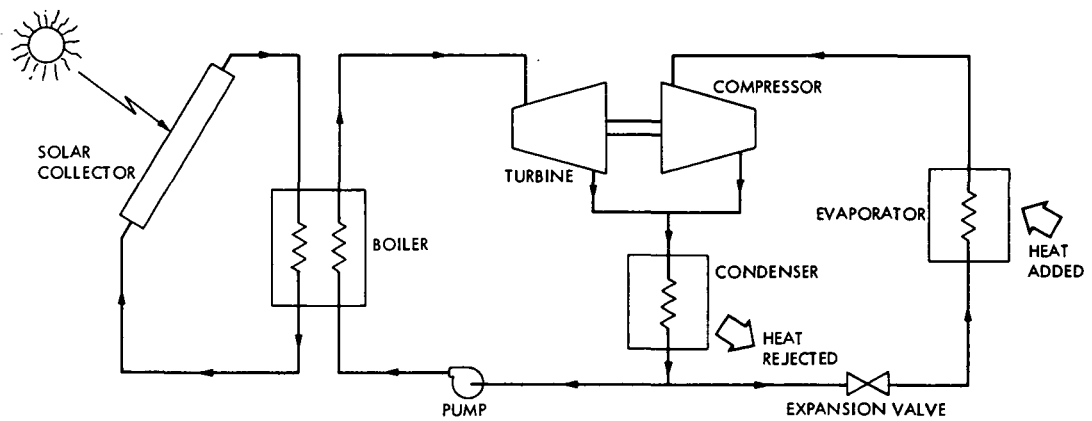


Fig. 2. Single fluid turbocompressor system

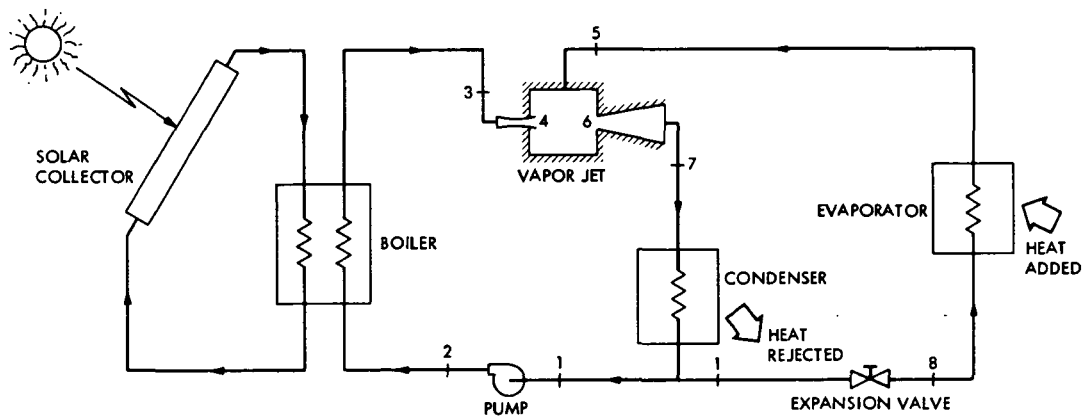


Fig. 3. Vapor jet refrigeration system

ORIGINAL PAGE IS
OF POOR QUALITY

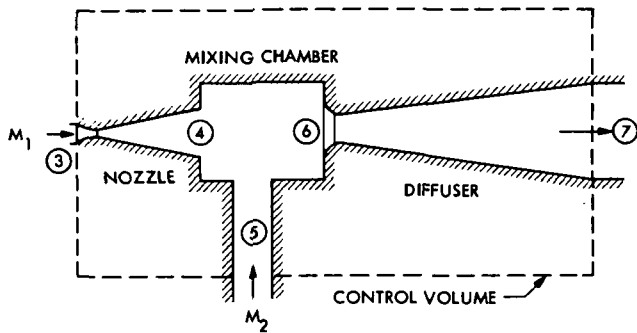


Fig. 4. Ejector control volume

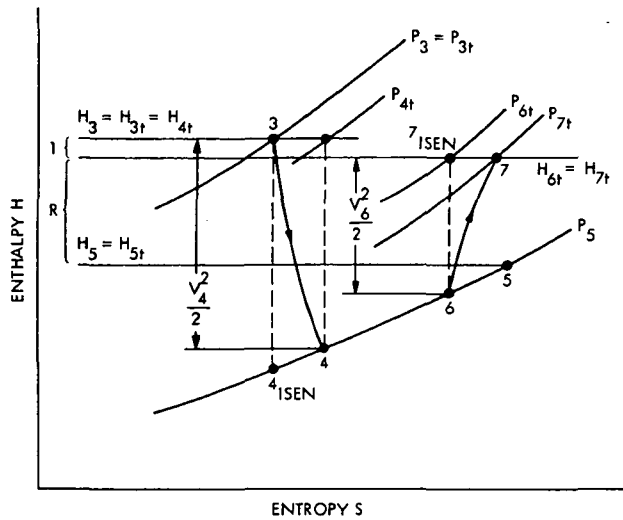


Fig. 6. Ejector presentation on the enthalpy-entropy (Mollier) diagram

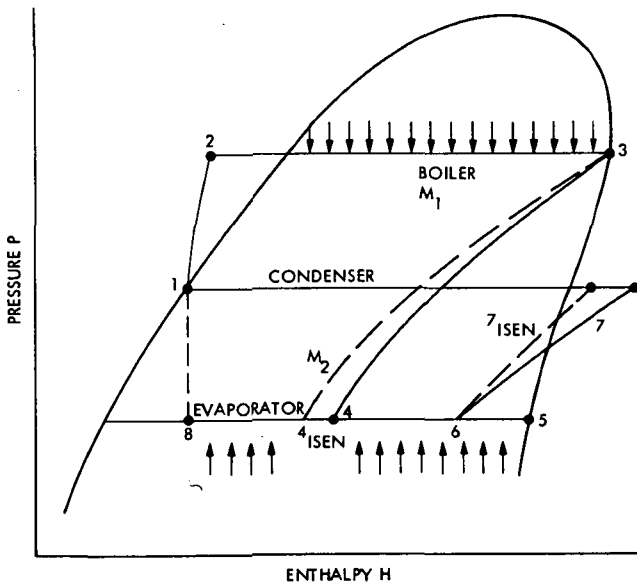


Fig. 5. Pressure-enthalpy diagram showing the complete jet refrigeration cycle

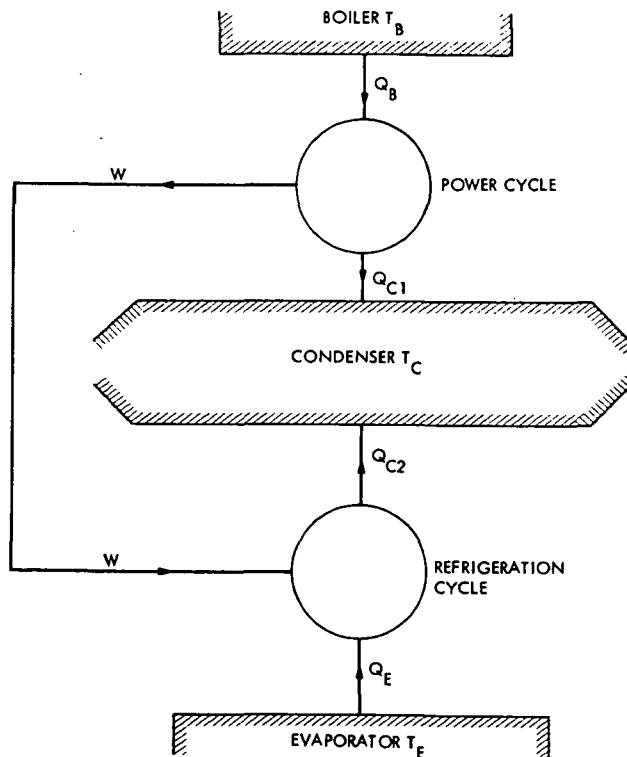


Fig. 7. Carnot cycle idealization for the jet refrigeration scheme

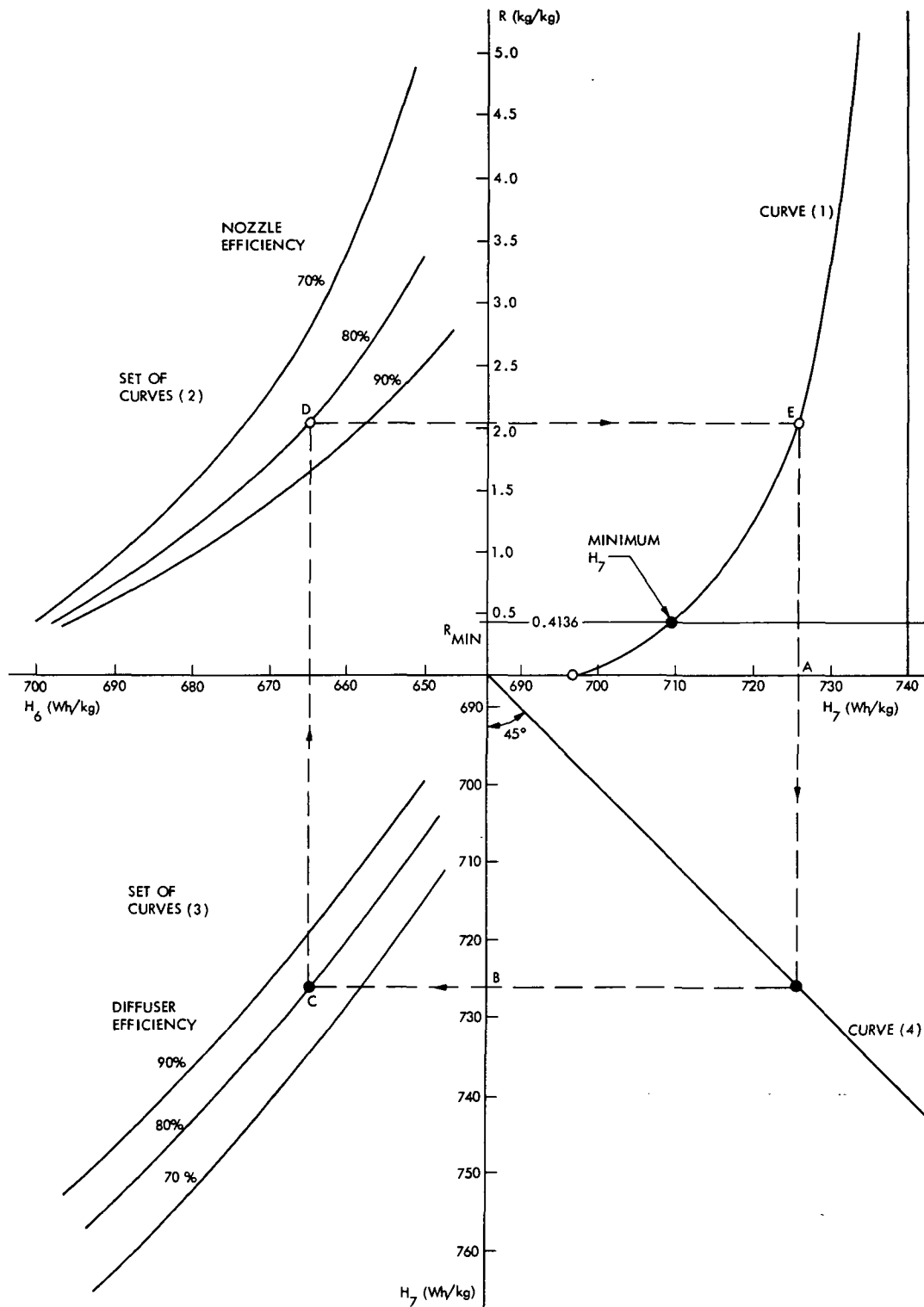


Fig. 8. Performance nomogram for the given operating conditions

ORIGINAL PAGE IS
OF POOR QUALITY

N78-11155

Estimated Displacements for the VLBI Reference Point of the DSS 14 Antenna

H. McGinness
DSN Engineering Section

The maximum displacement of the defined VLBI reference point is shown to be primarily a function of the ambient air temperature.

I. Introduction

When using the 64 meter azimuth-elevation antenna at DSS 14 for very long base interferometry (VLBI), it is necessary either to establish that a defined reference point will have displacements less than a tolerable amount, or to be able to evaluate these displacements as functions of measurable quantities. These quantities would include the environmental quantities of temperature and wind together with the runouts of the antenna axes. The reference point is defined as the nominal intersection of the azimuth and elevation axes. This article describes three different sets of measurements which relate to the thermal displacement of the reference point and shows how the total displacement magnitude can be estimated from the ambient temperature during the VLBI experiment.

II. Description of the Experiments

A. Alidade Leg Temperatures

The surface temperatures of the front right alidade leg, together with the ambient air temperature, were measured during the period of April 1976 to March 1977. On most days, three measurements were made near the hours 0600, 1500, and 2300. Six surface thermometers were distributed along an

inner surface of the leg (so as to avoid direct sunlight). Generally, the variation among the thermometer readings was small. A more detailed description of the gage locations can be obtained from Reference 1.

The vertical displacement of the reference point caused by ambient temperature changes will be proportional to the average temperature of the alidade legs. Therefore, it is necessary to compare the ambient air temperature range to the alidade leg surface temperature range and then subsequently estimate the relationship between the leg surface temperature and the leg average temperature. From the alidade leg temperature records, those showing large daily ambient ranges are displayed in Table 1. Table 1 shows the ambient difference between the hours 0600 and 1500, the leg surface temperature difference between these same hours, and the maximum observed range of the leg temperature. The last two columns of the table show the ratios of leg temperature ranges to ambient temperature range. It is believed that the minimum and maximum ambient temperatures usually occur near the hours 0600 and 1500 respectively. The conclusion is that on days having large ambient temperature excursions, the leg surface temperature usually undergoes somewhat less excursion. Since the average of the last column of Table 1 is 75%, this value will be applied subsequently when estimates for total displacements are made.

B. Displacements of Reference Point

During the operational shutdown of DSS 14 in May and June 1977, an experiment was conducted which directly measured most of the reference point displacement occurring over the time of the experiment. Figure 1 shows the configuration of the test apparatus. Two spring loaded styluses at right angles to each other were mounted in a frame positioned only 23 centimeters above the reference point. This frame was fixed to an inverted tripod, which in turn, was temporarily clamped to the frame of the intermediate reference structure Optical Assembly, which is firmly attached to the antenna dish back-up structure. Two pieces of carbon paper were mounted on another bracket attached to the Master Equatorial mirror frame, which in turn is firmly mounted on the top of the shielded instrument tower. Each piece of carbon paper was covered with plain white paper and was perpendicular to its stylus.

This experiment was begun at hour 1630 on May 27, 1977. The antenna elevation and azimuth angles were held at 88.95° and 246° respectively. The styluses were carefully retracted, the papers removed, and examined at hour 0940 on June 2, 1977. During this seven-day period, the ambient temperature ranged from 12° to 35° celsius. At DSS 12, approximately 18 kilometers away, there were winds of 13 meters per second with gusts up to 19 meters per second within the time period. Examination of the stylus traces showed that the total horizontal excursion along the Y axis was 2.0 mm, the total horizontal excursion along the X axis was 1.0 mm, and the total vertical excursion along the Z axis was 2.50 mm.

The styluses were reset against new paper at hour 0945 of June 2, 1977. The traces were examined at hour 1215 of June 4, 1977. During this 2-day period, the ambient air temperature ranged from 17° to 36° celsius. The total horizontal excursion along the Y axis was 1.75 mm, the total horizontal excursion along the X axis was 1.0 mm, and the total excursion along the vertical Z axis was 2.50 mm. Wind records were not obtained during this period; however, persons at the antenna said the winds were weak.

Daily ambient temperature ranges of 16.7, 16.7, 17.7, and 17.7°C occurred during the displacement experiment.

Between the hours 1515 and 1635 of June 4, 1977, the antenna was rotated a full turn about the azimuth axis. The wind during this period did not exceed 4.5 meters per second. The antenna was started and stopped several times and a noticeable motion of the vertical stylus was observed at each acceleration and estimated to be ± 0.25 mm. The vertical stylus was located approximately 7 mm from the center of rotation and produced a closed trace on the paper. Perfect antenna bearings would have produced a 7 mm radius perfect circle.

The actual trace varied approximately 0.50 mm from the best-fitted circle. The conclusion is that the combined effect of imperfect radial and hydrostatic thrust bearings produces a horizontal motion of ± 0.50 mm.

C. Instrument Tower Temperatures

Instrument tower temperatures were recorded in December 1966 as part of a tower motion study (Ref. 2). Over a 36-hour period, the ambient air varied from 4.0°C to 13.0°C and the lower portion of the steel tower varied from 17.1 to 18.9°C . During this time, the greatest temperature difference across 180° of the tower was 0.54°C and this figure was based on the average of three differential thermocouples at various elevations on the tower. Also, during this time the angular displacement of the tower top had a total excursion of 6.7 arc seconds. The experimentally determined relationship between the angular displacement, θ in arc seconds, and the horizontal displacement, X in millimeters, of the tower top is, from Ref. 3, $X = 0.0613\theta$. This relationship produces a horizontal displacement of 0.41 mm when the angular displacement is 6.7 arc seconds.

If it is assumed that there is a temperature difference ΔT of 0.54°C across the diameter of the tower and that the temperature is linear with respect to the radius, R , of the tower, the horizontal displacement, X , may be estimated by the following equation:

$$X = \frac{\beta \ell^2 \Delta T}{2R} \quad (1)$$

where β is the linear expansion coefficient, and ℓ is the length of the steel portion of the tower.

For $\beta = 10.8 \times 10^{-6}$ per degree Celsius, $\ell = 23$ m, $R = 3.5$ m, equation (1) yields the value of 0.44 mm which checks well with the above value of 0.41 mm.

The ratio, R , between the observed change in instrument tower temperature and the change in ambient air temperature is, using the above values:

$$R = \frac{18.9 - 17.1}{13 - 4} = 0.20 \quad (2)$$

For subsequent calculations, it will be assumed that the change in tower temperature will be 20% of the change in ambient air temperature. Furthermore, it will be assumed that the horizontal displacement of the instrument tower top will be 0.41 mm multiplied by the ratio of ambient air temperature

ORIGINAL PAGE IS
OF POOR QUALITY

change in degrees Celsius to the value 9 which was the ambient air temperature change during these observations. Such calculations are necessary to modify the measured displacements of the VLBI reference point so as to have the displacements referenced to earth rather than to the top of a moving instrument tower. It is practical to determine experimentally the displacements of the instrument tower top with respect to the bottom of the tower foundation, but this has not yet been done.

III. Heat Transfer Analysis

In order to obtain a relationship between the average temperature of the alidade leg and the ambient air temperature, it is necessary to obtain solutions for two transient heat transfer problems. The first problem is to estimate the time required for the interior of the leg to reach a certain temperature, given that the surface temperature has suddenly changed. The second problem is to estimate the surface temperature of the leg, given that the ambient air temperature has suddenly changed.

The alidade leg cross section and some of its properties are shown in Figure 2. Since the ratio of surface area to volume is large, it seems reasonable to treat it as a wall of finite thickness, initially at a uniform temperature θ_0 , and subsequently subjected to a sudden change in the temperature of both surfaces to θ_s . Reference 4 gives the following solution:

$$\theta = \theta_s - (\theta_s - \theta_0) f_3 \left(\frac{\alpha t}{d^2}, \frac{X}{d} \right) \quad (3)$$

where

θ is the temperature at distance X from the center of the wall

d is the wall thickness

α is the thermal diffusivity

t is the time

$f_3 \left(\frac{\alpha t}{d^2}, \frac{X}{d} \right)$ is a dimensionless function of the indicated dimensionless parameters and is given in graph form in Reference 4.

It may be seen that the temperature at the center of the leg flange (center of wall) is almost the same as the surface temperature after only a few minutes by evaluating equation (3) for $t = 120$ sec. Using the graph of Reference 4, the following value of f_3 is obtained:

$$\begin{aligned} f_3 \left(\frac{\alpha t}{d^2}, \frac{X}{d} \right) &= f_3 \left(\frac{0.000012(120)}{0.072^2}, \frac{0}{0.072} \right) \\ &= f_3(0.277, 0) = 0.10 \end{aligned}$$

The substitution of this value into equation (3) yields:

$$\theta = \theta_s - (\theta_s - \theta_0)0.10 = 0.90 \theta_s + 0.10 \theta_0 \quad (4)$$

which shows that the innermost temperature is more than 90% of the surface temperature after only two minutes from the sudden change of surface temperature. For the case at hand, the surface temperature will change sufficiently slowly to justify the assumption that the effective leg temperature is the same as its surface temperature.

Reference 4 also contains the solution for the problem of a wall of finite thickness initially at a uniform temperature θ_0 , subsequently exposed on both sides to a liquid or gas at constant temperature θ_f . The surface temperature, θ_s , is given by the following equation:

$$\theta_s = \theta_f + (\theta_0 - \theta_f) f_6 \left(\frac{\alpha t}{d^2}, \frac{hd}{k} \right) \quad (5)$$

where f_6 is a dimensionless function of the dimensionless Fourier and Biot numbers as indicated, where h is the surface conductance, k is the coefficient of conductivity of the alidade leg, α , t , and d are as defined before.

The surface conductance, h , is often times evaluated as follows:

$$h = 5.68 [1 + 0.74 V] \quad (6)$$

where V is the wind speed in meters/second and h is in watts/m² degree C. Using a wind speed of 4.47 m/sec (10 mph), the value of h is 24.5 watts/m² degree C.

Taking a time value of 4320 seconds and using the properties of the alidade leg shown in Figure 2, the Fourier and Biot numbers are respectively 10.0 and 0.039, for which values the function f_6 is 0.48 according to the curves of Reference 4. For these values equation (5) becomes:

$$\theta_s = \theta_f + (\theta_0 - \theta_f)(0.48) = 0.52 \theta_f + 0.48 \theta_0 \quad (7)$$

Reference 5 gives a convenient formula for the case of a body of any shape having a surface area A_s and a volume V , namely:

$$\theta_s = \theta_f + (\theta_0 - \theta_f) \exp\left(-\frac{h A_s t}{CV\rho}\right) \quad (8)$$

where C is the specific heat of the body and ρ is its mass density.

When equation (8) is evaluated for $t = 4320$ seconds, $h = 24.5$ watts/m² degree C, and the properties shown in Figure 2, the following is obtained:

$$\theta_s = \theta_f + (\theta_0 - \theta_f)(0.44) = 0.56 \theta_f + 0.44 \theta_0 \quad (9)$$

which checks reasonably well with equation (7) derived from the other reference. Using equation (8) and the above values of the parameters, there is obtained:

$$\theta_s = \theta_f + (\theta_0 - \theta_f) \exp(-0.000189 t) \quad (10)$$

which gives the surface temperature, θ_s , of the leg as a function of time. Here θ_0 is the initial uniform temperature of the leg and θ_f is the new ambient air temperature moving at a speed of 4.47 meters per second which is the average wind speed at DSS 14.

Equation (10) shows that there can be a considerable lag between the leg surface temperature and the ambient air temperature. For instance, two hours after a sudden air temperature change, equation (10) gives:

$$\theta_{s_{t=7200}} - \theta_f = (\theta_0 - \theta_f)(0.256) \quad (11)$$

$$\theta_0 - \theta_{s_{t=7200}} = \theta_0 - [\theta_f + (\theta_0 - \theta_f)(0.256)] \quad (12)$$

$$\theta_0 - \theta_{s_{t=7200}} = 0.744 (\theta_0 - \theta_f) \quad (13)$$

Thus after two hours, the difference between original and surface temperature is 74.4% of the difference between original and ambient air temperature. This tends to support the last column values of Table 1.

IV. Comparison of Measured and Calculated Vertical Displacements

From the measured leg temperatures of May and June of 1976, for cases where the ambient temperature daily excursions were approximately 15°C, it appears from Table 1 that the average leg surface temperature excursion averages 75% of the ambient excursion. The previous discussion shows that the effective leg temperature approximates very closely the surface temperature. From the December 1966 tower temperature records, it was concluded that the tower temperature excursion was 20% of the ambient temperature excursion.

During the reference point displacement test from June 2 to June 4, 1977, the ambient temperature excursion was 19°C. Using 75% of the ambient, the vertical displacement, ΔZ , is calculated as follows:

$$\Delta Z = 0.75 \Delta T \beta \ell \quad (14)$$

where ΔT is the ambient excursion, β is the coefficient of linear expansion, ℓ is alidade leg length.

Using the values from Figure 2, equation (14) becomes:

$$\Delta Z = 0.75 (19^\circ) 10.8 \times 10^{-6} (25000) = 3.85 \text{ mm} \quad (15)$$

which is the calculated vertical displacement with respect to ground.

The calculated vertical displacement of the instrument tower top with respect to ground is:

$$\Delta Z_{\text{TOWER}} = 0.20 \Delta T \beta L \quad (16)$$

where L is the length of the steel tower and is also 25000 mm.

$$\begin{aligned} \Delta Z_{\text{TOWER}} &= 0.20 (19^\circ) 10.8 \times 10^{-6} (25000) \\ &= 1.03 \text{ mm} \end{aligned} \quad (17)$$

The calculated relative vertical displacement between reference point and tower top is ΔZ_{REL} :

$$\Delta Z_{\text{REL}} = \Delta Z - \Delta Z_{\text{TOWER}} = 3.85 - 1.03 = 2.82 \text{ mm} \quad (18)$$

which is to be compared to the measured relative vertical displacement of 2.50 mm.

V. Estimation of Total Displacement of Reference Point

The analyses and tests described above allow an estimation of the magnitude of the total reference point displacement

with respect to ground to be made, provided the extreme daily mean ambient temperatures are known.

The following assumptions are made:

- (1) The maximum daily mean ambient temperature is 35°C and the temperature ranges from 27° to 43°C.
- (2) The minimum daily mean ambient temperature is 2°C and the temperature ranges from -6° to 10°C.
- (3) The reference point is defined as the intersection point of the antenna axes at an ambient temperature of 18.5°C and for a condition of zero wind.
- (4) The alidade leg temperature is assumed to be equal to the daily ambient mean plus (or minus) 75% of the instantaneous ambient variation from the daily ambient mean.
- (5) The maximum leg temperature is 35° + 0.75 (8) or 41°C which is 22.5°C above the reference temperature.
- (6) The minimum leg temperature is 2° - 0.75 (8) or - 4°C which is 22.5°C below the reference temperature.

The maximum vertical displacement, ΔZ from the zero reference condition is:

$$\Delta Z = \Delta T \beta \ell = 22.5 (10.8 \times 10^{-6}) 25000 = 6.1 \text{ mm} \quad (19)$$

The maximum horizontal displacement caused by temperature change, ΔH_T , is estimated by forming the ratio between the expected ΔT and the modified ΔT observed during the reference point displacement experiment. From the experiment, the resultant horizontal displacement was $\sqrt{1.75^2 + 1.0^2}$ or 2.02 mm over an ambient temperature range of 19°C. This, by assumption number 4 above, corresponds to a leg temperature range of (0.75) (19) or 14.25°C. Therefore, the expected maximum horizontal displacement of the reference point caused by thermal effects, ΔH_T , is:

$$\Delta H_T = 2.02 \left(\frac{22.5}{14.25} \right) = 3.19 \text{ mm} \quad (20)$$

The horizontal displacement caused by azimuth bearing runout, ΔH_B , is, from the experiment:

$$\Delta H_B = \pm 0.50 \text{ mm} \quad (21)$$

It is believed that the runout of the elevation bearing will not exceed this value and will be ignored.

The calculated horizontal displacement of the reference point caused by a 13.4 m/sec (30 mph) wind, when the dish is faced into the wind at the horizon look, is:

$$\Delta H_W = 1.6 \text{ mm} \quad (22)$$

Assuming that these three horizontal components of displacement are arithmetically additive, the resultant horizontal component, ΔH_R , is:

$$\begin{aligned} \Delta H_R &= \Delta H_T + \Delta H_B + \Delta H_W \\ &= 3.19 + 0.50 + 1.6 = 5.30 \text{ mm} \end{aligned} \quad (23)$$

The total displacement, D , is obtained by combining the vertical and horizontal components geometrically, obtaining:

$$D = \sqrt{(\Delta Z)^2 + (\Delta H_R)^2} = \sqrt{6.1^2 + 5.3^2} = 8.1 \text{ mm} \quad (24)$$

If the effects of the 13.4 meter per second wind and the bearing runout are ignored, the displacement, D_0 , is:

$$D_0 = \sqrt{(\Delta Z)^2 + (\Delta H_T)^2} = \sqrt{6.1^2 + 3.19^2} = 6.9 \text{ mm} \quad (25)$$

Since D_0 is almost as large as D , and since D_0 is proportional to the difference between the reference temperature of 18.5°C and the mean daily ambient of the day of a VLBI experiment, an approximate expression for the displacement of the reference point, D_A , is:

$$D_A = 6.9 (t - 18.5), \text{ in millimeters} \quad (26)$$

where t is the mean daily ambient temperature on the day of the VLBI experiment.

References

1. Hung, Phillips, and Zanteson, *JPL Deep Space Network Progress Report 42-36*, Page 41.
2. H. McGinness, *JPL Space Programs Summary 37-50*, Vol. II, Page 151.
3. H. McGinness, *JPL Technical Report 32-1526*, Vol. VI, Page 142.
4. Brown, and Marco, *Introduction to Heat Transfer*, McGraw Hill, 1958.
5. F. Kreith, *Principles of Heat Transfer*, Intext Educational Publishers, 1973, Page 141.

ORIGINAL PAGE IS
OF POOR QUALITY

Table 1. Air Temperatures

Day of Year 1976	Temperature Ranges, Celsius degrees				
	①	②	③	②	③
	Ambient Air Temp. Range for Hours 0600 and 1500 approx.	Alidade Leg Surface Temp. Range for Hours 0600 and 1500 ^a	Alidade Leg Surface Temp. Range for Hours 0600, 1500, and 2300 ^a	①	①
138	15	12.7	12.7	.85	1.00
143	13	6.7	11.3	.52	.87
145	11	5.8	5.8	.53	.53
146	18	12.7	16.7	.71	.93
149	12	8.8		.73	
151	17	11.0	14.0	.65	.82
157	15	13.3	14.0	.89	.93
158	15	13.9	13.9	.93	.93
168	11	8.0	8.0	.73	.73
169	15	10.8	10.8	.72	.72
172	13	8.4	8.4	.65	.65
177	13	6.3	13.3	.48	1.00
181	10	6.0	6.0	.60	.60
183	15	7.1	15.1	.47	1.00
185	18	10.3	10.3	.57	.57
186	13	6.9	14.5	.53	1.00
187	17	9.4	11.2	.55	.66
190	14	9.6	9.6	.69	.69
195	12	7.4	7.6	.62	.63
197	13	8.5	8.5	.65	.65
200	12	6.9	7.0	.58	.58
213	14	5.8	6.8	.41	.49
216	13	10.0	13.0	.77	1.00
217	12	5.8	8.3	.48	.69
222	15	8.5	11.3	.57	.75
233	18	10.5	13.7	.58	.76
241	13	6.0	7.4	.46	.57
243	12	4.7	8.2	.39	.68
245	15	8.6	10.5	.57	.70
			MEAN	.617	.754
			STD. DEVIATION	.136	.163

^aBased on average of all gages.

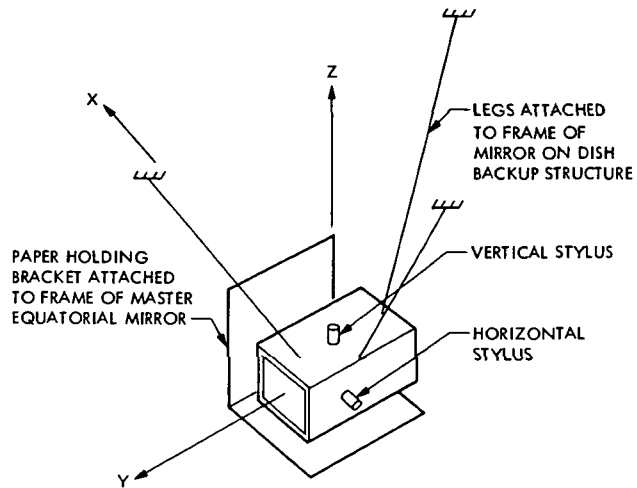


Fig. 1. Schematic diagram of test apparatus

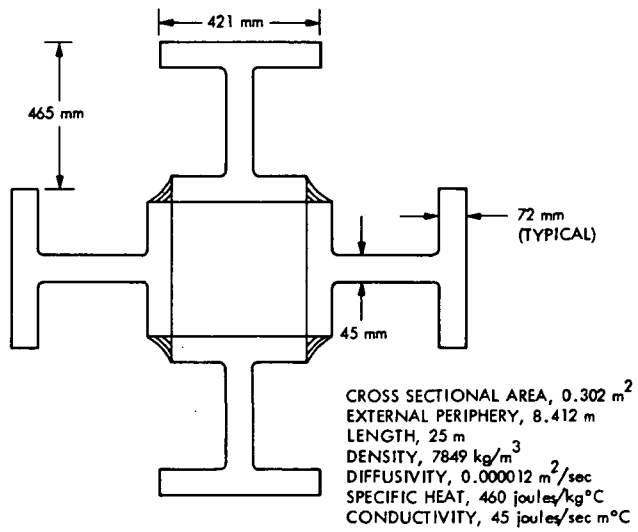


Fig. 2. Alidade leg cross section

N78-11156

Dynamic Modeling and Sensitivity Analysis of Solar Thermal Energy Conversion Systems

C. L. Hamilton
TDA Planning Office

Since the energy input to solar thermal conversion systems is both time variant and probabilistic, it is unlikely that simple steady-state methods for estimating lifetime performance will provide satisfactory results. The work described here uses dynamic modeling to begin identifying what must be known about input radiation and system dynamic characteristics to estimate performance reliably. Daily operation of two conceptual solar energy systems was simulated under varying operating strategies with time-dependent radiation intensity ranging from smooth input of several magnitudes to input of constant total energy whose intensity oscillated with periods from 1/4 hour to 6 hours. Integrated daily system output and efficiency were functions of both level and dynamic characteristics of insolation. Sensitivity of output to changes in total input was greater than one. These findings support the feeling that interplay of radiation dynamics and collector response times affects the quality of energy delivered, and therefore system performance.

I. Introduction

Considerable effort is currently devoted to development and demonstration of concepts for utilizing solar energy, anticipating an eventual significant contribution from that source to meeting society's energy needs. Widespread implementation of solar conversion systems, particularly those for production of solar thermal power, will not occur, however, until they can be designed to meet cost and performance specifications with reasonable assurance. There are intrinsic differences between solar-driven power systems and those the engineering community has traditionally dealt with. Conventional systems can be designed for predominantly steady-state operation at the appropriate capacity, then provided with the amount of fuel necessary to produce the output expected. Design of a solar

power system, in contrast, must be done not only on the basis of its intended output but also considering the energy input that can be expected during its lifetime. That energy input is location-dependent and time-variant; fluctuation of solar radiation intensity with time is only partly deterministic, with largely probabilistic effects occurring on time scales from minutes to seasons or longer.

One consequence of these characteristics is that solar-driven power systems must spend a significant portion of their operating time responding to input transients. Another consequence is that a long lead time is required to evaluate the solar resource at any one location. This paper addresses the thesis that performance analysis for solar thermal power systems must take into account the dynamic characteristics of both the

components and the input energy to lead to satisfactory results. There is a need to identify in a fundamental sense what aspects of system dynamic response are important to performance under what conditions. Among other things, this understanding would allow timely initiation of insolation measurement programs that would yield without excess expense information adequate to support reliable design of commercial solar power systems.

The work described here represents an initial sensitivity analysis whose goal was to determine how a solar thermal conversion system's output responds to variations in the energy input. One objective was to learn how much performance estimates might be affected by uncertainty in the magnitude of measured insolation values. Another was to gain some understanding of what effect variations in solar radiation intensity on several time scales (cloudy skies) might have on system performance, and how any such effects might depend on those system parameters that govern its dynamic response characteristics.

II. Description of the Study

Since fundamental understanding was sought here, rather than detailed analysis of any specific systems, a generalized approach to the problem was taken. Emphasis is on dynamic behavior, and we wish to examine interactions one step at a time. System behavior is considered here in terms of energy flow, starting with solar radiation intensity as a time-dependent input. The model system is envisioned as the simplest group of components that will take that input as heat and convert it to mechanical energy, the output. A computer-based dynamic model serves to relate mechanical energy production to radiation input as a function of time and the system's dynamic characteristics. Calculation of the energy outputs resulting from a representative set of driving functions yields data from which judgment can be made regarding the sensitivities we are interested in.

A. Dynamic Modeling Technique and Model System

The dynamic model used in this work represents a simplified application of a versatile methodology for time-dependent simulation of the energy transfer behavior of systems and their components (Ref. 1). The technique has been applied at various levels of detail, from rough examination of system behavior in response to general specifications on subsystem performances (Ref. 2) to detailed analysis of effects exerted on a component's transient and integrated performance by its thermal characteristics (Ref. 3). Functionally, the model performs an explicit forward finite difference analysis on energy flow in a system, which may be defined as made up of a collection of components or subsystems, or of segments of one

component. It is expressed in a framework that retains an intuitively understandable connection between parameters and their effects and is embodied in an interactive computer program of modular design for flexibility in adding or substituting a variety of components.

For purposes of this study, the system illustrated in the energy flow diagram, Fig. 1, was characterized for simulation. Representation of the solar collector is quite general; it is treated as a lumped energy transfer device. The parameters used to model the collector include: thermal capacitance (the aggregate heat capacity of those portions of the component in which energy will accumulate, like absorber structure and heat transfer fluid); a lumped heat loss coefficient (which need not be constant, but can be expressed as a function of collector temperature or internal energy); and characteristics of the heat transfer fluid (specific heat, flow rate, temperature at collector inlet—these also need not be constant). These descriptive parameters are all expressed as values per unit area of collector aperture, allowing treatment in the model of collectors with widely varying shapes, sizes and concentration ratios. An ideal engine operating at a fixed fraction of Carnot efficiency is assumed for conversion of the collected thermal energy to mechanical energy. Pertinent characteristics of the surroundings are ambient temperature and a sink temperature to which the Carnot engine rejects heat.

A simulation begins with an initial value for the internal energy content of the collector and proceeds according to the equation

$$CHS_{(T+DT)} = CHS_{(T)} + (CHIR_{(T)} - CHOR_{(T)})DT$$

in time steps that are small in comparison with both system response times and rate of change of radiation intensity.

CHS designates the heat stored in the collector at the time subscripted, $CHIR$ is the net rate at which energy enters the collector, $CHOR$ is the net rate for energy leaving the collector and DT is the time increment over which the rates apply. Both rates are determined for each time step from solar input rate, the collector parameters, and parameters describing the surroundings (in this case ambient temperature) pertaining at time T . At each time step, the temperature of fluid leaving the collector (on which engine efficiency will depend) can be determined from the corresponding value of CHS . Here it was assumed for simplicity that the temperature distribution between collector inlet and outlet is linear and the temperature defined in terms of CHS corresponds to the average between the two. Finer spatial resolution could be used, but would not affect qualitatively the findings produced.

As written, the computer code contains one set of characteristic system parameters and initial values. To run the program, inputs are provided in the form of an equation describing solar radiation intensity as a function of time plus any desired changes in the descriptive parameters. Simulation corresponding to one working day (12 hours from sunup to sunset) is carried out in each run; program outputs include solar radiation input integrated over the day, converter output integrated over that period, and integrated system efficiency (total output/total input) for the day.

B. Protocol for Parameter Variations

Our experimental system was first characterized in a baseline configuration, whose specifications are summarized in Table 1. The first five entries represent quantities that remained constant throughout this study. Those in the lower group comprise the collector-dependent parameters and operational characteristics that it was anticipated would affect system dynamic response. Each of these factors was varied in the course of the investigation.

While the system we are looking at is an admitted abstraction, the parameters used to describe it are not chosen entirely arbitrarily. Specifically, the values of collector thermal capacitance and heat loss coefficient used are consistent with the characteristics of existing collectors. The parameters incorporated into the baseline configuration correspond to a glass concentric tube collector produced by Owens-Illinois; they are derived from a separate detailed dynamic analysis of that collector. The concentric tube design is characterized by large thermal inertia and slow response. For an alternate set of collector parameters, thermal capacitance and loss coefficient values representing a fast responding collector, a NASA-Honeywell design, were also tested. Those were experimental values (Ref. 4). Both choices are examples of moderate performance, non-concentrating collectors.

Six variations on the baseline configuration were examined. For each of the two collector types (slow responding and fast responding) system performance was simulated under three different sets of hypothetical operating constraints. Two of those involved operation with collector heat transfer fluid flowing at a constant rate (chosen for each collector such that a steady solar input of 1 kW/m^2 would produce an equilibrium outlet temperature of 240°C). In the first constant flow rate case, the engine was programmed to accept all the energy delivered to it above 62°C , the inlet temperature, for conversion. A somewhat more realistic second case limited the engine to utilizing only energy above 100°C . The third set of operating constraints involved variation of the fluid flow rate to maintain the collector outlet temperature within a fairly narrow range ($175\text{-}200^\circ\text{C}$).

Daily performance for each modification of the baseline configuration was simulated using a systematically varied set of solar radiation inputs. Those included a basic "clear-day" function (the first loop of a sine function plus some third harmonic, Fig. 2) with amplitudes of 1, 0.8, 0.6, 0.4, and 0.3 or 0.2, intended to mimic days ranging from clear to uniformly hazy. To examine the effect of oscillations in radiation intensity, sine functions of amplitude 0.4 were imposed as a harmonic component on the basic function with amplitude 0.6; the periods of those harmonics varied from six hours down to 15 minutes. Figure 3 illustrates a sample of harmonic input.

III. Results

Figures 4 and 5 summarize the effects exerted on performance of our conceptual systems by their dynamic characteristics and operating strategies, in response to solar input containing only smooth diurnal variation. Integrated daily system efficiencies are plotted against integrated daily insolation; it can be seen that in all cases tested integrated efficiency is a non-linear function of total energy collected. This is a consequence of the fact that a heat engine's performance depends on both the quantity and quality (temperature) of the energy delivered to it, and reflects the impact of radiation intensity on the temperature increase that can be sustained in the collector heat transfer fluid.

Data from these runs were used to determine the sensitivity of system output estimates to uncertainty in the knowledge of total input. Plots of total output versus total input were constructed, and the range of system outputs corresponding to a variation of $\pm 10\%$ and $\pm 15\%$ around a nominal daily radiation input of 7 kWh/m^2 was extracted. The resulting sensitivities, expressed in terms of percentage output change divided by percentage change in input, are tabulated in Table 2. Owing to the non-linearity of the input-output curves the sensitivity values vary some, but all are 1.6 or greater. They indicate that a 10% inaccuracy in a measured value of daily insolation would lead to an estimate of system performance for that day that would be in error by at least 16%, even if the performance estimate were based on dynamic analysis.

The results discussed above apply to operation of two imaginary non-concentrating solar thermal energy conversion systems on days ranging from uniformly hazy to uniformly clear. Figures 6 and 7 show the effect on system performance of some short-term variations in radiation intensity, mimicking cloud passage in otherwise clear skies. In these graphs all points represent response to the same total radiation input, 5.38 kWh/m^2 . The variables represented are operating strategy and the period of oscillation for the superimposed harmonic in

the solar radiation function. Deviations in system efficiency from that produced by smooth input were plotted against the harmonic period. Here the influence of collector response time shows strongly. Performance of the system containing a fast-responding collector is more sensitive to the frequency of input transients, but both collectors produce the same general behavior. System efficiency is higher than expected from the total daily input when radiation intensity oscillates slowly compared to the collector's response time—the system can develop higher working temperatures when it has time to respond to the peaks in intensity. Conversely, intensity oscillations that are too rapid to follow tend to degrade system performance, since the collector cannot take advantage of the peaks to make up for energy missed during the troughs. In the case of the fast collector, system efficiency varies nearly $\pm 15\%$ from that seen with smooth input as the dynamics of input radiation change.

IV. Conclusions

In all the cases examined in this study, both the integrated daily system output and system efficiency as estimated by dynamic simulation were functions of both the level and dynamic characteristics of the insolation that was assumed to drive the system. System output showed a sensitivity of signifi-

cantly greater than one to variations in the magnitude of total radiation assumed, even without considering the additional effects produced by input transients. These findings support the intuitive feeling that the interplay of input radiation dynamics and collector response times can have a pronounced effect on the quality of energy delivered, and therefore on system performance. They also suggest that evaluation of system performance for design purposes must be based on accurate radiation figures, either measured data or a modeled equivalent, spaced at short time intervals. Most of the insolation data currently available or reconstructable is likely to be neither accurate enough nor detailed enough to support design of a commercial solar power plant (Ref. 5).

The effects observed here apply to an abstracted, highly simplified system concept. More work is needed to probe, systematically, other factors that have potential for affecting the dynamics of system performance. Realistic engine performance characteristics should be included, as well as the effects of storage subsystems. The ultimate goal is an understanding of which design parameters tend to exaggerate and which tend to attenuate the basic input-output sensitivities demonstrated here. With that understanding should come ability to design solar power systems for efficient performance at their designated sites, and to specify the requirements on insolation data necessary to support site selection and design.

References

1. Hamilton, C. L., "A Dynamic Model for Analysis of Solar Energy Systems," *The Deep Space Network Progress Report 42-27*, p. 41, 1975.
2. Hamilton, C. L., "An Experiment in Dynamic Modeling for a Complete Solar-Powered Energy System," *The Deep Space Network Progress Report 42-31*, p. 137, 1975.
3. Hamilton, C. L., "Dynamic Modeling for Evaluation of Solar Collector Performance," *The Deep Space Network Progress Report 42-34*, p. 141, 1976.
4. Simon, F. F., *Flat-Plate Solar Collector Performance Evaluation With a Solar Simulator as a Basis for Collector Selection and Performance Prediction*, NASA Technical Memorandum NASA TM X-71793, 1975.
5. Durrenberger, R. W., and Brazel, A. J., "Need for a Better Solar Radiation Data Base," *Science* Vol. 193, p. 1154, 1976.

**ORIGINAL PAGE IS
OF POOR QUALITY**

Table 1. Specifications for system baseline configuration

Parameters held constant	
Heat engine mechanical/cycle efficiency	0.8
Sink temperature	35°C
Ambient temperature	25°C
Heat transfer fluid specific heat	$9.99 \times 10^{-4} \frac{\text{kWh}}{\text{kg}^\circ\text{C}}$
Collector inlet temperature	62°C
Variable parameters	
Collector thermal capacitance ^a	$0.016 \frac{\text{kWh}}{\text{m}^2^\circ\text{C}}$
Collector heat loss coefficient ^a	Variable ^b
Heat transfer fluid flow rate	$4.5 \frac{\text{kg}}{\text{hm}^2}$
Temperature of energy accepted by converter	>62°C

^aFor alternate collector, thermal capacitance is 0.0026 kWh/m²°C, loss coefficient is 26.8×10^{-4} kW/m²°C.

^bVaries with collector temperature from 8.7×10^{-4} kW/m²°C at 62°C or less to 15.9×10^{-4} kW/m²°C at 102°C or greater.

Table 2. Sensitivity measures for system response to smooth radiation input of nominal 7 kWh/m²

$\frac{\Delta I}{I}$	$(S = \frac{\Delta O/O}{\Delta I/I})$	Fast collector		Slow collector	
		Constant flow	Constant temperature	Constant flow	Constant temperature
-0.10		1.65	1.9	1.62	1.7
+0.10		1.74	1.62	1.69	1.61
-0.15		1.71	1.78	1.69	1.64
+0.15		1.71	1.65	1.64	1.64

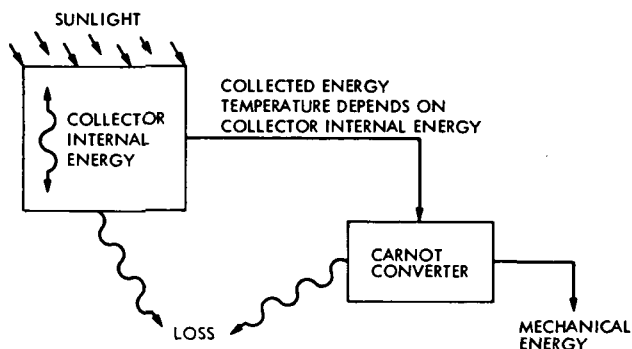


Fig. 1. Abstracted system for sensitivity analysis, diagrammed in terms of energy flow

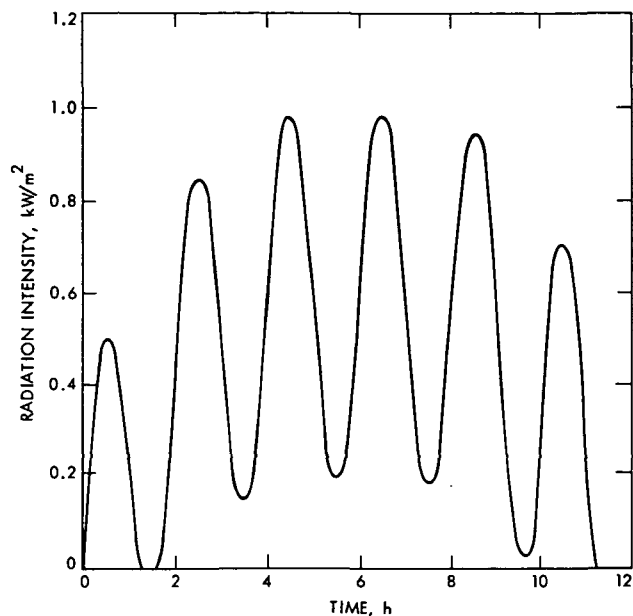


Fig. 3. Sample of input function with superimposed harmonic of 2-hour period

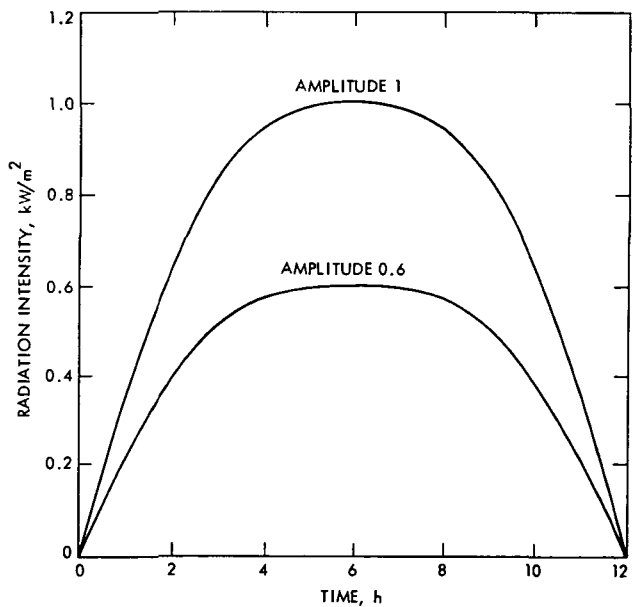


Fig. 2. Basic smooth input function for solar radiation intensity

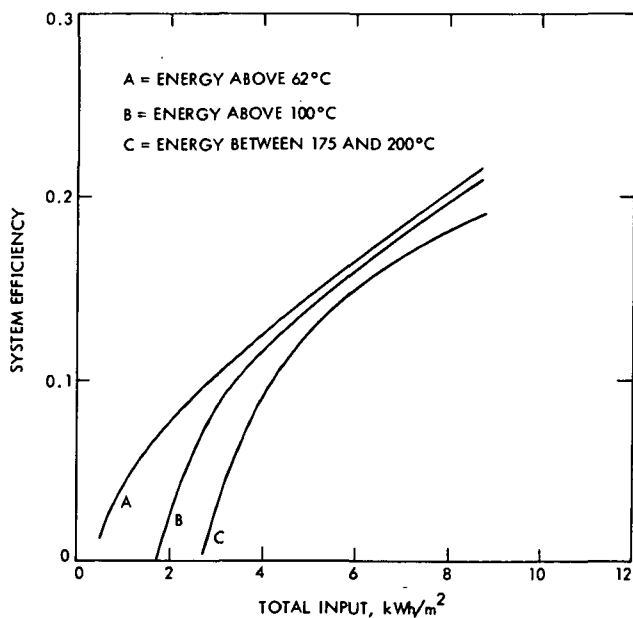


Fig. 4. Integrated system efficiency versus smooth daily radiation input for system containing slow responding collector

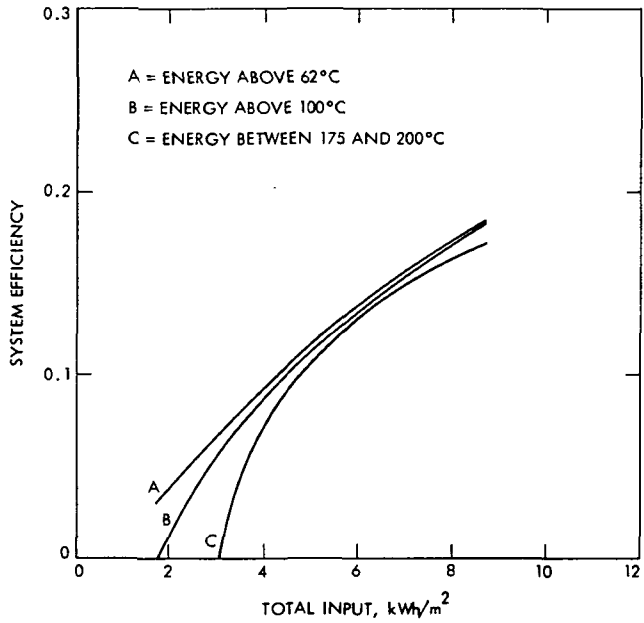


Fig. 5. Integrated system efficiency versus daily radiation input for system containing fast responding collector, smoothly varying input curve

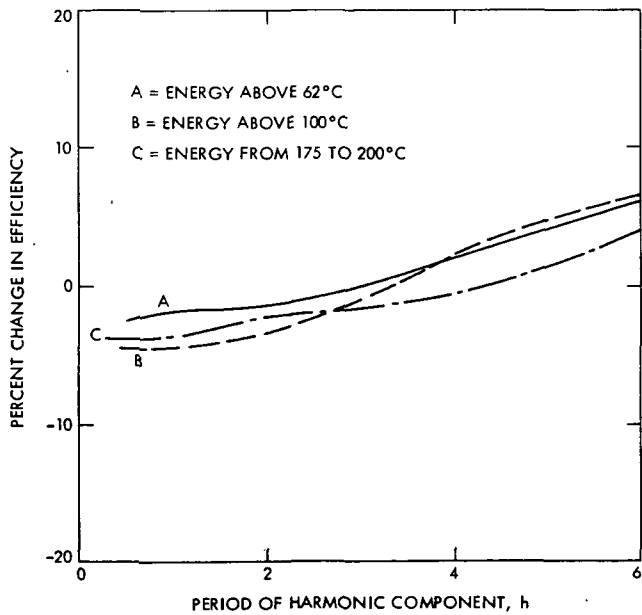


Fig. 6. Effect of input transients on efficiency of system with slow collector

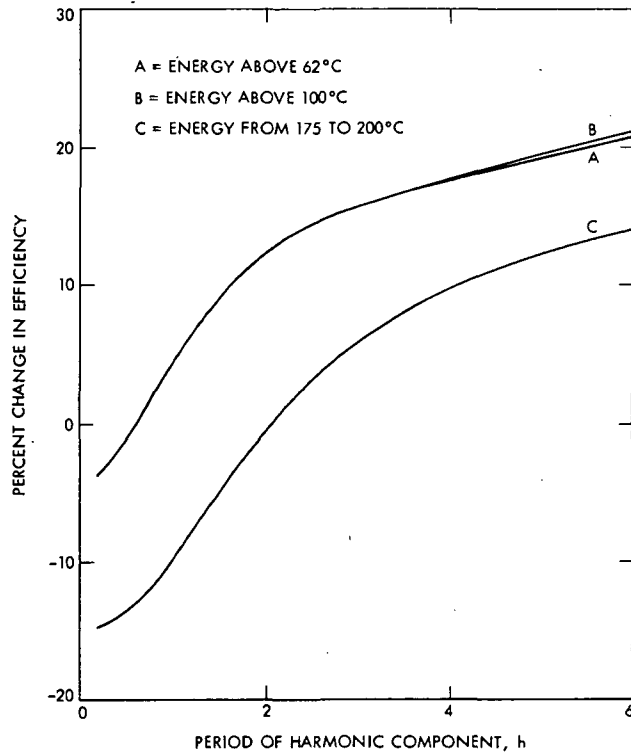


Fig. 7. Effect of input transients on efficiency of system with fast collector