Low-Noise Systems in the Deep Space Network

Edited by Macgregor S. Reid

Jet Propulsion Laboratory
California Institute of Technology

DEEP SPACE COMMUNICATIONS AND NAVIGATION SERIES

DEEP SPACE COMMUNICATIONS AND NAVIGATION SERIES

Issued by the Deep Space Communications and Navigation Systems
Center of Excellence
Jet Propulsion Laboratory
California Institute of Technology

Joseph H. Yuen, Editor-in-Chief

Published Titles in this Series

Radiometric Tracking Techniques for Deep-Space Navigation
Catherine L. Thornton and James S. Border

Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation
Theodore D. Moyer

Bandwidth-Efficient Digital Modulation with Application to Deep-Space Communications
Marvin K. Simon

Large Antennas of the Deep Space Network
William A. Imbriale

Antenna Arraying Techniques in the Deep Space Network
David H. Rogstad, Alexander Mileant, and Timothy T. Pham

Radio Occultations Using Earth Satellites: A Wave Theory Treatment
William G. Melbourne

Deep Space Optical Communications
Hamid Hemmati, Editor

Spaceborne Antennas for Planetary Exploration
William A. Imbriale, Editor

Autonomous Software-Defined Radio Receivers for Deep Space Applications
Jon Hamkins and Marvin K. Simon, Editors

Low-Noise Systems in the Deep Space Network
Macgregor S. Reid, Editor

Low-Noise Systems in the Deep Space Network

Edited by Macgregor S. Reid

Jet Propulsion Laboratory
California Institute of Technology

DEEP SPACE COMMUNICATIONS AND NAVIGATION SERIES

Low-Noise Systems in the Deep Space Network

February 2008

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

Table of Contents

Foreword

The Deep Space Communications and Navigation Systems Center of Excellence (DESCANSO) was established in 1998 by the National Aeronautics and Space Administration (NASA) at the California Institute of Technology's Jet Propulsion Laboratory (JPL). DESCANSO is chartered to harness and promote excellence and innovation to meet the communications and navigation needs of future deep-space exploration.

DESCANSO's vision is to achieve continuous communications and precise navigation—any time, anywhere. In support of that vision, DESCANSO aims to seek out and advocate new concepts, systems, and technologies; foster key technical talents; and sponsor seminars, workshops, and symposia to facilitate interaction and idea exchange.

The Deep Space Communications and Navigation Series, authored by scientists and engineers with many years of experience in their respective fields, lays a foundation for innovation by communicating state-of-the-art knowledge in key technologies. The series also captures fundamental principles and practices developed during decades of deep-space exploration at JPL. In addition, it celebrates successes and imparts lessons learned. Finally, the series will serve to guide a new generation of scientists and engineers.

Joseph H. Yuen
DESCANSO Leader

Preface

The Jet Propulsion Laboratory (JPL) is an operating division of the California Institute of Technology (Caltech) in Pasadena, California. JPL's history started in the early 1930s when a graduate student in Caltech's Guggenheim Aeronautical Laboratory, Frank Malina, made a proposal to his professor, Theodore von Karman, for a topic for his thesis. The proposal was to design, build, and test a rocket motor. His proposal was accepted, and Malina started work. He quickly picked up two assistants, and together they made good progress, so good in fact that they were forbidden from any further experiments (and the associated noise) at Caltech. They moved their equipment to a dry arroyo, just outside the Pasadena city limits, about 12 km from the Caltech campus, in what is now the city of La Cañada. This is the present site of JPL.

Successful experiments continued throughout the 1930s, and by the start of the second World War, von Karman had set up the work as a new laboratory that later was called the Jet Propulsion Laboratory. JPL's experiments were successful and picked up U.S. Army funding, which continued throughout World War II and beyond. At this time ballistic missiles were under development, and JPL was also involved with radio tracking and control of the missiles. This tracking system was the forerunner of the present worldwide Deep Space Network (DSN) for tracking spacecraft.

In 1957, the Soviet Union startled the world by launching Sputnik, the first artificial satellite. The following year, JPL became part of the newly formed National Aeronautics and Space Administration. Since then, JPL and its DSN have been leading or supporting exploration of every planet in the Solar System and many other astronomical bodies.

One of today's objectives of JPL is to further our understanding of the origin and evolution of the Solar System, and the origin and evolution of life in the Universe. This is undertaken by robotic spacecraft missions to the planets,

their moons, the asteroids, and the comets. In addition, JPL supports spaceborne observatories probing the limits of the Universe. Furthermore, the probes for this exploration are yielding data at steadily increasing data rates. The DSN's objectives, therefore, are to acquire telemetry data from spacecraft, transmit commands to spacecraft, track spacecraft position and velocity, perform very-long-baseline interferometry observations, conduct radio and radar astronomy, measure variations in radio waves for radio science experiments, gather science data, and monitor and control the performance of the network.

The DESCANSO books-series editor, Joseph H. Yuen, provides a definition and description of the technology of the DSN. This book is one in this series and describes the low-noise microwave systems that form the front end of all the DSN ground stations. The microwave front end is key to establishing the sensitivity and capability of the receiving chain, and therefore, of the entire ground station. The receiving system sensitivity and capability are defined by G/T where G is the antenna gain and T is the overall noise temperature of the entire receiving chain, usually called system operating noise temperature, T_{op}. To improve the station's receiving capability, it is necessary to improve G/T. This can be done by increasing antenna gain or by decreasing T_{op}. In the past the DSN has both increased G and reduced T_{op}. It has been somewhat more cost effective and efficient to decrease T_{op} than it has been to increase G.

As the microwave front end of the station is crucial to establishing the sensitivity and capability of the receiving chain, it is therefore incumbent on the designers of the receiving system to expend considerable effort in reducing T_{op} and in calibrating and maintaining the low-noise front end. The more accurately T_{op} can be defined, the narrower the tolerances on the design of a spacecraft mission that can be accepted. This reduction in a project's design tolerances on spacecraft power can yield a considerable reduction in cost, it can enable an increase in the science data rate for the same spacecraft power, or it can provide some compromise between the two. Hence, improving calibration accuracy of the receiving chain is vitally important. Accurate noise-temperature calibrations are also necessary for the maintenance of the low-noise performance of the station. In a similar manner, the accuracy in the measurement of antenna gain is also important. This book also describes current antenna calibrations, some of which are new.

The challenge of successful radio communication at planetary distances is unique and challenging. An extremely low-noise front end is a requirement for the sensitivity that must be, and is, routinely achieved within the DSN. This book is a description of the various low-noise systems in the DSN; their development, calibration, and operation; and how these systems have been used

for tracking and for science. Overall system noise-temperature calibration of the front end is considered in one chapter, and important types of low-noise receiver front end are treated in detail in other chapters, as is the atmosphere and the calibration of antenna gain.

The book will be useful to designers and operators of communications systems, radio- and radar-astronomy stations, space-research facilities, and interferometry observatories. Analyses are presented in detail with rigor, and key equations are summarized for easy reference. The use of precise definitions (more precise than usually used in industry) in system noise temperature calculations is unusual. These definitions are based on IEEE standards and definitions not always used by practitioners. No other book goes into this level of detail. However, all analyses and equations are explained in full with examples from field measurements. The detailed explanations mean that limited engineering knowledge is sufficient to understand the text, but the rigor is there if required. High school algebra and undergraduate level calculus are generally sufficient to understand this book, although some parts may require graduate level calculus. However, the analyses are given with examples and working programs in detail so that no calculations or mathematics are required for the statistics and averages of calibration errors of measurements.

Macgregor S. Reid,
Pasadena, California
February 2008

Acknowledgments

This book presents results of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

The editor and co-authors acknowledge the constant advice, help, comments, and encouragement from Joseph H. Yuen, the general editor of the DESCANSO series. In addition, they acknowledge help from Patricia A. Ehlers and Roger V. Carlson, the JPL editors, and from Judi Dedmon, the JPL layout specialist. Dan Bathker provided much advice on DSN antennas in general and the 70-m antennas in particular. Tom Otoshi contributed many detailed analyses of the DSN microwave front-end components and systems.

The authors of Chapter 2 give thanks to Robert C. Clauss for help in early draft versions, and for co-authoring numerous DSN Progress Reports used as a foundation for this chapter. The questions and comments from David L. Hills were most helpful, and Douglas S. Hofine and Gary Bury at the Goldstone Communications Complex supplied many data sets. Manuel M. Franco and Lyle J. Skjerve provided many data sets and collections from Goldstone antennas. Samuel M. Petty supplied much of the information for Section 2.4 on Low Noise Amplifiers, and Stephen D. Slobin verified most of the algorithms and calculations in Section 2.5 Receiving Systems.

The authors of Chapters 3 and 4 wish to particularly thank Dr. Walter Higa, who led JPL's maser development work for more than twenty years, beginning in 1955. He assembled and coached a team responsible for all aspects of maser research, development, and operations in the field environment of large DSN antennas. The many technologies involved included quantum physics, microwave engineering, and thermodynamics with emphasis on cryogenic refrigerator design. We thank the many scientists, engineers, technicians, and specialists at JPL, in the DSN, at universities, at Airborne Instruments

Laboratories, and at Bell Telephone Laboratories for their countless contributions to the development and successful operation of these masers.

Chapter 5 author, Dr. Bautista, would like to thank Sander Weinreb and Sam Petty who originally conceived and proposed development of the cryogenic HEMT for the Voyager at Neptune Encounter. He would also like to express his gratitude to Bert Fujiwara and Steve Montanez who performed the micro-circuit assembly and Jose Fernandez who led infusion of the high electron mobility transistor technology into the Deep Space Network.

Chapter 6 author, Dr. Shambayati, would like to thank Dr. Stephen D. Slobin of the Jet Propulsion Laboratory for his invaluable help in explaining JPL's ground weather model to him.

Chapters 7 and 8 author, Dr. Rochblatt, wishes to thank those who contributed for this work: first and foremost Paul Richter for his system analysis and for pioneering the raster scan technique at JPL; Phil Withington and Herschel Jackson for their contribution to the development of the OTF-Mapping R&D system; and Manuel Vasquez and Jesus Calvo from the Madrid Deep Space Tracking Station (MDSCC) in Spain for the development of the operational antenna calibration instrumentation. In addition the author would like to thank Stephen Slobin, Art Freiley, Graham Baines, Charles Stelzried, Manuel Franco, and Pablo Perez for their contributions via technical discussions. Also, thanks go to W. Van Snyder for assistance with the integration of Eq. (7.4-13). Dr. Rochblatt thanks Lawrence Teitelbaum and the JPL Radio Astronomy and Radar Group and the technical staff at the Goldstone Venus Development Site for their dedicated support of the Goldstone-Apple Valley Radio Telescope project and the Jupiter Patrol observations at Goldstone.

Dr. Rochblatt acknowledges the support of the following individuals from JPL who contributed to the success of the holography program: Dan Bathker, Boris Seidel, Phil Withington, Herschel Jackson, Yahya Rahmat-Samii, Carl Jones, Lawrence Ansley, and Peter Kinman. He would also like to acknowledge Eikontech Ltd. and the Interferometrics Corp. for contributing to the success of this program.

February 2008

Contributors

J. Javier Bautista received a BS in physics in 1971 from St. Mary's University (San Antonio, Texas) and an MS and a PhD in low-temperature physics in 1974 and 1979, respectively, from Michigan State University, (East Lansing, Michigan). At the Jet Propulsion Laboratory (JPL), Pasadena California, he is a principal engineer in the Station Communication Ground Systems. In 1986 he founded the JPL Cryo-Electronics Front-end Equipment Group to enhance and improve JPL's cryogenic, low noise amplifier (LNA) and circuit capabilities. The group's primary responsibility was to develop prototype, high electron mobility transistor (HEMT) based LNAs for the Deep Space Network (DSN), and the DSN's Prototype Large Array of Small Aperture Antennas. He is currently responsible for the design and development of HEMT-based cryogenic, LNAs for 1 to 40 GHz for the DSN. This task includes development of characterization techniques for low noise amplifiers, devices, and components.

Robert C. Clauss received an AA from Pasadena City College (California) after attending military electronics schools while in the United States Marine Corps. He retired in 2006, after 47 years at JPL, where he had worked as a technician, engineer, supervisor, and research program manager. His work with microwave systems and equipment included the development of masers at frequencies from 1 to 40 GHz, design work for spacecraft transponders and transmitters, and mentoring undergraduate students who helped design, build, and track SURFSAT, an Earth-orbiting satellite. He has received 8 patents, 18 NASA monetary awards for technology development, and the NASA Exceptional Service Medal.

Arthur J. Freiley received a BS in electrical and electronic engineering from California Polytechnic University (Pomona, California). He joined JPL in 1970 as a member of the Antenna and Propagations Group working on the research and development of large ground-based reflector antennas technology and calibrations for NASA-JPL Deep Space Network (DSN). Mr. Freiley later assumed the duties and responsibilities of cognizant development engineer for the Antenna Microwave Subsystem for the DSN, developing microwave feed systems (complete with computer control systems). He has also served as the task leader for the Goldstone Solar System Radar (GSSR) megawatt transmitter system development. More recently he has served as the antenna microwave system engineer, developing system requirements, and validating and verifying system performance against those requirements. He has also served as the system engineer for the Goldstone-Apple Valley Radio Telescope (GAVRT) Educational Outreach program.

Macgregor S. Reid received his BSc (1953), BSc Eng (1956), MSc Eng (1962), and PhD (1969) (all in engineering) from the University of the Witwatersrand (Johannesburg, South Africa). Dr. Reid worked at JPL from 1969 until his retirement in 1998. For the last 10 years of his career, he was technical executive assistant to the JPL director. He served a three-year term as Vice President of the American Institute of Aeronautics and Astronautics (AIAA), Washington, District of Columbia. Dr. Reid was Formation Chairman of the International Organization for Standardization (ISO) Subcommittee on Space Systems and Operations for 9 years, and he represented the United States on ISO's Committee on Aircraft and Space Operations for 12 years. He has served on many Advisory Committees, both national and international, and he is the author of more than 80 publications in the technical literature, three of which have won awards. In recognition of his international work, an asteroid was named after Dr. Reid in 1998 (6894 MacReid). He was awarded medals from AIAA and NASA, and six monetary awards from NASA for the invention of new technology.

David Rochblatt is recognized internationally as one of the top experts on large-reflector antennas and metrology. Mr. Rochblatt has been a design engineer, the manager for R&D ground antennas systems, and the antenna calibration expert for the NASA-JPL-DSN for more than 25 years. In 1995 he received the NASA Exceptional Achievement Medal for developing the Microwave Antenna Holography System. Mr. Rochblatt consulted at the Russian Space Agency, the U.S. Naval Research Laboratory (Pomonkey, Maryland), the University of Naples (Italy), the University of Massachusetts at Amherst, and Cedars Sinai Medical Centre (Los Angeles, California). Also, he has been an invited speaker at a number of IEEE workshops. Mr. Rochblatt holds a BSEE and an MSEE from the University of California at Los Angeles

(UCLA) (received in 1978 and 1980, respectively) and is a senior member of the IEEE. He has published 70 journal and conference papers and received more than a dozen NASA Achievement Awards. His other interests include deep-space telecommunication, inverse scattering problems, and phase retrieval applications for microwaves and optical communications.

Shervin Shambayati received his BS in applied mathematics and engineering from California State University, Northridge, in 1989 and his MSEE, engineer degree and PhD from the University of California, Los Angeles, in 1991, 1993, and 2002, respectively. Dr. Shambayati joined the JPL Deep Space Telecommunications Group in 1993. With that group, Dr. Shambayati participated in the development and testing of the Deep Space Network Galileo Telemetry Receiver (DGT) and the 34-m Arraying Task. In 1997, Dr. Shambayati transferred to the Information Processing Group at JPL where he has been working ever since. With that group he participated in the Mars Global Surveyor Ka-band Link Experiment II (MGS/KaBLE II) and DS1 in-flight Ka-band tests. Most recently, Dr. Shambayati was the principal investigator for the Mars Reconnaissance Orbiter Ka-Band Demonstration. Dr. Shambayati's main research areas include evaluation of weather effects on the end-to-end performance of the deep-space link; end-to-end link design and spacecraft downlink operations over Ka-band links.

James S. Shell received his BSE in engineering physics and his MS degree in physics, in 1975 and 1976 respectively, both from the University of Michigan, Ann Arbor, Michigan. He received his PhD in physics in 1982 from Michigan State University (East Lansing, Michigan). His research focused on the magnetic properties of dilute magnetic alloys at low temperatures. Prior to coming to JPL in 1985, he worked at Henry Ford Hospital (Detroit) in the area of ultrasound imaging, and at Imperial College (London) researching low temperature properties of metals. Since coming to JPL, he has worked primarily on low noise microwave amplifiers (including ruby masers and high electron mobility transistors) and cryogenics for the DSN. He is currently a member of the Spacecraft Transponder Group at JPL.

Dr. Charles Stelzried received a BS in engineering (1957) and an MS in engineering (1959) from the University of California at Los Angeles, California, and his PhD in engineering (1969) from the University of Southern California (Los Angeles, California). He worked at the Jet Propulsion Laboratory, Pasadena, California, from 1953 to his retirement in 2005. Dr. Stelzried helped pioneer DSN low noise receiving systems starting with the first JPL Venus Radar Experiment in 1961 at Goldstone, California. He was a technical group supervisor in the Telecommunications Engineering Division from 1965 to 1981 as well as a radio science team member with the NASA

Mariner 10, Viking, and Helios missions. From 1981 to 1986 he was the DSN manager for the Ulysses, Magellan, Giotto, Venus Balloon, and Vega Pathfinder deep space missions. Dr. Stelzried is a fellow of the IEEE and a member of Tau Beta Pi, AAAS, Sigma Xi, and the International Union of Radio Science (URSI) Commission A. He has authored more than 170 JPL DSN Progress Reports and 5 patent disclosures (including the multicone configuration used in the DSN 70-m antennas), and 60 refereed publications. He has numerous JPL/NASA awards, including the NASA Medal for Exceptional Scientific Achievement.

Chapter 1
Introduction

Macgregor S. Reid

The Deep Space Network (DSN) is a National Aeronautics and Space Administration (NASA) facility. It is managed, technically directed, and operated by the Jet Propulsion Laboratory (JPL) of the California Institute of Technology (Caltech). The objectives of the DSN are to maintain communications with spacecraft. This consists of collecting telemetry data from the spacecraft, transmitting command data to them, providing spacecraft trajectory data to mission operations and scientists, and monitoring and controlling the network performance. Other objectives are to gather science data from spacecraft, to measure variations in transmitted radio waves from spacecraft for radio science experiments, and to perform very-long-baseline interferometry observations. The DSN is the largest and most sensitive scientific telecommunications and radio navigation network in the world.

The DSN consists of three deep-space communications complexes (DSCC) separated by approximately 120 degrees (deg) of longitude around the world:

1) Goldstone, California
2) Near Madrid, Spain, and
3) Near Canberra, Australia

This placement allows continuous 24-hour observation of spacecraft as the Earth rotates, and it provides sufficient overlap for transferring the spacecraft radio link from one complex to the other.

The precursor of the DSN was established in January 1958 when JPL, under contract to the U.S. Army, deployed radio tracking stations in Nigeria, Singapore, and California to track the first successful U.S. satellite, Explorer 1. Army funding for JPL continued until the National Aeronautics and Space

Administration was officially established in October 1958. Two months later the Army contract with Caltech, which provided funding for JPL, was transferred to a NASA contract, and JPL's primary funding continues through NASA to this day. The Laboratory's employees always have been, and remain to the present, employees of Caltech [1].

At about this time in 1958 the concept of a separate Earth-based communications network was established. Shortly afterward, NASA created the Deep Space Instrumental Facility (DSIF) as a separate Earth-based communications facility that would support many space flight missions simultaneously, thereby avoiding the impractical duplication of a specialized space communications network for each flight project. A separate facility, located on the JPL campus, was the Space Flight Operations Facility (SFOF), which communicated with all the DSIF stations through a Ground Communications Facility (GCF), also a separate organization from the DSIF. The network was given responsibility for its own research, development, and operation, in support of all its users. In 1963 the DSIF was combined with the multi-mission part of the SFOF and the GCF to form the DSN. Under this concept it became a world leader in the development of large antennas, low-noise receivers, tracking, telemetry and command systems, digital signal processing, and deep-space radio navigation. These functions are incorporated in a state-of-the-art telecommunications system that can be upgraded to meet the requirements of new missions while maintaining support for current missions. This book concentrates on the low-noise front-end of all the DSN receivers and the calibration of the large antennas, whereas other technologies are covered by other books in this Deep Space Communications and Navigation series.

Each of the three DSCCs around the world has one 70-meter (m) antenna, one 26-m antenna, and several 34-m antennas.

A single antenna with its associated equipment is called a Deep Space Station (DSS) and the entire location is called a Deep Space Communications Complex (DSCC). All the stations are remotely operated from a centralized Signal Processing Center (SPC) at each DSCC. Once data has been processed at a DSCC, it is transmitted to the SFOF for further processing and distribution to the various science teams over conventional ground communications networks.

Figure 1-1 is a photograph of the 70-m antenna at Goldstone. In addition to spacecraft tracking, this antenna is used for radio and radar science. The 70-m antennas in Spain and Australia are essentially identical, and they are used for spacecraft tracking and some radio science. The 70-m antennas each have three Cassegrain feedcones in the center of the antenna, which contain the receive and transmit equipment portion of the station. Figures 2-1, 2-2, and 2-3 show the three feedcones, termed a "tricone", on the reflector surface. A slightly

Fig. 1-1. 70-m antenna at Goldstone DSCC.

asymmetric Cassegrain subreflector is used, which can be rotated about the symmetric focal axis of the main reflector. Rotation provides a large outdoor, mechanical, ultralow-noise microwave "optical" or freespace bandswitch for "connection" to any one of the three feedcones. Each feedcone is placed on the Cassegrain focus-circle as formed below the subreflector. Perfect main reflector boresight pointing results for any one of the three feeds used, the asymmetric

subreflector and tricone design causes almost no final or exit beam "scan" loss. In practice only two of the three feedcones are brought into focus for operations. The third is used for engineering purposes and science observations. The S-band feedcone is always directed to the X-band feedcone's dichroic plate and reflected toward the subreflector.

Nearly uniform feed illumination of the main reflector is achieved by the use of shaped dual reflector principles in the asymmetric environment. The combination of switchable tricone microwave feed systems with shaped high-gain efficiency techniques, results in highly flexible large antennas capable of high-performance operation in several internationally-allocated deep space microwave frequency bands with high performance.

X-band receive only, or X-band receive and transmit, operations are conducted with a high pass dichroic frequency-selective reflector placed on top of the X-band (XTR) feedcone, in the retracted position (Fig. 2-2). The XTR feedcone system can transmit with the dichroic plate either extended or retracted. Simultaneous S-band and X-band operations, with or without S-band transmit, are conducted with the dichroic plate in the extended position (Fig. 2-3). X-band passes through the extended dichroic plate while S-band is reflected towards an ellipsoidal reflector on top of the S-band polarization diversity (SPD) feedcone. Thus, by means of two relatively small (sub-sub) reflectors, simultaneous use of two of the three tricone feedcones is obtained, with near-perfect pointing (co-axial S- and X-beam), and essentially no beam scan loss.

The DSN consists of three 70-m antennas, three 34-m High Efficiency (HEF) antennas, six Beam Waveguide (BWG) antennas, one 34-m High Speed Beam Waveguide (HSB), and three 26-m antennas, distributed among the three complexes.

The microwave feed systems of the BWG antennas are placed indoors in a subterranean room, mostly protected from local weather effects, which results in improved link performance during rain. When combined with the uniform illumination shaping technique similar to that of the 70-m antennas, the DSN 34-m BWG antennas are similarly capable of operation with high performance and flexibility of operational frequencies, (Fig. 7-3). A four-reflector, very low-loss 3-m diameter "mirror relay" system in a shielded tube, is used to transfer signals to and from the Cassegrain focus of the main reflector and the secondary focus in the subterranean room. The room contains additional ellipsoidal flat and dichroic reflectors to combine several frequency-band feed systems, which include the low-noise preamplifiers, the high power transmitter, and the control and calibration equipment.

In the HEF antennas, a constant-beamwidth, common aperture S- and X-band (S-/X-band) feedhorn is used at the Cassegrain focus, in a conventional outdoor single-feedcone arrangement. These antennas are called High

Efficiency antennas because they were the first DSN antennas to use the uniform aperture illumination technique.

The DSN 26-m antennas are planned for decommissioning in 2008. They are currently used for tracking near-Earth missions, and this function will be transferred to the 34-m antennas after the decommissioning.

One of the 34-m antennas at the Goldstone DSCC was decommissioned in 1996 and is no longer considered a DSN tracking asset. However, it remains a NASA/DSN asset. The Goldstone Apple Valley Radio Telescope (GAVRT) program was set up as a JPL outreach venture between the Lewis Center for Educational Research (LCER) and JPL, to bring hands-on science to America's classrooms. GAVRT uses the decommissioned, but DSN maintained, antenna at the Goldstone site to conduct radio astronomy observations of scientifically significant objects by remote control, to give the students a real-world experience of scientific discovery, (see, for example, Section 7.6.2 in Chapter 7). By the end of the year 2006, 176 schools in 27 U.S. states, 3 U.S. territories, and 13 foreign countries with U.S. schools, have been involved with the GAVRT program. Much of the equipment and many of the measurement techniques described in this book are applied to the GAVRT observations and system calibrations.

The communication between the ground network and a spacecraft at planetary distances presents special problems, both unique and challenging. Communications performance scales as $1/R^2$, and the extreme distances involved in the exploration of the Solar System make ultralow-noise preamplifiers necessary. To put this in perspective, consider first a typical geostationary communications satellite at an altitude of 40,000 km (4×10^4) above the Earth's surface. The outer planets are many times farther from the Earth, for example the range to Neptune is about 4.5×10^9 km. The distance difference between these two examples varies by more than a factor of 100,000, which implies that the communications challenge for the space research spacecraft is ten orders of magnitude larger than that of the geostationary link.

This book describes the low-noise microwave systems that form the front end of all the DSN ground stations. The microwave front end of each antenna is key to establishing the sensitivity, polarization, frequency diversity, and capabilities of the receiving chain, and therefore of the entire ground station. The receiving system sensitivity and capability is defined by G/T, where G is the antenna gain and T is the overall noise temperature of the entire receiving chain, usually called system operating noise temperature, T_{op}. To improve the station's receiving performance, it is necessary to improve G/T, which can be done by increasing antenna gain, by decreasing T_{op}, or both. In the past the DSN has both increased G and reduced T_{op}. It has been somewhat easier, and

therefore more efficient and cost effective, to decrease T_{op} than it has been to increase the antenna gain.

As the microwave front end of the station is crucial to establishing the sensitivity and capability of the receiving chain, it is therefore incumbent on the designers of the receiving system to expend considerable effort in reducing T_{op} and in calibrating and maintaining the low-noise front end. The more accurately T_{op} can be defined, the narrower the tolerances on the design of a spacecraft mission that can be accepted. The reduction in a project's design tolerances on spacecraft power is worth a considerable amount of money, and enables more science data for the same spacecraft power, and hence the importance of the calibration of the receiving chain. Accurate calibrations are also necessary for the maintenance of the low-noise performance of each antenna system. All missions and radio science benefit from the high gain and extreme sensitivity of the performance of the DSN.

The cost of receiving data from distant spacecraft in dollars per bit is proportional to the receiving system's noise temperature. Each time the system noise temperature is reduced by half, the maximum transferable data rate is doubled, and the cost per bit is reduced by half. Over the years the DSN has developed remarkable innovations to reduce system noise temperature, and many examples have been documented. A 20-percent, or 3-kelvin (K), reduction in the S-band T_{op} enabled the reception of 117 thousand bits per second (kbps) from Mariner 10 during the Mercury flyby in 1974 [2]. This resulted in continuous image coverage of Mercury along the spacecraft's flight path. Without the T_{op} reduction, the mission would have been forced to use a 22-kbps data rate, and 80 percent of the data would have been lost. Support of the Galileo mission is a similar example. Furthermore, an X-band T_{op} reduction of about 20 percent in 1980 contributed to higher data rates from the Voyager spacecraft at Saturn, Uranus, and Neptune. The implementation of the X-/X-/Ka-band feed system in the BWG subnet reduced the X-band diplexed (transmit and receive) T_{op} from 36 K to 18 K for a 3-decibel (dB) improvement in system noise performance [3]. The same technology improved the 70-m X-band T_{op} from 39 K to 15 K for the Cassini mission to realize per day data downlink at Saturn [3].

Mission planners have the option of increasing the data rate from a deep-space mission and thereby reducing the tracking time needed to return the same volume of data, or increasing the total volume of data received. The percentage reduction in tracking time can be the same percentage by which the T_{op} is reduced. A reduction in the tracking time results in reduced costs to a mission for tracking support. Use of the time saved with an increase in the data return

enables a greater percentage of data to be returned than the percentage reduction in tracking time; a 20-percent reduction in T_{op} enables a 25-percent increase in data rate. The value of a T_{op} reduction can be based on the costs of using DSN antennas to track deep-space missions. Hourly rates for the use of DSN antennas have been estimated for mission planning purposes. These rates, combined for the DSN's various antennas were used to show that a 20-percent reduction in T_{op}, 4 K at X-band, could save a few tens of millions of dollars per year [4]. A reduction in T_{op} also means a higher antenna utilization time.

Another way to estimate the value of a T_{op} reduction considers the initial cost and the operations cost of a fully equipped 34-m BWG antenna over, say, a 20-year period. The construction and implementation cost of approximately \$33M combined with a \$2M per year operations cost over the 20-year lifetime gives a total lifetime cost of about \$73M. Adding five such 34-m antennas to the DSN increases the total data reception capability by about 25 percent. Reducing the X-band system noise temperature by 4 K, or 20 percent, increases the DSN's total data reception capability by about 25 percent. By this measure, a 1-K reduction at X-band throughout the DSN is worth about \$91M, and, in this case, a 20-percent reduction in T_{op} appears to be worth about \$365M [4].

The conclusion is clear; noise in the DSN is extremely expensive. The DSN has devoted considerable time and effort to reducing T_{op} and to the concomitant activity of accurately calibrating T_{op}. Interesting and innovative concepts and methods have been developed in both these projects. This book describes and discusses these developments.

Chapter 2 begins with a standard, but brief, review of noise in receiving systems in general, and goes on to describe all the noise element contributions, both internal and external, that make up the total noise in the system operating noise temperature. All noise contributions are detailed in their measurement and calibration, with emphasis on the achievement of maximum accuracy and precision. Each step in the process is analyzed with mathematical rigor, so that the reader who wishes to follow the logic trail throughout the system can do so compactly. On the other hand, the method of measurements and calibrations is described fully, including gain fluctuations and system linearity. This gives an overall view of what needs to be done, how it is done, and why it is important. Step-by-step details allow an operator to carry out the measurements and calibrations and derive the results of the statistical error analyses with little, or even no consideration of the mathematics or theory. Many carefully selected references are provided for those readers who want more detail or who like to read around each topic.

The ruby maser microwave preamplifiers in the DSN are linear, phase-preserving, and ultralow-noise. They are the most sensitive and lowest noise amplifiers used in the field. Yet over many years in the DSN they have proved to be rugged and stable, as well as providing a natural spectrum filtering capability few other amplifiers can match in low-noise systems. Chapter 3 introduces ruby masers in general and the relevant properties of ruby. The mechanism of microwave amplification by stimulated emission of radiation (maser) is described with some theory to give an overview of the entire amplification process. The remarkable low-noise properties of masers is shown to be derived primarily from the amplification process, but also from the innovative engineering developed by the DSN to integrate the masers into the front-end subsystem. Cavity masers, reflected wave masers, and comb-type masers are included. Detailed mathematical theory of the maser process is provided for completeness, including the noise temperature performance and gain dependence of masers, together with references for further study if needed.

Chapter 4 completes the maser discussion with a description and analysis of the JPL-developed closed-cycle refrigerator (CCR), and other refrigeration systems used to support the masers. The chapter describes radiation shields, heat leaks, and other engineering details needed for the analysis and operation of CCRs in the field.

Chapter 5 is a comprehensive discussion of high electron mobility transistors (HEMT), and it includes theory, practical operation, calibrations, device fabrication and optimization, noise modeling and characterization, amplifier development, and subsystem measurements. The chapter gives a complete overview of the entire subject of HEMT amplifiers with mathematical derivations, details of materials and fabrication, amplifier design and operation, and measurements and calibrations. The chapter can be used both as a HEMT textbook for detailed study, as well as an introduction to the use of HEMTs for operators and managers who only need an overview of their fabrication and use in the field.

The atmosphere has a significant effect on the communications link between the ground and the spacecraft, especially at the higher operating frequencies. The atmosphere attenuates the signal and adds noise temperature to the ground receiving system, and changing weather conditions change these parameters. Sometimes large fluctuations of the parameters are experienced. Chapter 6 is a description and discussion of how the DSN handles the problem of changing weather conditions. The DSN uses direct weather measurements by water vapor radiometers (WVR), along with meteorological forecasts and a JPL-derived standard surface weather model. The WVRs are described, and the theory of the derivation of the required parameters is presented. The development of the JPL standard weather model is described, including how it is used together with meteorological forecasts.

This book is focused on the system operating noise temperature, and the need for and the importance of, the measurement and calibration thereof. However, as stated above, the system sensitivity and quality are defined by G/T, where G is the antenna gain and T is the total system noise temperature. The book would not be complete without a discussion of the measurement and calibration of antenna gain and efficiency. Chapter 7 is an overview of this extensive subject, which will be treated in more detail in a future book in this series. This chapter uses the principles of noise temperature measurements, as described in Chapter 2, to derive antenna gain, systematic pointing corrections, subreflector focus, and calibration of radio sources used in the antenna gain assessment.

With new technology developed at JPL, it has been possible to obtain measurement precision yielding as much as an order of magnitude improvement over previous methods in the determination of antenna aperture efficiency, and factors of five or more in the determination of pointing errors and antenna beamwidth. This raster-scan improvement has been achieved by performing continuous, rapid, raster scans of both extended and point sources. This mapping technique for antenna and radio source calibration is called "on-the-fly" mapping. The chapter begins with a description of the general requirements for DSN antenna calibrations, followed by a discussion of current methods and their shortcomings. The final section describes the new mapping technique with some results. Two of the most important parameters that determine antenna gain are the alignments of the main reflector panels and antenna stability. These are best measured by coherent holographic techniques, which are described in the final chapter, Chapter 8.

When applied to reflector antennas, microwave holography is a technique that uses the Fourier transform between the complex far-field radiation pattern of an antenna and the complex aperture distribution. The result of this operation is a data set of aperture phase and amplitude distributions. These data are used to precisely characterize crucial performance parameters, which include panel alignment, subreflector position, antenna aperture illumination, directivity at various frequencies, and the effects of gravity deformation. Thus, the holography technique provides a method for analysis, evaluation, and RF performance improvement of large reflector and beam waveguide antennas. Microwave holography is one of the most economical techniques for increasing the performance of the large DSN antennas in terms of cost-to-performance ratio. Chapter 8 describes the design and use of the holographic measurement system and its use in the DSN to improve, optimize, and maintain its performance to prescribed specifications. The chapter describes the design of the holographic instruments, and the mathematical algorithm and software for the development of the holographic measurement system. Microwave holography is one of the important items in the DSN suite of calibration instruments and methodologies.

To the extent possible, this book adheres to Institute of Electrical and Electronics Engineers (IEEE) standard terminology definitions [5]. This is an important consideration for clarity, understanding, and cross-referencing throughout the text. An example of the difficulty that arises when authors do not use standard terminology is given in Mumford and Scheibe's book, Noise; Performance Factors in Communications Systems [6], where they list nine different definitions of amplifier noise factor in terms of noise temperature used by various authors. This clearly shows the importance of using IEEE noise standards [5,7] for the definition of receiving system noise terms.

Each chapter in this book can be read in three levels. Rigorous mathematical theory of the devices, their operation, and calibration methods, is one level. An overview of the design, and in some cases the fabrication of the devices, how they are constructed and used in the field, their measurements and calibrations, form a second level. This level will be useful for engineers and supervisors. In addition, the book can be read on a third level of advice, instructions, and methods. The information presented at this third level was derived from much experience in the field, and this level will be useful for technicians and operators.

References

[1] C. R. Koppes, *JPL and the American Space Program*, Yale University Press, New Haven, Connecticutt and London, United Kingdom, 1982.

[2] W. R. Corliss, *The History of the Deep Space Network*, NASA, CR 151915, National Aeronautics and Space Administration, Washington, District of Columbia, 1976.

[3] J. Bautista, R. Clauss, S. Petty, and J. Shell, "DSN Low Noise Amplifiers in the New Millennium," *TMOD Technology and Science Program News*, Issue 13, Rev. 1, Jet Propulsion Laboratory, Pasadena, California, January 2001. http://tmot.jpl.nasa.gov/index.html

[4] J. Layland and L. Rauch, "The Evolution of Technology in the Deep Space Network: A History of the Advanced Systems Program," *The Telecommunications and Data Progress Report 42-130, April–June 1997*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–44, August 15, 1997, (also available as JPL publication 95-20 at http://tmot.jpl.nasa.gov/, click on Program Overview Information). http://ipnpr.jpl.nasa.gov/progress_report/

[5] *The Authoritative Dictionary of IEEE Standard Terms*, 7th ed., *IEEE Standards 100*, Hoboken, New Jersey, 2000.

[6] W. W. Mumford and E. H. Scheibe, *Noise; Performance Factors in Communications Systems*, Horizon House–Microwave Inc., Dedham, Massachusetts, 1968.

[7] "IRE Standards on Electron Tubes: Definition of Terms, 1962 (62 IRE 7.S2)," *Proceedings of the IEEE*, vol. 51, pp. 434–435, March 1963.

Chapter 2
System Noise Concepts with DSN Applications

Charles T. Stelzried, Arthur J. Freiley, and Macgregor S. Reid

2.1 Introduction

The National Aeronautics and Space Administration/Jet Propulsion Laboratory (NASA/JPL) Deep Space Network (DSN) has been evolving toward higher operational frequencies for improved receiving performance for the NASA deep space exploration of our Solar System. S-band (2.29–2.30 gigahertz [GHz]), X-band (8.40–8.50 GHz), and Ka-band (31.8–32.3 GHz) are the 2006 de ep space downlink (ground receive) frequency bands. These internationally allocated microwave downlink frequency bands are listed in Table 2-1. The DSN is considering the use of higher microwave frequencies as well as optical frequencies for the future. Higher frequencies generally provide improved link performance, thus allowing higher data rates.

The communications link performance is c ritically dependent upon the receiving system antenna gain (proportional to a ntenna area) and the noise temperature performance; a system figure of merit of gain-to-temperature ratio (G/T) is de fined in S ection 2.5.1. This chapter provides the definitions and calibration techniques for measurements of the low noise receiving systems. Noise in a receiving system is defined as an undesirable disturbance corrupting the information content. The sources of noise can be separated into external and internal noise. Sources of external noise [1–5] include Cosmic Microwave Background (CMB), Cosmic Microwave Foreground (CMF), galactic and radio sources, solar, lunar, and planetary sources, atmospherics (includes lightening), atmospheric absorption, man-made noise, and unwanted antenna pickup. The

Table 2-1. 2006 DSN allocated operational deep space downlink microwave frequencies.

DSN Microwave Band	Mid-Band Frequencies (GHz)	Allocated Frequencies (GHz)
S	2.295	2.29–2.30
X	8.425	8.40–8.45
Ka	32.05	31.8–32.3

CMB noise [6] is considered to be the residual radiation, currently measured to be 2.725 K, [7] from the origin of the universe (Big Bang). CMF noise includes emission from diffuse Galactic and unresolved point sources [8] collected by the antenna. Man-made noise or radio frequency interference (RFI) includes coherent signal generation as well as ignition systems, spark discharges, and transmission of noise signals emanating from power transmission lines. The CMF contribution is small compared to the CMB for DSN applications. The CMB and the atmosphere are the dominant external noise sources for the DSN microwave receiving systems.

Sources of internal noise include [9, 10, and 4 (p. 4)] thermal, shot, current, and Barkhausen noise. Thermal noise is the minimum internal noise source for microwave receiving systems. The noise performance of DSN low-noise amplifiers (LNAs) is typically expressed in terms of noise temperature (Section 2.2.6). For example, LNAs with a noise figure range of 1.010 to 1.014 (ratio) are more conveniently characterized with a noise temperature range of 3 to 4 kelvins.

Thin sheet Kapton at the feedhorn apertures retains the nitrogen pressure in the feedhorns and associated waveguide components. On the 70-meter (70-m) tricone-equipped antennas, blowers are installed at the X-band feedhorn aperture in an attempt to reduce signal attenuation and increased noise contribution due to rain or dew effects on the horn covers. Rain adhering to the weathered Kapton feedhorn cover causes measurable degradation effects at all DSN frequencies, including S-band. An important strategy is to maintain fresh Kapton horn covers to minimize the degradation. These precautions are not necessary with the DSN 34-m beam waveguide (BWG) antennas since the feedhorns are not exposed to weather.

In this chapter, system noise temperature contributions and measurement techniques are presented and analyzed for typical receiving systems of the DSN. The antenna, LNA, and receiver assemblies comprise the receiving system. The DSN adheres to the Institute of Electrical and Electronics Engineers (IEEE) noise temperature definitions [11,12]. Mismatch effects are minimized with the use of matched components in the LNA input configuration [13].

The large antennas of the DSN required for deep space communications are described in [14,15] and Chapter 1 of this book. The 70-m antennas are the

largest and most powerful in the DSN with three [16] Cassegrain feedcones ('tricone') shown in Fig. 2-1. The X-band transmit receive (XTR) feedcone has a high-pass dichroic frequency selective reflector (i.e., it pa sses X-band and reflects S-band), shown in the retracted position (Fig. 2-2) for X-band only

Fig. 2-1. Goldstone 70-m antenna tricones S-band polarization diversity (SPD) feedcone located behind the X-band/K-band radar (XKR) and X-band transmit/receive (XTR) feedcone.

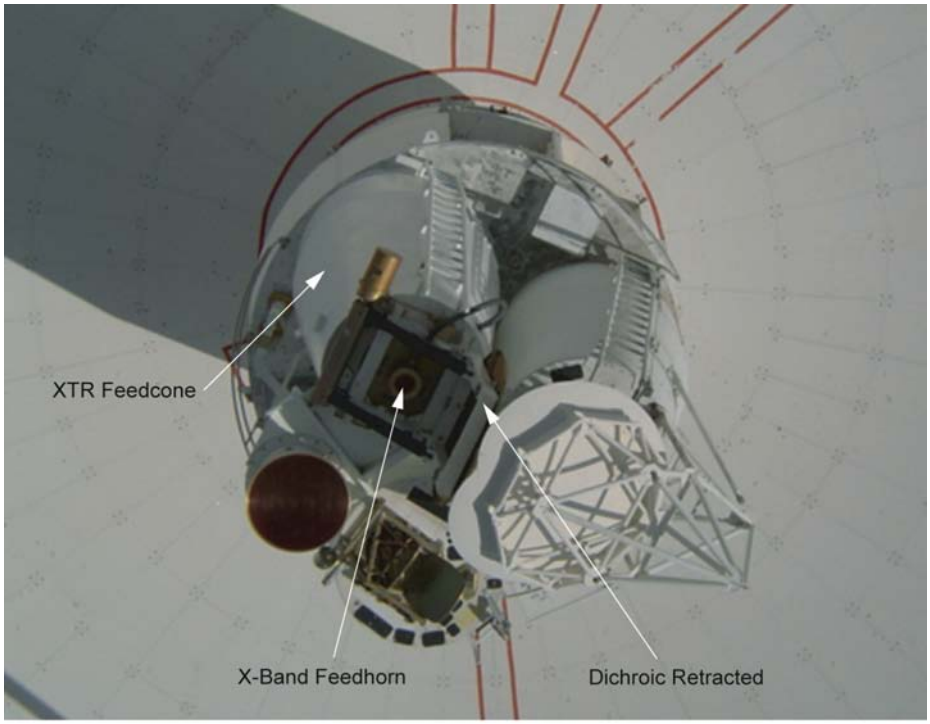


Fig. 2-2. Goldstone 70-m antenna tricones showing XTR feedcone with S-band/X-band dichroic retracted as viewed from the subreflector.

operations. This reflector is moved to the extended position (Fig. 2-3) for simultaneous S-X-band operation. Thus, by means of two relatively small (sub-sub) reflectors, simultaneous use of two of the three feedcones is obtained, with nearly coincident beam (co-axial S-X beam) pointing and essentially no beam scan loss. This XTR feedcone configuration is used in Section 2.5 as an example of DSN noise temperature system calibration. These noise temperature calibrations of the LNA with a standard calibrated zenith pointing feedhorn (test a), a completed feedcone assembly in a ground zenith configuration (test b) and the completed system on the antenna (test c) results in a fully calibrated system. The system noise temperatures of the operational systems are verified with routine periodic measurements.

The system operating noise temperature (T_{op}) for a DSN receiving system is measured by switching the LNA between a calibration ambient load and the antenna. T_{op} arises from multiple contributions defined at the feedhorn aperture.

$$T_{\mathrm{op}} = T_{\mathrm{sky}} + T_{\mathrm{ant}} + T_{\mathrm{feed}} + T_{\mathrm{LNA}} + T_f \qquad (2.1\text{-}1)$$

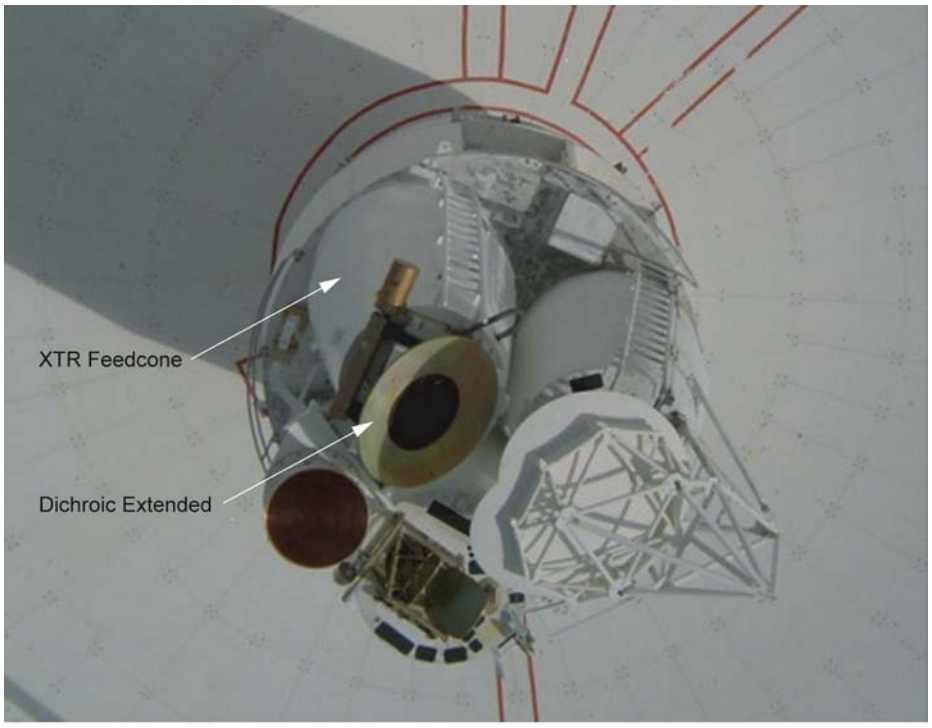


Fig. 2-3. Goldstone 70-m antenna tricones showing XTR feedcone with S-band/X-band dichroic extended as viewed from the subreflector.

T_{op} is composed of the sky noise temperature and the ground antenna and microwave system

$$T_{op} = T_{sky} + T_{AMW} \tag{2.1-2}$$

where

T_{op} system operating noise temperature, K

$T_{sky} = T_{atm} + T_{CMB} / L_{atm}$ = combined noise temperature contribution of T_{atm} and T_{CMB}, K

T_{CMB} = CMB noise temperature, K

T_{atm} = noise temperature contribution due to the atmosphere, K

L_{atm} = atmospheric loss, ratio

T_{ant} = noise temperature contribution due to the antenna, K

T_{feed} = noise temperature of the microwave feed components, K

T_{LNA} = LNA noise temperature, K

T_f = follow-up amplifier noise temperature, K

$T_{\text{AMW}} = T_{\text{ant}} + T_{\text{feed}} + T_{\text{LNA}} + T_f$ = antenna microwave system effective input noise temperature, K

2.2 Noise Temperature Concepts

2.2.1 Thermal Noise

The concept of thermal noise is stated by W. Mumford [4 (p. 3)] as: "The random motion of the free electrons in a conductor caused by thermal agitation gives rise to a voltage at the open ends of the conductor. In most conductors, the frequency components of this noise cover the whole radio spectrum uniformly."

Thermal or Johnson [9] noise power density available from a source such as an antenna or a resistive termination (load) is given by [11 (p. 735)]

$$N = \text{k}\, T_p \tag{2.2-1}$$

where

N = noise power density, W/Hz

k = Boltzmann's constant = 1.38065×10^{-23}, J/K

T_p = resistive termination physical temperature, K

A reduction of noise power at higher frequencies according to Planck's radiation law is ignored for present DSN applications. Accurate system operating noise temperature results (as analyzed and verified in Section 2.2.3) are obtained without application of Planck's radiation law noise temperature reduction. This simplifies the analysis and calculations, and is in accordance with IEEE standards [11 (p. 734)].

Connecting a resistive source to the input of a noiseless amplifier (with available power gain G and noise bandwidth B) provides amplifier output noise power N_o [4 (p. 12)]

$$N_o = \text{k}\, T_p\, BG \tag{2.2-2}$$

where

N_o = amplifier output noise power with resistive source and noiseless amplifier, W

$B = (1/G_m)\int G(f)df$ = noise bandwidth [4 (p. 13)], Hz

f = operating frequency, Hz

$G = G(f)$ = available power gain, ratio

G_m = maximum available power gain [4 (p. 8)], ratio

k = Boltzmann's constant = 1.38065×10^{-23}, J/K

T_p = resistive termination physical temperature, K

The standard ambient noise temperature is $T_o = 290$ K, which from Eq. 2.2-1 gives an available noise power of –203.975 dB relative to 1 W/Hz-K. Similarly, $T = 1$ K results in a noise power of –228.599 dB relative to 1 W/Hz-K (or –198.6 dBm relative to 1 mw/Hz-K).

2.2.2 System Operating Noise Temperature

The preceding discussion and definitions lead to a system operating noise temperature concept. System operating noise temperature is important for the DSN, and for other operational systems, for the determination of the operating communications system performance between the spacecraft and the ground antennas. Although this discussion focuses on the downlink analysis, the concepts apply to the uplink analysis as well.

Consistent with Eq. (2.2-2), and Fig. 2-4, system operating noise temperature (T_{op}) is given by [4 (p. 34)]

$$T_{\text{op}} = \frac{P_o}{\text{k}BG} \tag{2.2-3}$$

where

T_{op} = system operating noise temperature, K

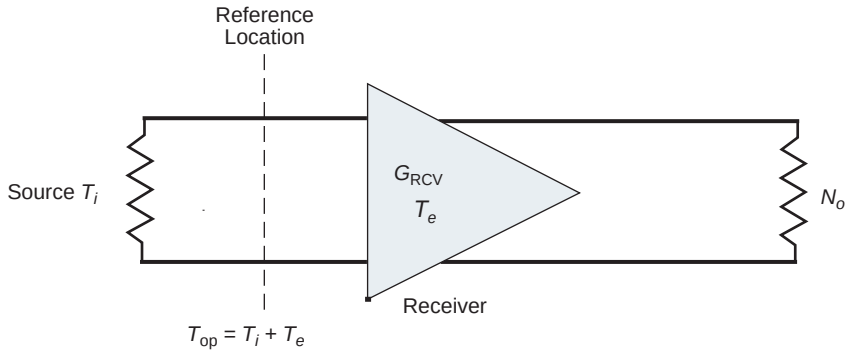


Fig. 2-4. Receiving system consisting of an input source (typically an antenna or ambient resistive load) and an amplifier with a single reference location for defining the system noise temperature, T_{op}, in terms of T_i and T_e.

P_o = receiver output noise power, W

T_{op} is always defined for a specific reference location as shown in Fig. 2-4. T_{op} is composed of both an incident input noise temperature and the receiver effective input noise temperature [11 (p. 766)].

$$T_{\text{op}} = T_i + T_e \tag{2.2-4}$$

where

T_{op} = system operating noise temperature, K

T_i = input source noise temperature, K

T_e = receiver effective input noise temperature, K

2.2.3 Planck's Radiation Law Noise Power Reduction

The purpose of this section is to evaluate Planck's radiation law noise power reduction at higher frequencies on measurements of system noise temperature T_{op}. The correction for T_{op} that results from the reduced noise power is analyzed for a range of frequencies and temperatures. For microwave frequencies below 100 GHz, this correction is not necessary, nor is it used for DSN calibrations [18].

The noise temperature reduction for a resistive calibration termination as a function of frequency is shown in Fig. 2-5 using Planck's radiation law noise power reduction [4 (p. 76)] is given by

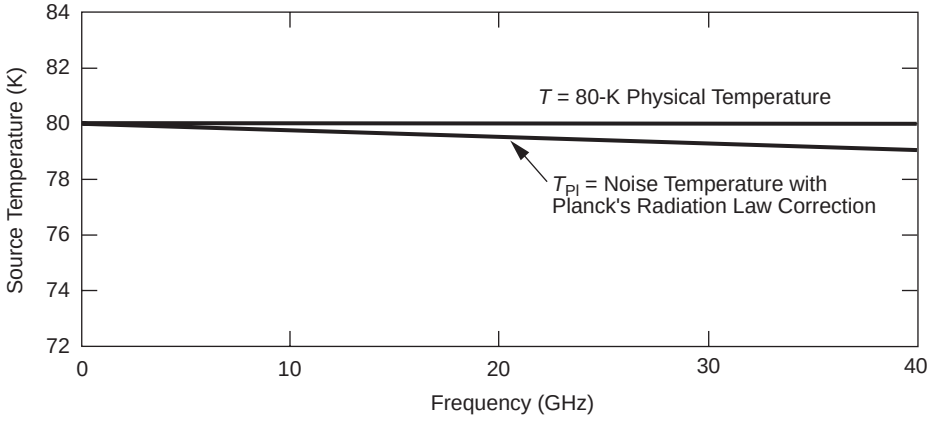


Fig. 2-5. Source noise temperature at a physical temperature of 80 K (horizontal line) and showing Planck's radiation law noise temperature reduction versus frequency.

$$T_{\mathrm{Pl}} = T\left[\frac{x}{e^{x}-1}\right] \tag{2.2-5}$$

where

T_{Pl} = source noise temperature with Planck's radiation law correction, K

T = source physical temperature, K

h = Planck's constant = 6.626069×10^{-34}, Js

f = frequency, Hz (unless noted otherwise)

x = hf/kT ~ 0.048 f(GHz)/T(K), ratio

From Eq. (2.2-5) at 32 GHz, Planck's radiation law reduces the noise temperature of a resistive termination with a physical temperature of 80 K by 0.765421 K. However, an accurate system temperature measurement results without the Planck's correction as shown in the following equations. This results from a lower measured receiver effective input noise temperature, which nearly cancels the higher antenna noise temperature.

As discussed, and with terms defined in Section 2.6, a receiver effective input noise temperature calibration consists of measuring the '*Y*-factor' power ratio by switching the receiver input between the cold and hot loads.

$$Y_{ch} = \frac{T_h + T_e}{T_c + T_e} \tag{2.2-6}$$

Solving for T_e in terms of the known cold and hot loads noise temperatures T_c and T_h

$$T_e = \frac{T_h - Y_{ch}T_c}{Y_{ch} - 1} \tag{2.2-7}$$

Also, system noise temperature, as discussed and defined in Section 2.6, can be determined by measuring the *Y*-factor power ratio by switching the receiver input between the hot load and the antenna

$$T_{\mathrm{op}} = \frac{T_h + T_e}{Y_{ah}} \tag{2.2-8}$$

Combining Eqs. (2.2-7) and (2.2-8)

$$T_{\mathrm{op}} = \frac{Y_{ch}\left(T_h - T_c\right)}{Y_{ah}\left(Y_{ch} - 1\right)} \tag{2.2-9}$$

using $Y = Y_{ch} / \left(Y_{ah} \left(Y_{ch} - 1 \right) \right)$, then

$$T_{\text{op}} = Y \left(T_h - T_c \right) \tag{2.2-10}$$

where

$Y = Y_{ch} / [Y_{ah}(Y_{ch} - 1)]$, ratio

$Y_{ch} = P_h / P_c =$ hot (ambient) and cold loads *Y*-factor, ratio

$Y_{ah} = P_h / P_a =$ hot (ambient) load and antenna *Y*-factor, ratio

The *Y*-factor ratio values are obtained from the calibration measurements and do not de pend on whether or not it is pla nned to account for Planck's correction. The difference in T_{op} using the physical temperature of the loads versus Planck's correction is

$$\Delta T_{\text{op}}, \text{K} = T_{\text{op}} - T_{\text{opPl}} \tag{2.2-11}$$

$$\Delta T_{\text{op}}, \text{K} = Y \left[\left(T_h - T_c \right) - \left(T_h - T_c \right)_{\text{Pl}} \right] \tag{2.2-12}$$

$$\Delta T_{\text{op}}, \% = 100 \frac{\left(T_h - T_c \right) - \left(T_h - T_c \right)_{\text{Pl}}}{\left(T_h - T_c \right)} \tag{2.2-13}$$

where

$\left(T_h - T_c \right) =$ calibration loads physical temperatures difference, K

$\left(T_h - T_c \right)_{\text{Pl}} =$ calibration load noise temperature difference with Planck's reductions, K

Figure 2-6 shows a plot of ΔT_{op} as a percentage versus frequency. These curves are well behaved. The percentage difference in T_{op} due to using the load's physical temperatures versus Planck's reduction corrections for this example with 80-K and 290-K cold and hot c alibration loads is only 0.00085 percent at 32 GHz. (The difference is la rger at higher frequencies and lower cold load temperatures). The difference is le ss than 0.026 percent at 40 GHz with a 4-K cold load, the worst case shown for this range of parameters.

Expanding Planck's radiation law noise power reduction Eq. (2.2-5) [19 (p. 132)] in a Taylor series, assuming $x^2 < 4\pi^2$ where $x =$ h*f*/*k*T

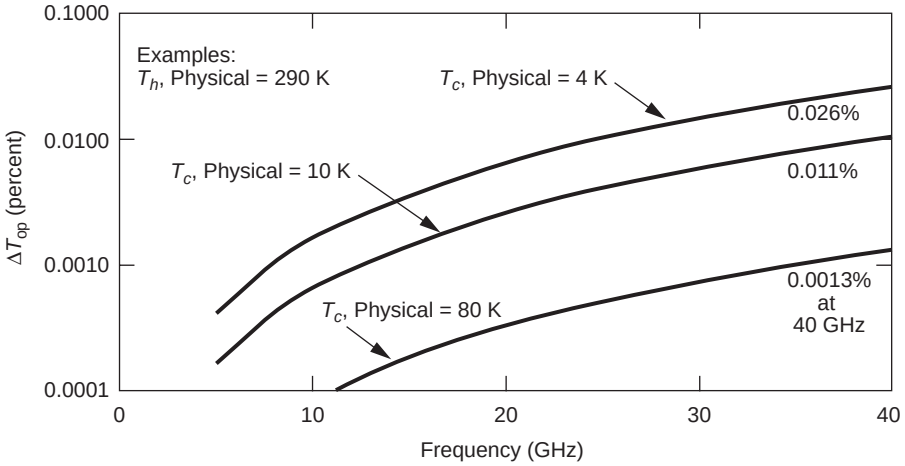


Fig. 2-6. System noise temperature measurement difference, ΔT_{op} in percent versus frequency, due to Planck's radiation law correction using a 290-K hot load and a range of cold load temperature.

$$T_{\mathrm{Pl}} \simeq T\left[1-\left(\frac{x}{2}\right)+\left(\frac{x^2}{12}\right)+\cdots\right] \tag{2.2-14}$$

$$T_{\mathrm{Pl}} \simeq T-\frac{\left(\frac{\mathrm{h}f}{\mathrm{k}}\right)}{2}+\frac{\left(\frac{\mathrm{h}f}{\mathrm{k}}\right)^2}{12T}+\cdots \tag{2.2-15}$$

Using Eq. (2.2-13) with Eq. (2.2-15) ignoring higher order terms results in cancellation of the first two terms. This leaves only the third small term

$$\Delta T_{\mathrm{op}},\% \simeq \frac{100\left(\frac{\mathrm{h}f}{\mathrm{k}}\right)^2}{12T_cT_h} \tag{2.2-16}$$

Eq. (2.2-16) shows that the error ΔT_{op} (in percent) is increased at higher frequencies and lower calibration load temperatures. The value for ΔT_{op} (in percent) provides the error term for the system error analysis due to neglecting Planck's correction. For the DSN this error term is much smaller than the total errors due to other causes. At 32 GHz with $T_c = 80$ K and $T_h = 290$ K, Eq. (2.2-16) gives T_{op} (percentage = 0.00085) in agreement with the previous example with Eq. (2.2-13) and shown in Fig. 2-6.

The Planck correction with frequency is not needed for DSN operational microwave frequencies. The DSN reports measurements and analyses of system noise performance, at microwave frequencies, using the IEEE definition [11 (p. 735)] with the noise temperature of the calibration loads defined by their physical temperatures. This includes treating the CMB noise temperature as a constant 2.725 K over the range of DSN operational frequencies, all less than 100 GHz.

2.2.4 Translating Noise Temperature Reference Locations

System noise temperatures of a receiving system are defined for a specific reference location. Separate incident input and receiver effective noise temperatures are associated with the system noise temperature for that same given reference location. Typically, the individual noise temperature contributions are determined at other reference locations and must be translated to the common reference location. The equations for the various needed translations follow. It is assumed that the system is linear and that the components are 'matched.' For the three separate reference locations 1, 2, and 3 shown in Fig. 2-7, system noise temperatures by definition are

$$T_{\mathrm{op}1} = T_{i1} + T_{e1} \tag{2.2-17}$$

$$T_{\mathrm{op}2} = T_{i2} + T_{e2} \tag{2.2-18}$$

$$T_{\mathrm{op}3} = T_{i3} + T_{e3} \tag{2.2-19}$$

T_i contains all the noise sources to the input side, and T_e contains all the noise sources to the output side of the reference location. With a component loss L (ratio, equal to or greater than 1) between the noise reference locations 1 and 2, and gain G (ratio, equal to or greater than 1) between the noise reference locations 2 and 3

$$T_{\mathrm{op}1} = LT_{\mathrm{op}2} \tag{2.2-20}$$

$$T_{\mathrm{op}2} = \frac{T_{\mathrm{op}1}}{L} \tag{2.2-21}$$

$$T_{\mathrm{op}3} = GT_{\mathrm{op}2} \tag{2.2-22}$$

G (ratio) and L (ratio) are always equal to or greater than 1, so that $T_{\mathrm{op}1}$ and $T_{\mathrm{op}3}$ are always equal to or greater than $T_{\mathrm{op}2}$.

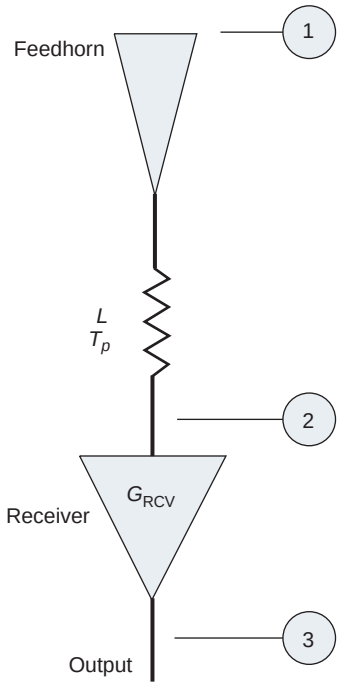


Fig. 2-7. Receiving system noise temperatures are defined at specific locations: Feedhorn aperture at reference location 1, receiver input at reference location 2, receiver output at reference location 3.

From Mumford [4 (p. 23)], the noise temperature translation equation between reference locations 1 and 2 separated by a loss L, ratio, with T_p, K equal to the physical temperature of the loss

$$T_{e1} = LT_{e2} + (L-1)T_p \tag{2.2-23}$$

Solving for T_{e2}

$$T_{e2} = \frac{T_{e1}}{L} - \left(1 - \frac{1}{L}\right)T_p \tag{2.2-24}$$

With Eqs. (2.2-4, -20, and -23)

$$T_{i1} = LT_{i2} - (L-1)T_p \tag{2.2-25}$$

Solving for T_{i2}

$$T_{i2} = \left(\frac{T_{i1}}{L}\right) + \left(1 - \frac{1}{L}\right) T_p \tag{2.2-26}$$

In the above four equations, the noise temperature contribution of the components with loss can be written $T_{L1} = (L-1)T_p$ and $T_{L2} = T_{L1} / L = (1-1/L)T_p$.

For a receiving system with an LNA, feed loss and follow-up amplifiers, the noise temperature contributions are

$$T_{e1} = T_{\text{LNA1}} + T_{\text{feed1}} + T_{f1} \tag{2.2-27}$$

and

$$T_{e2} = T_{\text{LNA2}} + T_{f2} \tag{2.2-28}$$

where

T_{LNA1} = LNA noise temperature at the feedhorn aperture, reference location 1, K

$T_{\text{LNA2}} = (L_{\text{feed1}} - 1)T_p$ = LNA noise temperature at the LNA input, reference location 2, K

T_{feed1} = loss noise temperature contribution at the loss input, reference location 1, K

T_{f1} = the follow-up amplifiers noise temperature contribution, reference location 1, K

T_{f2} = the follow-up amplifiers noise temperature contribution at the reference location 2 (LNA input), K

The technique for solving noise temperature translations is to use the above Eqs. (2.2-17 through 2.2-26) in terms of T_i and T_e, obtaining T_i from the input noise sources and T_e from the receiving system noise sources as in Eqs. (2.2-27 and 2.2-28).

2.2.5 Noise Temperature and Loss Contributions

Low loss microwave components at room temperature (~ 290 K) contribute ~ 6.7 K noise temperature for 0.1 dB attenuation. Therefore, it is important to keep losses low to achieve lower noise temperature systems. Accurate measurements (Section 2.6) of the transmission components are essential for

the development of low noise components with minimum loss. System design and component selection are equally important.

2.2.6 Receiver Noise Temperature and Noise Figure

Although the noise figure (sometimes called noise factor) concept is not usually used with low noise systems (Section 2.1), the relationship is provided here for convenience. Noise temperature and noise figure for single response receivers are related by [4 (pp. 54, 55, Table 1 first listed definition)]

$$F = 1 + \left(\frac{T_e}{290} \right) \tag{2.2-29}$$

Solving for T_e

$$T_e = (F - 1)\, 290 \tag{2.2-30}$$

where

F = receiver noise figure, ratio

T_e = effective receiver input noise temperature, K

Mumford [4 (pp. 54, 55, Table 1)] lists nine separate definitions from the literature for the amplifier noise figure. This clearly shows the importance of noise standards.

2.3 Antennas

2.3.1 Antenna Noise Temperature

Following Eq. (2.2-1), the antenna noise temperature T_a is defined

$$T_a = \frac{N_a}{\mathrm{k}} \tag{2.3-1}$$

where

T_a = antenna noise temperature, K

N_a = noise power density delivered by the antenna into a matched termination, W/Hz

k = Boltzmann's constant = 1.38065×10^{-23}, J/K

Antenna noise temperature is given by the integral of the product of the surrounding external physical temperature distribution convolved with the antenna response. As given by Rusch [20 (p.73)], in terms of the antenna external blackbody temperature and antenna gain, both in the direction Ω,

$$T_a = \int_{4\pi} T(\Omega) G(\Omega) d(\Omega) \tag{2.3-2}$$

where

$T(\Omega)$ = equivalent blackbody temperature of area $d(\Omega)$ in direction Ω, K

$G(\Omega)$ = antenna gain in direction Ω, ratio

In terms of the normalized antenna radiation pattern $P(\Omega)$

$$T_a = \frac{\int_{4\pi} T(\Omega) P(\Omega) d(\Omega)}{\int_{4\pi} P(\Omega) d(\Omega)} \tag{2.3-3}$$

Both downlink DSN antenna gain and noise temperature are defined and referenced at either the LNA input or at the feedhorn input (Section 2.5). The maximum antenna gain relative to an isotropic radiator is given by

$$G_m = \frac{4\pi A_e}{\lambda^2} \tag{2.3-4}$$

where

G_m = maximum antenna gain relative to isotropic radiator, ratio

$G_m(\text{dBi})$ = maximum antenna gain relative to isotropic radiator $= 10\log(G_m)$

$A_e = \varepsilon A_{p.} = \pi D_e^2 / 4 =$ antenna effective area, m^2

ε = antenna gain efficiency (less than 1), ratio

$A_p = \pi D_e^2 / 4 =$ physical area, m^2

λ = wavelength, m

D_p = antenna physical diameter, m

D_e = antenna effective diameter, m

It is difficult to determine the noise temperature contribution of the antenna to the system noise temperature by integration of the convolution of the antenna pattern with the noise source distributions, Eq. (2.3-2). There are multiple noise temperature sources, from various antenna structures, contributing to the total antenna noise temperature variation with antenna elevation angle. The forward and rear antenna spillovers intercept the ground and sky differently as a function of elevation angle. The antenna reflecting surfaces contribute to the

system noise temperature as a function of operating frequency from the ohmic I^2R losses of reflection. The leakage through the surface holes and cracks of the reflecting panels, also contribute to the antenna noise temperature [21–23].

A practical approach used by the DSN for determining the antenna contribution to the system noise temperature is to measure the increased noise temperature between the LNA receiving system on the ground and when it is mounted on the antenna. The noise temperature contribution of the DSN antennas has been measured in this manner for many decades for different frequencies and LNA receiving systems. The atmospheric contributions for the ground and antenna-mounted feedcone configurations are measured separately and accounted for, so the increased noise due to the antenna is determined quite accurately.

2.3.2 DSN Antennas

A performance handbook is available [24] for the Deep Space Network (DSN). The handbook modules provide performance values of the uplink and downlink telecommunications link parameters between spacecraft and the DSN. These modules are updated as new capabilities and data (such as the antenna gain and noise temperature) become available. These modules document the DSN antennas at each of the operating frequencies in more detail than appropriate for this chapter. Current downlink values for the Goldstone, California, station antennas parameters are listed in Table 2-2. The Australian and Spanish antennas have similar values; these include the effect of the higher atmospheric losses.

The front-end loss for a DSN antenna feed assembly is determined by the difference between the system noise temperature at the feedhorn aperture

Table 2-2. DSN Goldstone large ground antennas downlink performance for 25-percent weather and zenith antenna-pointing elevation angle.

Antenna	Freq. Band	Gain (dBi)	Noise Temp. (K)	G/T (dB)	HPBW (deg)
34-m BWG	S	56.8	36.8	41.1	0.23
	X	68.0	33.0	52.9	0.063
	Ka	78.5	31.0	63.6	0.017
34-m HEF	S	56.0	38.0	40.2	0.23
	X	68.1	19.8	55.1	0.063
	Ka	NA	NA	NA	NA
70-m	S	63.4	22.0	50.0	0.11
	X	74.4	20.6	61.3	0.031
	Ka	NA	NA	NA	NA

(system reference location 1), and at the LNA input (system reference location 2). Reference location 1 at the feedhorn aperture is shown in Figs. 2-2 (X-band feedhorn) and 2-8 (feedhorn aperture). Reference location 2 at the LNA input is shown in Fig. 2-12. The DSN calibration procedure for feedcone assembly noise temperature measurements is described in Section 2.5. For precise calibrations of systems with distributed loss between these reference locations it is necessary to translate the various noise temperatures to a common system reference location. This is accomplished with the equations summarized in Table 2-3.

2.3.3 Antenna External Noise Sources

The Cosmic Microwave Background (CMB) and the atmosphere contribute the minimum external incident noise temperature contributions to the DSN antennas. The effects of the ground environment can be reduced by antenna design.

As discussed in Section 2.2.3, the DSN treats the cosmic noise contribution as frequency independent consistent with the Rayleigh-Jeans approximation for the range of frequencies in use. The cosmic microwave foreground (CMF) temperature as discussed in Section 2.1, is small compared to the CMB and is currently not included or considered as a DSN antenna noise contributor.

Fig. 2-8. Goldstone 70-m antenna XTR cone X-band feedhorn aperture reference location 1.

Table 2-3. Noise Temperature Equations Summary.

Items	Equations	Text Eqns.
	General:	
(1)	$M = G/T$	2.5-1
(2)	$T_{\text{op}} = T_i + T_e$	2.2-4
(3)	$T_{\text{op1}} = LT_{\text{op2}}$	2.2-20
(4)	$T_{i1} = LT_{i2} - (L-1)T_p$	2.2-25
(5)	$T_{i2} = T_{i1}/L + (1-1/L)T_p$	2.2-26
(6)	$T_{e1} = LT_{e2} + (L-1)T_p$	2.2-23
(7)	$T_{e2} = T_{e1}/L - (1-1/L)T_p$	2.2-24
(8)	$L = (T_p + T_{e1})/(T_p + T_{e2})$	2.6-10
	DSN Application:	
(9)	$T_{\text{op1}} = L_{\text{feed}}T_{\text{op2}}$	2.2-20
(10)	$T_{i1} = L_{\text{feed}}T_{i2} - T_{\text{feed1}} = T_{\text{sky1}} + T_{\text{ant1}} + T_{\text{dichroic1}}$	2.2-25
(11)	$T_{i2} = T_{i1}/L_{\text{feed}} + T_{\text{feed2}} = T_{\text{sky2}} + T_{\text{ant2}} + T_{\text{dichroic2}} + T_{\text{feed2}}$	2.2-26
(12)	$T_{e1} = L_{\text{feed}}T_{e2} + T_{\text{feed1}} = T_{\text{LNA1}} + T_{\text{feed1}} + T_{f1}$	2.5-3
(13)	$T_{e2} = T_{e1}/L_{\text{feed}} - T_{\text{feed2}} = T_{\text{LNA2}} + T_{f2}$	2.5-4
(14)	$T_{\text{feed1}} = (L_{\text{feed}} - 1)T_p$	2.5-18
(15)	$T_{\text{feed2}} = (1-1/L_{\text{feed}})T_p$	2.5-9
(16)	$T_{\text{sky1}} = L_{\text{feed}}T_{\text{sky2}}$	*
(17)	$T_{\text{ant1}} = L_{\text{feed}}T_{\text{ant2}}$	*
(18)	$T_{\text{feed1}} = L_{\text{feed}}T_{\text{feed2}}$	2.5-8
(19)	$T_{\text{LNA1}} = L_{\text{feed}}T_{\text{LNA2}}$	2.5-6
(20)	$T_{f1} = L_{\text{feed}}T_{f2}$	2.5-7
(21)	$L_{\text{feed}} = (T_p + T_{e1})/(T_p + T_{e2})$	2.5-16
(22)	$T_e = (T_h - Y_{ch}T_c)/(Y_{ch} - 1)$	2.2-7
(23)	$T_f = (T_h + T_e)/Y_{oo} - T_{\text{cryo}}/G_{\text{LNA}}$	2.6-8
(24)	$T_f = [T_h + T_{LNA} - (Y_{oo}T_{\text{cryo}}/G_{\text{LNA}})]/(Y_{oo} - 1)$	2.6-9
(25)	$T_{\text{op}} = (T_h + T_e)/Y_{ah}$	2.2-8

*Not specifically referenced in text

2.3.3.1 Atmosphere. T_{sky} is the combined measured incident noise temperature of the atmosphere and the CMB as defined at the feedhorn aperture (reference location 1), for each of the multiple receiving systems on each DSN antenna. Sky noise temperature measurements are important for noise temperature calibrations and can be measured with a water vapor radiometer (WVR) as discussed in Chapter 6, a surface weather model [24, module 105] or antenna "tipping" measurements.

The following describes an antenna tipping analysis for T_{sky}. Since the CMB noise is incident on and attenuated by the atmosphere, the sky noise temperature incident on the antenna is expressed by

$$T_{\text{sky}} = T_{\text{atm}} + \frac{T_{\text{CMB}}}{L_{\text{atm}}} \tag{2.3-5}$$

where

T_{sky} = sky noise temperature due to the attenuated CMB and atmosphere, K

T_{atm} = noise temperature contribution of the atmosphere, incident on the antenna, K

L_{atm} = loss due to the atmosphere, ratio

T_{CMB} = 2.725, K

The atmospheric noise temperature contribution due to its loss is given by

$$T_{\text{atm}} = \left(1 - \frac{1}{L_{\text{atm}}}\right) T_{p\text{atm}} \tag{2.3-6}$$

where

$T_{p\text{atm}}$ = physical temperature of an equivalent uniform atmosphere, K

Substituting T_{atm} into Eq. (2.3-5)

$$T_{\text{sky}} = \frac{T_{\text{CMB}}}{L_{\text{atm}}} + \left(1 - \frac{1}{L_{\text{atm}}}\right) T_{p\text{atm}} \tag{2.3-7}$$

Accurate system noise temperature calibration measurements require "clear weather." Clear weather, is defined as having a cumulative distribution (CD) of 0.25 or less; in this chapter a CD of 0.25 is used. The equation, $T_{p\text{atm}} = 255 + 25$ CD [24, (105B, p. 10, Eq. (1))], provides an estimate of

261.25 K during clear weather. This value for T_{patm} combined with T_{CMB} and measurement of L_{atm} provides an estimate of T_{sky} with Eq. (2.3-7).

The system noise temperature, T_{op} defined at the feedhorn aperture reference location (Fig. 2-8) for a ground antenna and receiving system is given by

$$T_{op} = T_{sky} + T_{AMW} \tag{2.3-8}$$

where

$T_{AMW} = T_{ant} + T_{feed} + T_{LNA} + T_f$ = antenna microwave system noise temperature, K.

Accounting for the difference in the antenna noise temperature change with elevation angle using $\Delta T_{ant} = \Delta T_{AMW}$ (Fig. 2-9), the difference in the sky noise temperature, ΔT_{sky}, between elevation angles 1 and 2 is given by

$$\Delta T_{sky} = \Delta T_{op} - \Delta T_{ant} = \left(T_{patm} - T_{CMB}\right)\left(\left(\frac{1}{L_{atm1}}\right) - \left(\frac{1}{L_{atm2}}\right)\right) \tag{2.3-9}$$

where

$T_{op1,2}$ = system noise temperature at antenna elevation angles 1 or 2, K

$L_{atm1,2}$ = atmosphere loss at elevation angles 1 or 2, ratio

$T_{ant1,2}$ = antenna noise temperature at elevation angles 1 or 2, K

$\Delta T_{op} = T_{op2} - T_{op1}$, K

$\Delta T_{ant} = T_{ant2} - T_{ant1}$, K

The propagation attenuation through the atmosphere at elevation angle EL [26 (p. 22)] for a homogeneous atmosphere, in terms of the zenith atmospheric attenuation, A_{atmZ}, using a flat-Earth model, is given (in dB) by

$$A_{atm} = \frac{A_{atmZ}}{\sin \mathrm{EL}} \tag{2.3-10}$$

For elevation angles 1 and 2

$$A_{atm1} = \frac{A_{atmZ}}{\sin \mathrm{EL}_1} \tag{2.3-11}$$

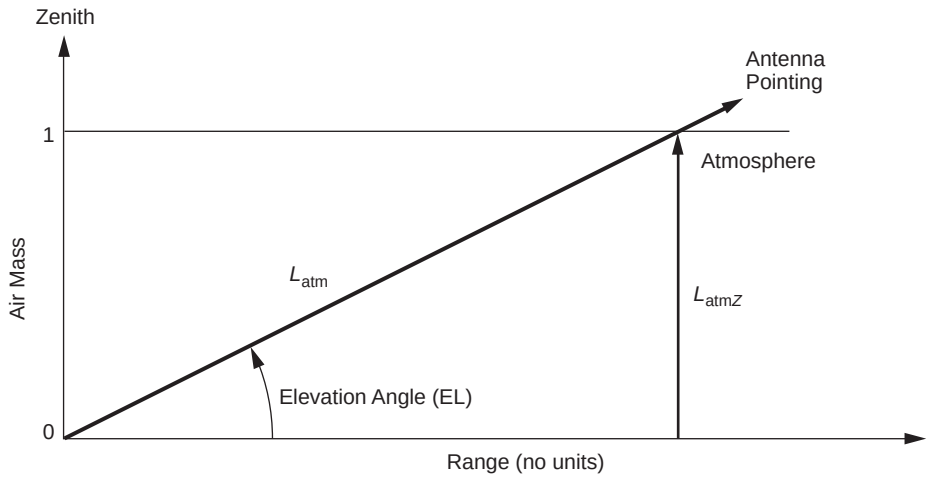


Fig. 2-9. Representation of ground antenna pointing through the atmosphere at elevation pointing angle, EL.

$$A_{\text{atm2}} = \frac{A_{\text{atmZ}}}{\sin \text{EL}_2} \tag{2.3-12}$$

Combining with Eq. (2.3-9) and using $L = 10^{(A/10)}$ results in

$$\Delta T_{\text{sky}} = \Delta T_{\text{op}} - \Delta T_{ant} = (T_{p\text{atm}} - T_{\text{CMB}}) \left(10^{\left(\frac{-A_{\text{atmZ}}}{10 \sin \text{EL}_1} \right)} - 10^{\left(\frac{-A_{\text{atmZ}}}{10 \sin \text{EL}_2} \right)} \right) \tag{2.3-13}$$

A solution can be obtained with elevation angles $\text{EL}_1 = 90$ deg (zenith) using $L_{\text{atm}} = L_{\text{atmZ}}$ for 1 air mass (AM) and $\text{EL}_2 = 30$ deg using $L_{\text{atm}} = \left(L_{\text{atmZ}}\right)^2$ for 2 air masses

$$\Delta T_{\text{sky}} = \Delta T_{\text{op}} - \Delta T_{\text{ant}} = \left(T_{p\text{atm}} - T_{\text{CMB}}\right) \left(\left(\frac{1}{L_{\text{atmZ}}} \right) - \left(\frac{1}{L_{\text{atmZ}}} \right)^2 \right) \tag{2.3-14}$$

Solving this quadratic equation for $1 / L_{\text{atmZ}}$ and inverting

$$L_{atmZ} = \frac{2}{1 + \sqrt{1 - 4Q}} \tag{2.3-15}$$

where

$Q = \left(\Delta T_{\text{op}} - \Delta T_{\text{ant}}\right) / \left(T_{p\text{atm}} - T_{\text{CMB}}\right)$, ratio

$T_{\text{op1}} = T_{\text{op90}} = T_{\text{op}}$ at 90-deg elevation angle, K

$T_{\text{op2}} = T_{\text{op30}} = T_{\text{op}}$ at 30-deg elevation angle, K

$T_{\text{ant1}} = T_{\text{ant90}} = T_{\text{ant}}$ at 90-deg elevation angle, K

$T_{\text{ant2}} = T_{\text{ant30}} = T_{\text{ant}}$ at 30-deg elevation angle, K

ΔT_{ant} needed for Q in Eq. (2.3-15) for the DSS 13 Goldstone research 34-m antenna operating at X-band (8.425 GHz) obtained from a published analysis [43, Table IV], ignoring both the atmosphere and the "bypass" noise contributions removed from this antenna, has a value of 0.89 K. From Eq. (2.3-9), $\Delta T_{\text{ant}} = \Delta T_{\text{op}} - \Delta T_{\text{sky}}$ is used to calibrate ΔT_{ant}. ΔT_{op} was measured with the DSS 13 antenna on 2005-320 to be 2.432 K (17 measurements between 90 and 30 deg elevation angle), and ΔT_{sky} was measured with the WVR to be 2.217 K during this same period. These values result in $\Delta T_{\text{ant}} = 0.215$ K. Although this is in disagreement with the analysis result of 0.89 K, the measurement value of 0.215 K is used in the following.

The quadratic solution, Eq. (2.3-15) is used with these measured ΔT_{op} and ΔT_{ant} data and with $T_{p\text{atm}} = 261.25$ K and $T_{\text{CMB}} = 2.725$ K giving averaged values of $A_{\text{atmZ}} = 0.0377$ dB and $T_{\text{skyZ}} = 4.961$ K. The measurement resolution (scatter) for A_{atmZ} has a standard deviation of 0.0015 dB for the 17 individual measurements. The calibrated value $\Delta T_{\text{ant}} = 0.215$ K could be used with this same antenna and Eq. (2.3-15) to calibrate the atmosphere at X-band on other days as needed independent of the WVR. Further analysis and measurements are needed for agreement between the measurements and the analysis for ΔT_{ant}.

For future improved atmospheric loss calibrations without the WVR, it seems worthwhile to consider a tipping radiometer system using a horn antenna designed for low sidelobes so that the very small value for ΔT_{ant} compared to the atmosphere can be neglected for the atmospheric calibration using the above quadratic solution for 30- and 90-deg elevation angle measurements.

Figure 2-10 shows the DSS-13 R&D Venus research station's antenna T_{op} and T_{AMW} values as a function of antenna (a) elevation angle and (b) air mass for clear weather conditions for 2005-347. A 0.0387-dB zenith atmospheric loss for this calculation obtained from the WVR at the same operating frequency results in a very small value for ΔT_{ant}. This is indicated by the value for

T_{AMW}, which is nearly independent of elevation angle. T_{AMW} data as a function of antenna elevation angle will be available for all the DSN antennas in future revisions of the DSN 810-005 document [24].

T_{AMW} is an important parameter describing the DSN receiving systems ground performance independent of weather. T_{AMW} is measured for each DSN antenna at each operational frequency and documented in the DSN 810-005 document [24]. T_{AMW} combined with T_{sky}, Eq. (2.3-8), provides the operational system noise temperature as a function of actual or statistical weather data and for any given antenna elevation angle.

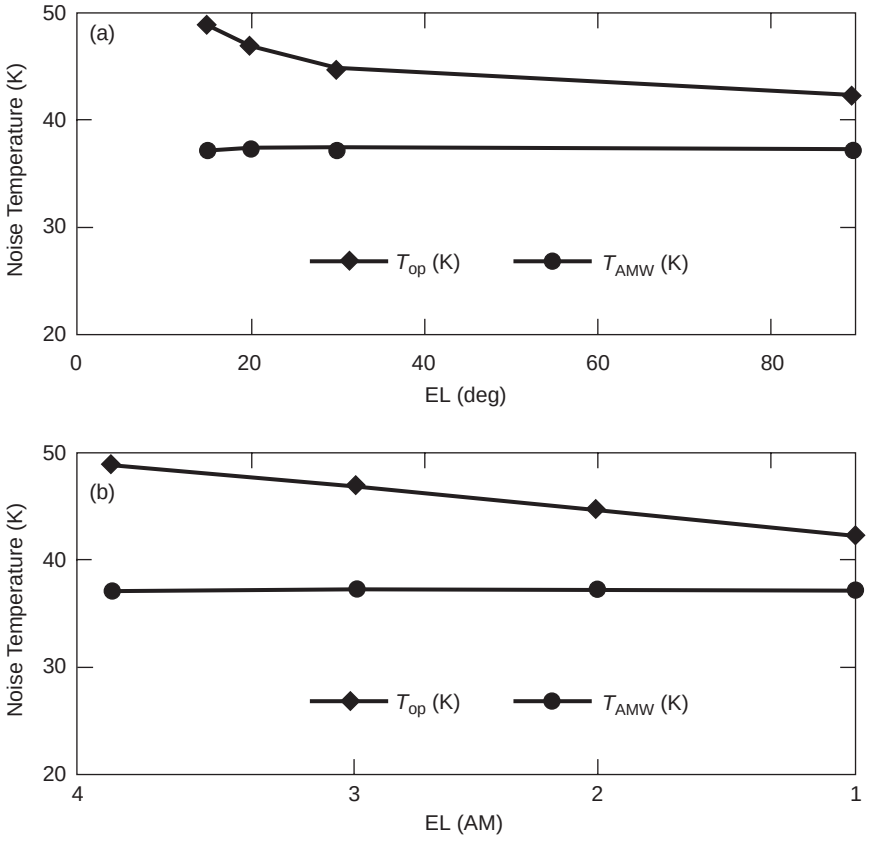


Fig. 2-10. Goldstone DSS 13 34-m BWG antenna X-band system noise temperature T_{op} and T_{AMW} versus antenna elevation angle, 2005-347 doy as functions of (a) elevation angle and (b) air mass.

2.3.3.2 External Noise Sources

Solar noise is the strongest source of observable deep space external noise [24 (module 105 Rev. B, p. 15)]. This noise source should be taken into account when tracking within a few degrees of the Sun. (It is usua lly considered negligible for DSN antennas at Sun–Earth–Probe (SEP) angles greater than 4 deg.) Figure 2-11 shows the Goldstone High Efficiency (HEF) 34-m antenna X-band system noise temperature increases due to t he Sun at various offset angles. This shows the largest increases occurring perpendicular to the antenna quadripod subreflector support structure [5 (module 105 Rev. B, Fig. 12, p. 33)]. An early report shows predicted and measured receiving system noise performance degradations when tracking spacecraft close to the solar limb [27] at S-band (2.297 GHz) with the 64-m Goldstone antenna prior to the upgrade to 70- m diameter. The phase of the 11-year solar cycle is important. The X-band brightness temperature for an active Sun can be two to four times as high as the quiet Sun.

The **Moon** appears as an approximately 240-K blackbody disk w ith an apparent diameter of about 0.5 deg, similar to the Sun as both seen from Earth. Due to the similar apparent diameters, the solar curves [24 (module 105,

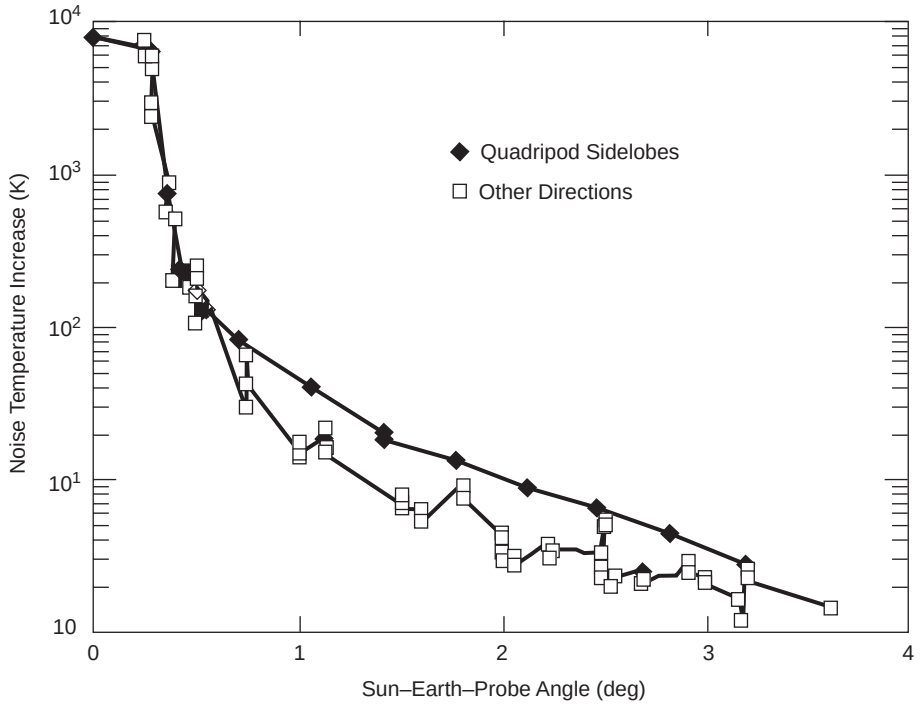


Fig. 2-11. Goldstone High Efficiency (HEF) antenna X-band system noise temperature increases due to the Sun at various offset angles, showing larger increases perpendicular to the antenna quadripod directions.

Rev. B, p. 18)] may be used for lunar calculations, with the noise temperature values scaled by their 'on point' values. The clear-sky system noise is included and must be subtracted out before scaling. At antenna pointing offset angles greater than 2 deg, the lunar noise temperature contribution is negligible.

The **Planets** increase the noise temperature within the antenna beam according to the formula [24 (105, Rev. B, p. 19)]

$$T_{\text{pl}} = \left(\frac{T_k G d^2}{16 R^2} \right) e^{-2.77\left(\frac{\theta}{\theta_o} \right)^2} \tag{2.3-16}$$

where

T_{pl} = increased system noise temperature due to a planet in the antenna beam, K

T_k = blackbody disk temperature of the planet, K

d / R = planet diameter, d, relative to the planet distance, R, ratio

θ / θ_o = angular distance from planet center to antenna beam center, θ, relative to the antenna full half-power beamwidth (HPBW), θ_o, ratio

G = antenna gain, corrected for atmospheric attenuation, ratio

The above calculation of a planet noise temperature contribution to a receiving system noise temperature assumes that the radiating source is small compared to the antenna beamwidth. The constants needed for this calculation are provided in Table 2-4, obtained from the DSN 810-005 document [24 (module 105 Rev. B, Table 20)]. This updates the Venus blackbody disk temperatures at X-and Ka-bands according to the de Pater publication [28]. A graphical approach which is valid for source sizes either less than or greater than the antenna beamwidth developed by Kantak and Slobin [29] as an alternative to Eq. 2.3-17 is also described in module 105, Rev. B [24].

Galactic noise is a frequency-dependent noise contributor from our galaxy (the Milky Way) ranging from a maximum of about 10,000 K at 100 MHz to less than 1 K at 1 GHz [24 (module 105 Rev. B, p. 21)]. This noise source should be taken into account for DSN S-band and X-band frequencies when the antenna beam is pointed toward the Galactic center; it is usually ignored at Ka-band (32 GHz).

Atmospherics and Radio Frequency Interference (RFI) can be disruptive and are not usually predictable. Atmospherics and intermittent RFI cause non-repeatability in system noise temperature measurements. Averaging the results and discarding 'outlier' data points can reduce the effect. A stable RFI signal can be detected and characterized within the receiving system

Table 2-4. Planetary parameters and X-band /Ka-band noise temperatures at mean minimum distance from Earth.

	Mercury	Venus	Earth	Mars	Jupiter	Saturn	Uranus	Neptune	Pluto
Equatorial Diameter (km)	4880	12,104	N/A	6794	142,984	120,536	51,118	49,532	2274
Mean Distance from Earth (10^6 km):									
Min.	91.7	41.4	N/A	78.3	628.7	1279.8	2721.4	4354.4	5763.9
Max.	207.5	257.8	N/A	377.5	927.9	1579.0	3020.6	4653.6	6063.1
Mean Distance from Sun:									
(10^6 km)	57.9	108.2	149.6	227.9	778.3	1429.4	2871.0	4504.0	5913.5
AU	0.387	0.723	1.000	1.523	5.203	9.555	19.191	30.107	39.529
Blackbody Disk Temp (K)	625	625 (X-band) 415 (Ka-band)	250–300[1]	180	152	155	160	160	160
T_{Planet} (K) at Mean Minimum Distance:									
X-Band:									
70-m (74.4 dBi gain)	3.05	91.96	N/A	2.33	13.53	2.37	0.10	0.04	0.00
34-m (68.3 dBi gain)	0.75	22.57	N/A	0.57	3.32	0.58	0.02	0.01	0.00
Ka-Band:									
34-m (78.8 dBi gain)	8.39	253.29	N/A	6.43	37.27	6.52	0.27	0.10	0.00

[1] Ocean (250 K) and Land (300 K).

Table taken from 810-005, Module 105, Rev. B, Change 12, May 26, 2006, Table 20 [24].

bandwidth using a spectrum analyzer; the effect may be reduced and or eliminated by c hanging filters and the operating frequency. This source of interference might not cause erratic measurements but can cause a biased result.

2.4 Low-Noise Amplifiers

2.4.1 Receiver Effective Noise Temperature

The antenna feedhorn, microwave front-end components, low-noise amplifiers (LNAs), and follow-on amplifiers generate internal noise temperature contributions to the front-end-assembly receiving system. These internal contributions are determined and translated according to the equations in Section 2.2 and applied to the receiving system in Section 2.5.

The receiver effective input noise temperature defined at the LNA cryogenic package input is given by (Sections 2.2 and 2.6).

$$T_e = T_{\mathrm{LNA}} + T_f, \tag{2.4-1}$$

where

T_e = receiver effective input noise temperature, K

T_{LNA} = LNA noise temperature, K

T_f = LNA follow-up amplifier noise temperature, K

Maser amplifiers, discussed in C hapter 3, provide the lowest amplifier noise temperatures and have been widely used in the DSN since its inc eption. HEMT amplifiers, discussed in Chapter 5, are a more recent innovation, and have nearly the low noise performance of masers. They are now also widely used in the DSN. A photograph of the X-band HEMT amplifier used with the Goldstone 70-m antenna XTR feedcone is shown in Fig. 2-12 [30].

2.4.2 Noise Temperature of Cascaded Amplifiers

The effective input noise temperature for a string of cascaded amplifiers is mostly determined by the front-end amplifiers. A receiver consisting of cascaded amplifiers, each with a separate gain and effective input noise temperature has an overall effective input noi se temperature given by [4 (p. 22)].

For cascaded amplifiers

$$T_e = T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1 G_2} + \cdots + \frac{T_{\mathrm{en}}}{G_1 G_2 \ldots G_{n-1}} \tag{2.4-2}$$

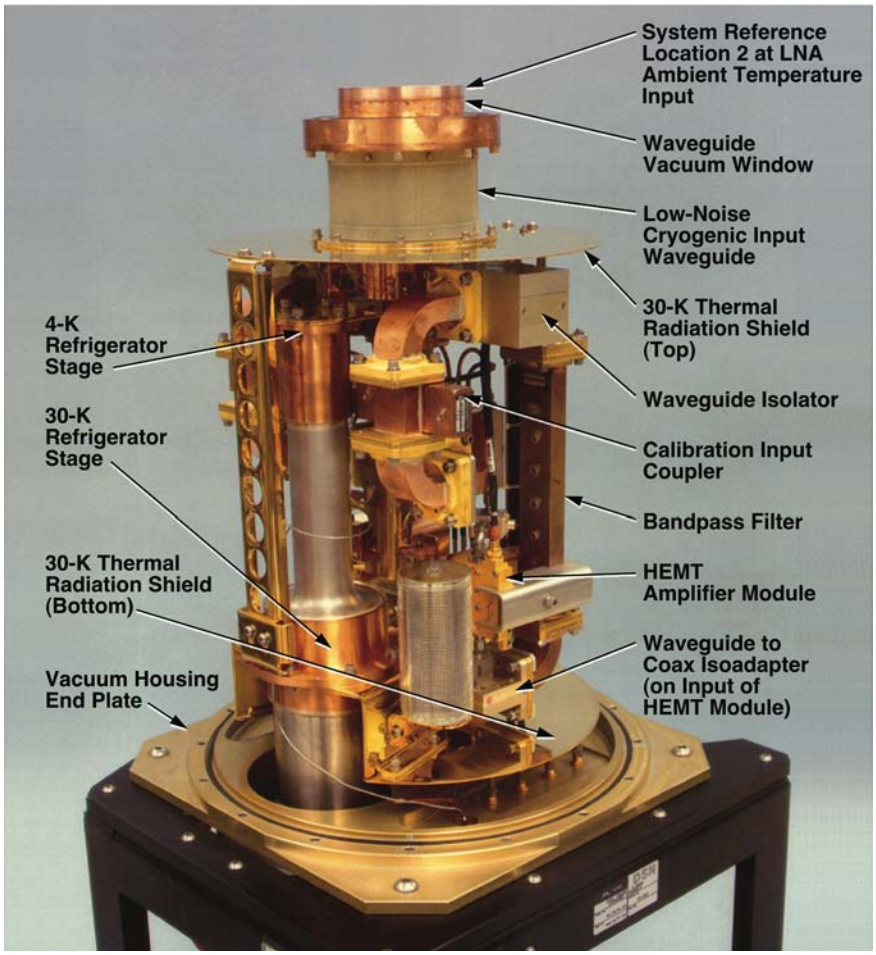


Fig. 2-12. Goldstone 70-m antenna XTR cone X-band HEMT amplifier.

where

T_e = effective input noise temperature of cascaded amplifiers, K

T_{e1} = effective input noise temperature of amplifier 1 at its input, K

T_{e2} = effective input noise temperature of amplifier 2 at its input, K

T_{e3} = effective input noise temperature of amplifier 3 at its input, K

T_{en} = effective input noise temperature of amplifier n at its input, K

G_1 = gain of amplifier 1, ratio

G_2 = gain of amplifier 2, ratio

G_n = gain of amplifier n, ratio

For DSN receiving systems, T_e can be determined with just the first two amplifiers

$$T_e = T_{\text{LNA}} + \frac{T_F}{G_{\text{LNA}}} \tag{2.4-3}$$

where

T_F = effective input noise temperature of LNA "follow-on" or "post" amplifier defined at its input location, K

The receiver noise temperature, T_e, is dominated by the LNA noise temperature, since the DSN LNAs have high gain, relative to the follow-on amplifiers and is generally optimized, minimizing both T_{LNA} and T_f for a well designed system

$$T_f = \frac{T_F}{G_{\text{LNA}}} << T_{\text{LNA}} \tag{2.4-4}$$

where

T_f = effective input noise temperature of the post amplifier defined at the LNA input, K

Good performance (low value for T_e) requires both low noise and high gain for the LNA. The "*Y*-factor" method for measuring T_f is analyzed in Section 2.6 with the results given in Eqs. (2.6-8 and 2.6-9).

2.5 Receiving Systems

2.5.1 Receiving System Figure of Merit

The maximum data rate capability for a communications system [31,32 (p. 2-1)] is proportional to the receiving system antenna gain, which itself is proportional to collecting area, and inversely proportional to the receiving system noise temperature. The receiving system figure of merit, *M*, is defined as

$$M = \frac{G}{T} \tag{2.5-1}$$

where

M = receiving system figure-of-merit, ratio

M, dB = 10 log (M)

G = antenna gain relative to isotropic radiator, ratio

$T = T_{\text{op}}$ = system operating noise temperature relative to 1 K, ratio

For a linear receiving system consisting of components with gain and loss, the figure of merit, M, does not change as, a function of the reference location within the receiving system. Moving the reference location within a linear receiving system changes the system gain and noise temperature equally across each component, so that G/T remains constant. Improvement in the receiving system figure of merit requires an increase of the antenna gain (Section 2.3) and/or a reduction in the system noise temperature, T_{op}. Lowering T is in many cases a cost-effective method of increasing G/T. Obtaining either lower system noise temperature or increased antenna area is expensive. System noise-temperature performance is valuable and important for achieving maximum data rates; and therefore, it deserves to be accurately determined and maintained.

2.5.2 Receiving System Operational Noise Temperature

System operating noise temperature is defined as the sum of the external input and internal receiver effective input noise temperatures,

$$T_{\text{op}} = T_i + T_e \tag{2.5-2}$$

where (all terms defined at the same system reference location)

T_{op} = system operating noise temperature, K

T_i = input noise temperature, K

T_e = receiver effective input noise temperature, K

The sum of the input and receiver effective noise temperatures includes contributions from the CMB, atmosphere, ground, antenna, antenna feedhorn, microwave front-end assembly, LNA, and follow-on amplifier. Some of these contributions, such as the LNA, are determined separately from the assembled system. Adding these individual noise temperature contributions requires that each be referenced to the same location.

For DSN bookkeeping it is convenient to use the components T_{CMB}, T_{atm}, T_{ant}, and T_{dichroic} for T_i and T_{LNA}, T_{feed}, and T_f for T_e, at both the feedhorn aperture reference location 1 and the LNA input reference location 2, separated by the feed loss, L_{feed}. For the receiving system noise temperature reference locations

$$T_{e1} = T_{\text{LNA1}} + T_{\text{feed1}} + T_{f1} \tag{2.5-3}$$

and

$$T_{e2} = T_{\text{LNA2}} + T_{f2} \tag{2.5-4}$$

where the noise temperature contribution due to the feed loss at reference location 1 is given by

$$T_{\text{feed1}} = (L_{\text{feed}} - 1)T_p \tag{2.5-5}$$

T_{e2} has no T_{feed} noise temperature contribution since the feed is located "upstream" from the receiver LNA input, reference location 2. The T_{feed} noise contribution becomes part of T_{i2} at the LNA input. With T_{feed} defined separately, the components of T_e are each related by the feed loss, L_{feed}

$$T_{\text{LNA1}} = L_{\text{feed}} T_{\text{LNA2}} \tag{2.5-6}$$

$$T_{f1} = L_{\text{feed}} T_{f2} \tag{2.5-7}$$

and

$$T_{\text{feed1}} = L_{\text{feed}} T_{\text{feed2}} \tag{2.5-8}$$

where the noise temperature contribution due to the feed loss at reference location 2 is given by

$$T_{\text{feed2}} = \left(1 - \frac{1}{L_{\text{feed}}}\right) T_p \tag{2.5-9}$$

When these DSN terms are manipulated and combined for the resultant T_{e1} and T_{e2} values, Eqs. (2.2-23 through 2.2-26) provide the correct translation relationships between the reference 1 and 2 locations. The translation equations are summarized in Table 2-3.

2.5.2.1 Calibration Example. The X-band Transmit and Receive (XTR) feedcone assemblies installed in the DSN 70-m antennas were calibrated beginning in 2000. This section provides a description of the calibration, including details of the noise temperature measurements.

The feeds are installed in the three equipment feedcone assemblies mounted on the dish surface, in a circle, near the center of the dish. From a given feed,

the signal proceeds by waveguide to the front-end equipment located in the designated feedcone assembly. In response to the NASA Cassini mission request for high power X-band uplink, the X-band receive operational (XRO) feedcone assembly containing the X-band maser LNAs was replaced with the X-band transmit receive (XTR) feedcone assembly on the three DSN 70-m antennas. This was accomplished with the addition of a transmit junction to accommodate the X-band uplink capability. A retractable S-/X-band dichroic is installed just above the feedhorn. With the dichroic plate retracted, the system is capable of receiving or transmitting X-band independently or simultaneously. With the dichroic plate extended, the system is capable of

1) Receive or transmit S-band independently or simultaneously
2) Receive or transmit X-band independently or simultaneously
3) Receive S-and X-band simultaneously while transmitting either S-band or X-band

When the XRO feedcone assembly was replaced by the XTR feedcone assembly, an additional loss was incorporated from the transmit junction. Furthermore the maser LNAs in the XRO feedcone assembly were replaced by the high electron mobility transistor (HEMT) LNAs in the XTR feedcone assembly. HEMTs have higher noise temperatures than the maser LNAs. However, the overall receiving system noise temperature with the XTR feedcone assembly was lower than that with the XRO feedcone assembly. This was achieved with development of a diplexed feed [33 (X-band portion)] where the transmit junction is combined with the feedhorn design and by cooling selected waveguide "front-end" components in the low-temperature cryogenic package. A photograph of the X-band HEMT LNA is shown in Fig. 2-12.

A noise temperature comparison of the XRO and XTR feedcone assemblies system noise temperatures mounted on the Goldstone 70-m antenna is shown in Fig. 2-13 for X-band, zenith pointing, and clear weather. The noise temperature values shown have been updated since the original publication [30], consistent with the DSN no longer using Planck's radiation law noise power reduction at higher frequency as discussed in Section 2.2.3. Therefore, the CMB noise temperature has been changed from an earlier publication value of 2.5 K for X-band to 2. 725 K for this updated figure. A sky brightness temperature (combined CMB and atmosphere noise contribution) of 4.8 K is used as compared with the original 4.6 K [30] value. The 0.2 K increase used for the CMB value and a similar decrease of 0.2 K in the receiver system noise temperatures are consistent with the Rayleigh-Jeans approximation. These changes cancel so that the final total system noise temperatures for both feedcones are unchanged.

The noise temperature calibration for DSN front-end assemblies typically uses an ambient calibration load first extended over and then retracted from the

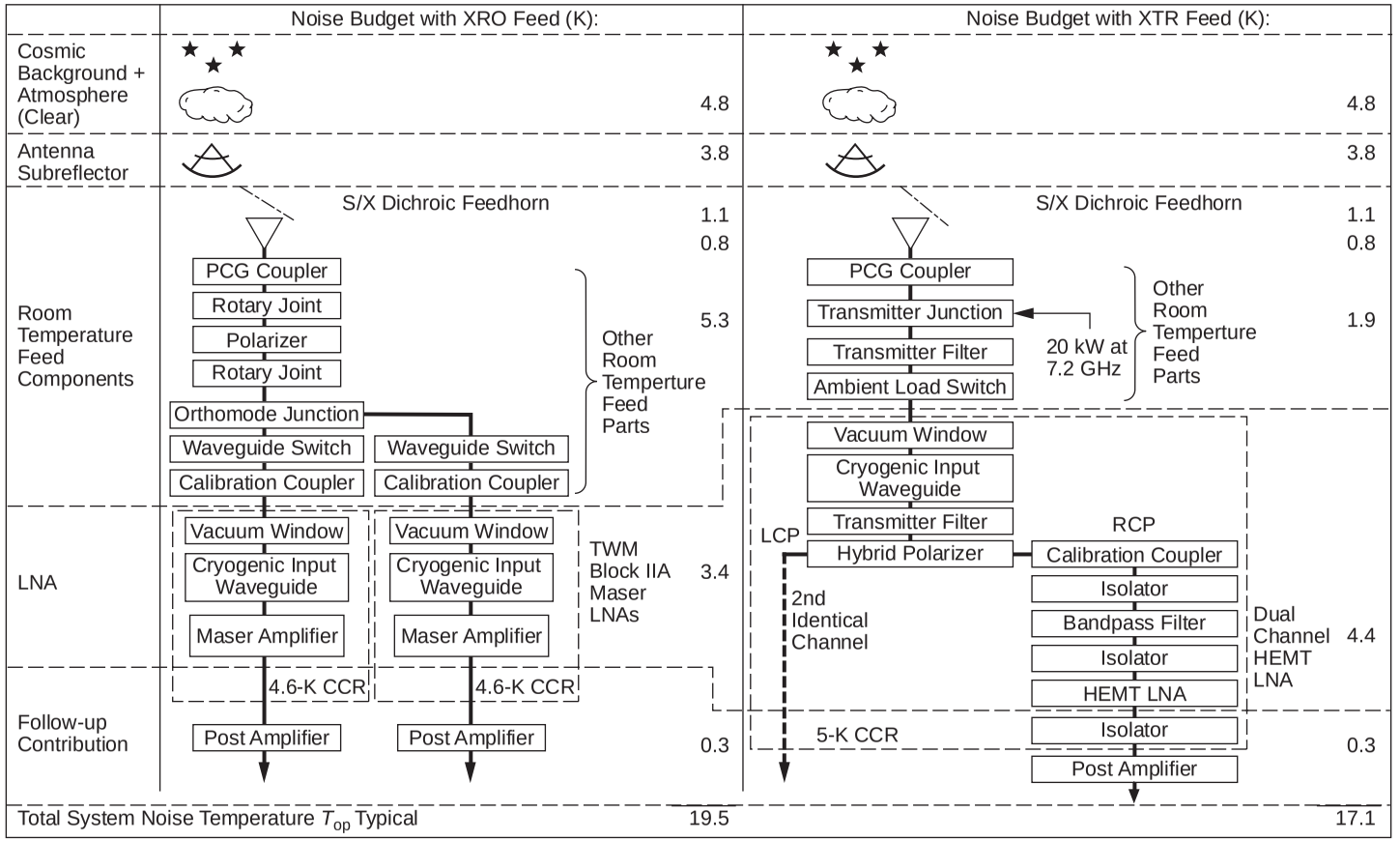


Fig. 2-13. Goldstone 70-m antenna new XTR diplexed feedcone noise temperature performance defined at feedhorn aperture reference location 1 compared with replaced XRO feedcone at X-band, zenith and clear weather.

feedhorn aperture input for BWG antennas, or a waveguide switch ambient load for non-BWG antennas as shown in Fig. 2-14 (a) and (b). The sensor for the physical temperature readout is located in the ambient load, in close proximity to the feed assembly. For this reason, the noise temperature calibrations use this same physical temperature, T_p, for both the calibration load and the feed assembly. With the waveguide losses and the ambient calibration load at the same physical temperature, there is no difference for noise measurements between locating the calibration load at the feedhorn aperture, the waveguide switch, or the LNA input. A complete noise temperature calibration sequence for a DSN antenna front-end assembly comprises three separate measurements, a, b, and c. This calibration sequence is required for each newly implemented microwave assembly. Some of these calibrations are also required following maintenance, upgrade, and/or repair. These measurements are performed pointing toward zenith on a clear day, and each requires T_{sky} input (Section 2.6). The three calibration categories are:

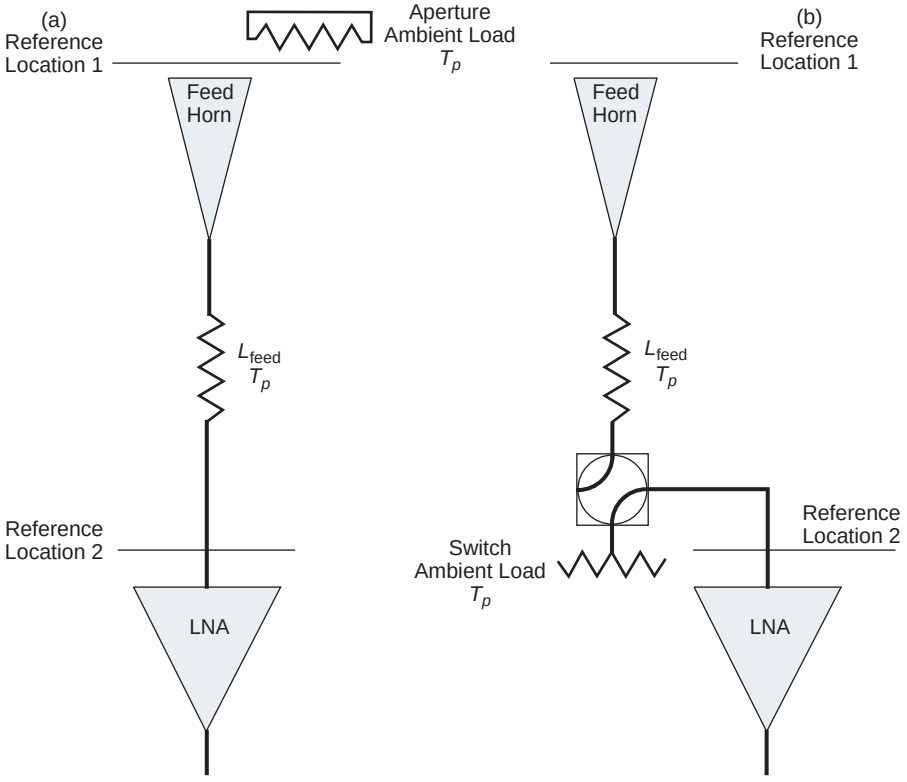


Fig. 2-14. Noise temperature calibration configurations for DSN feed assembly and LNA receiving systems: (a) aperture ambient load BWG antennas and (b) switch ambient load non-BWG antennas.

1) **LNA noise temperature (calibration a):** Determine the noise temperature of the LNA, T_{LNA2} defined at the LNA ambient temperature input, system reference location 2, Fig. 2-15.
2) **Feed assembly loss (calibration b):** Determine the feed assembly loss, L_{feed}, Fig. 2-16.
3) **System noise temperature (calibration c):** With the front-end assembly installed on the antenna, determine the system noise temperature parameter values for T_{op}, T_{AMW}, T_{ant}, and T_{UWV}, all defined at the feedhorn aperture, system reference location 1 shown in Fig. 2-17.

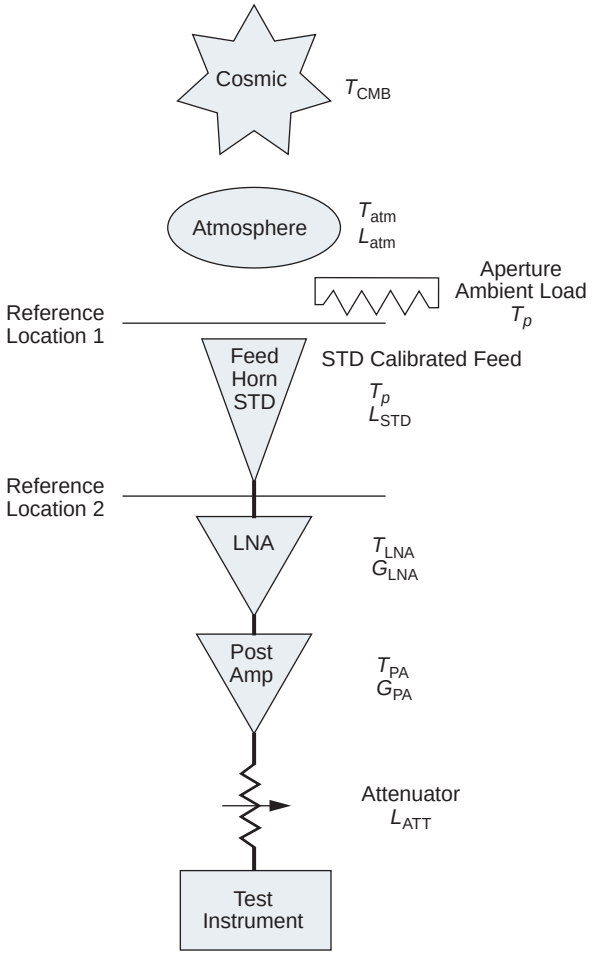


Fig. 2-15. Calibration a: DSN LNA noise temperature calibration configuration.

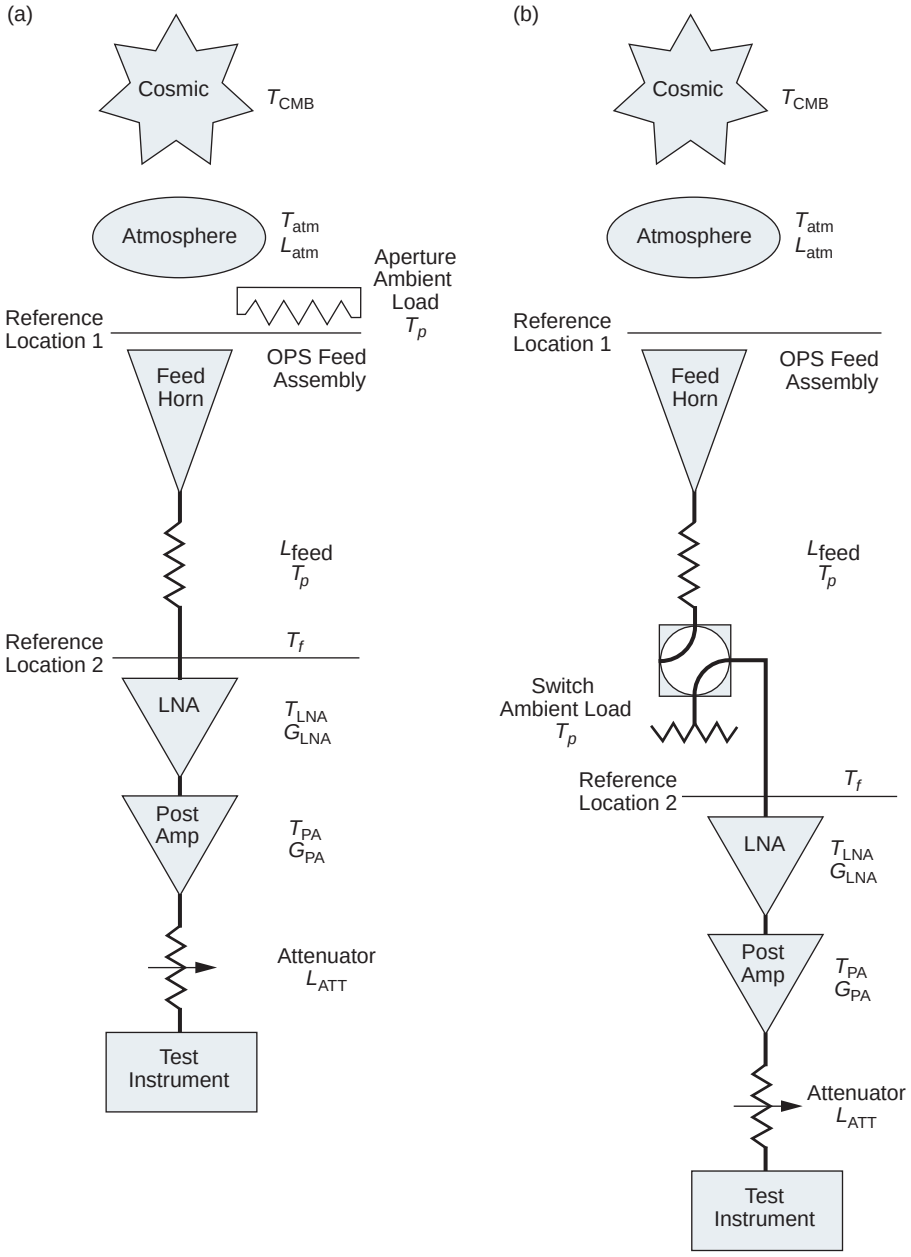


Fig. 2-16. Calibration b: DSN operational feed assembly noise temperature calibration configuration: (a) aperture load and (b) switch ambient load.

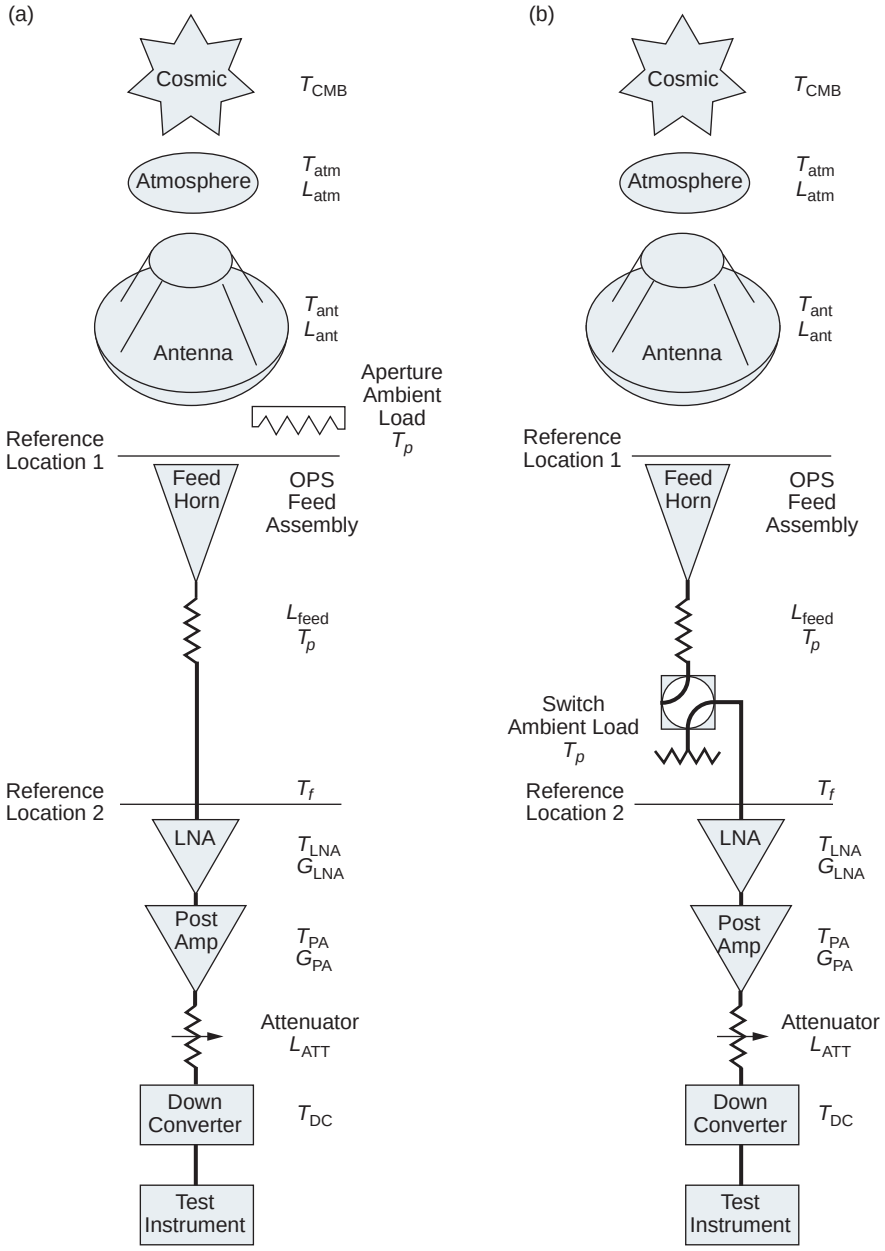


Fig. 2-17. Calibration c: DSN operational feed assembly noise temperature calibration configuration: (a) aperture ambient load and (b) switch ambient load.

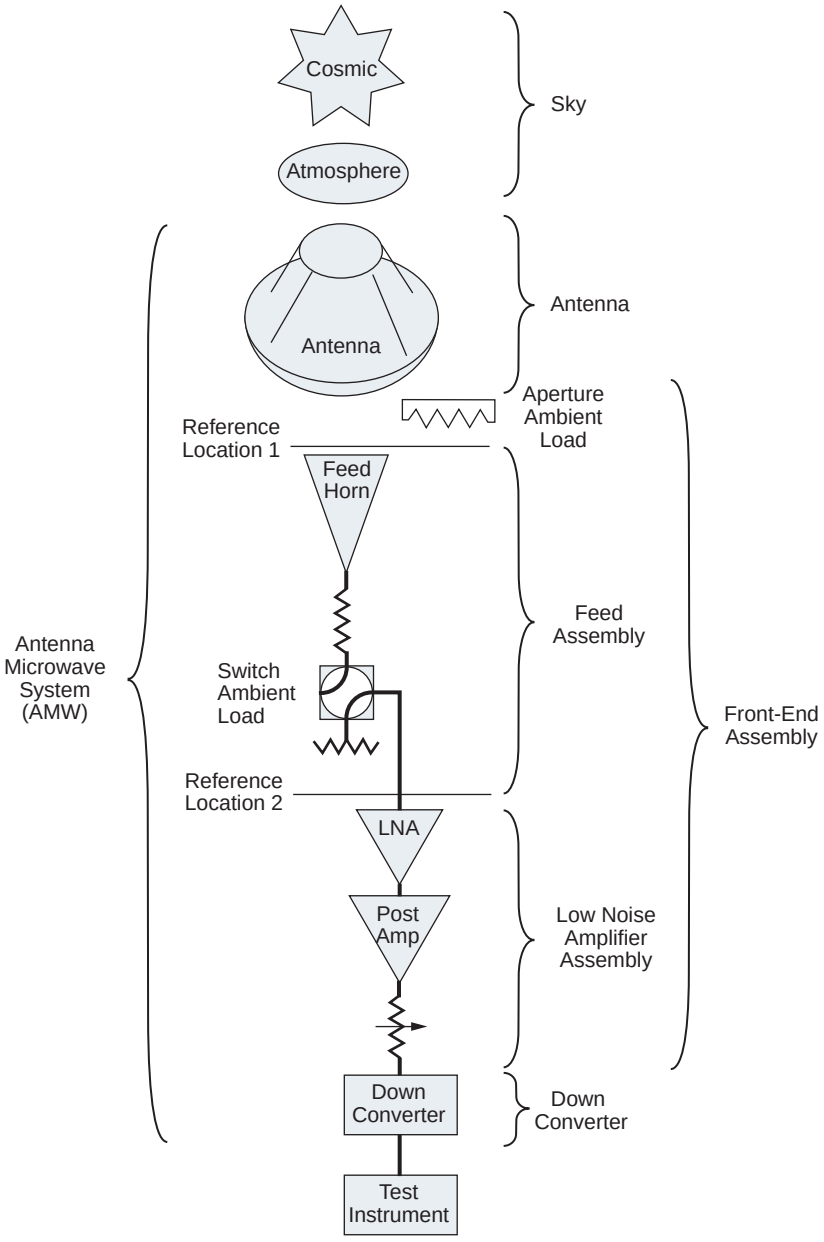


Fig. 2-18. DSN front-end assembly noise temperature calibration configuration components for aperture ambient load or switch ambient load.

2.5.2.2 LNA Noise Temperature (Calibration a). A JPL calibrated standard horn connects directly to the LNA (substituting for the operational feed assembly). The calibrated standard (std) feedhorn viewing the sky pointed to zenith is used as the cold load to measure T_e. The external input noise temperature, T_{i2}, to the LNA is given by

$$T_{i2} = T_{\text{sky2}} + T_{\text{std2}} \tag{2.5-10}$$

Substituting for T_{sky2}

$$T_{i2} = \frac{T_{\text{sky1}}}{L_{\text{std}}} + T_{\text{std2}} = 4.76 + 2.72 = 7.48 \text{ K} \tag{2.5-11}$$

where (values for this example)

T_{i2} = combined sky noise and standard feed loss defined at LNA input reference location 2, K

T_{sky2} = sky noise temperature defined at the LNA input reference location 2, K

$T_{\text{sky1}} = 4.800$ = sky noise temperature due to the combined atmosphere and CMB, defined at the standard (std) feedhorn aperture reference location 1, K

$T_{\text{std2}} = (1 - 1/L_{\text{std}})T_p = 2.7243$ = noise temperature of the calibrated standard feedhorn loss at the LNA input reference 2, K

$L_{\text{std}} = 1.009253\,(0.040 \text{ dB})$ = loss of the calibrated standard feedhorn between the aperture and the LNA input, ratio

$T_p = 297.1500\,(24.0000 \text{ C})$ = physical temperature of the feedhorn loss, K

Using T_{i2} from Eq. (2.5-11) with the *Y*-factor equation, Item 22 of Table 2-3, for T_{e2} (LNA and follow-on amplifier) noise temperature, gives

$$T_{e2} = \frac{T_p - Y_{ah}T_{i2}}{Y_{ah} - 1} = 4.70 \text{ K} \tag{2.5-12}$$

where (values for this example)

T_{e2} = receiving system effective noise temperature at LNA input reference location 2, K

Y_{ah} = 24.7742 (13.9400 dB) = hot (ambient) load and antenna *Y*-factor, ratio

T_p = 297.1500 (24.0000 C) = physical temperature of the hot (ambient) load, K

The follow-on noise temperature contribution (T_f) given by Eq. (2.6-8), with terms defined there and ignoring the smaller second term for this application with Y_{oo} = 977.23722 (29.9000 dB) is given by

$$T_{f2} = \frac{T_p + T_{e2}}{Y_{oo}} = 0.309 \text{ K} \tag{2.5-13}$$

and with Eq. (2.5-12) gives

$$T_{\text{LNA2}} = T_{e2} - T_{f2} = 4.704 - 0.309 = 4.395 \text{ K} \tag{2.5-14}$$

Although the DSN calibrates their LNA noise temperatures with an unobstructed view of the sky using a calibrated standard horn pointed toward zenith in place of a cold load, this calibration could be accomplished in the laboratory using two calibrated loads, a hot (ambient) load and a calibrated cooled load. The application of a standard horn as the cold load for calibrating LNAs has been used in the DSN for a number of years.

2.5.2.3 Feed Assembly Loss (Calibration b). The operational feed assembly is ready for evaluation after calibration and installation of the LNA in the front-end assembly. The receiver noise temperature is calibrated at the feedhorn aperture reference location 1, similar to calibration a, with Eq. (2.5-12). For the XTR feedcone pointed toward zenith

$$T_{e1} = \frac{T_p - Y_{ah} T_{i1}}{Y_{ah} - 1} = 7.497 \text{ K} \tag{2.5-15}$$

where (values for this example)

T_{e1} = receiving system effective input noise temperature defined at the feedhorn aperture reference location 1, K

T_p = 297.1500 (24.0000 C) = physical temperature of the ambient load, K

$T_{i1} = T_{\text{sky1}}$ = 4.8000 = input noise temperature due to the combined atmosphere and CMB, defined at the feedhorn aperture reference location 1, K

$Y_{ah} = P_h / P_a = 24.7738\ (13.93992\ \text{dB}) =$ hot (ambient) load and antenna *Y*-factor, ratio

The XTR front-end feed assembly loss can be calculated using Eq. (2.6-10) with the T_{e1} value from Eq. (2.5-15) = 7.497

$$L_{\text{feed}} = \frac{T_p + T_{e1}}{T_p + T_{e2}} = 1.0092296 \text{ ratio } (= 0.03990 \text{ dB}) \qquad (2.5\text{-}16)$$

This uses $T_{\text{LNA2}} + T_{f2}$ for T_{e2} and the first two numerator terms of Item 24 in Table 2-3 for T_{f2}

$$T_{f2} = \frac{T_p + T_{\text{LNA2}}}{Y_{oo} - 1} = 0.31609 \text{ K} \qquad (2.5\text{-}17)$$

where (for Eqs. 2.5-16 and 2.5-17)

L_{feed} = feed assembly loss, ratio

$T_{e2} = T_{\text{LNA2}} + T_{f2} = 4.395$ (from Eq. (2.5-14)) + 0.31609 (from Eq. (2.5-17)) = 4.711 = receiving system effective input noise temperature defined at the LNA input reference location 2, K

T_{f2} = follow-up amplifier noise temperature contribution defined at the LNA input reference location 2, K

$T_p = 297.1500\ (24.0000\ \text{C}) =$ physical and noise temperature of the ambient load, K

$Y_{oo} = 954.99259\ (29.8000\ \text{dB}) =$ LNA on and off *Y*-factor, ratio

The noise temperature contribution of the feed assembly loss, defined at the feedhorn aperture is

$$T_{\text{feed1}} = (L_{\text{feed}} - 1)T_p = (1.0092296 - 1)\ 297.1500 = 2.743\ \text{K} \qquad (2.5\text{-}18)$$

This completes the front-end assembly evaluation at the ground site, prior to installation on the antenna.

2.5.2.4 System Noise Temperature (Calibration c). With the installation of the front-end assembly on the antenna, the system noise temperature, T_{op}, and the noise contributions from the major components of the system are determined using items 9 and 25 of Table 2-3.

$$T_{\text{op1}} = L_{\text{feed}}\left(\frac{T_p + T_{e2}}{Y_{ah}}\right) = 17.12 \text{ K} \tag{2.5-19}$$

where

$T_{\text{op1}} = T_{\text{op}} =$ system noise temperature, front-end assembly on antenna, defined at the feedhorn aperture reference location 1, K

$L_{\text{feed}} = 1.0092296$ (0.03990 dB) = feed assembly loss, ratio

$T_p = 297.1500$ (24.00 C) = feed and load physical temperature, K

$T_{e2} = T_{\text{LNA2}} + T_{f2} = 4.3950 + 0.2690 = 4.6640 =$ receiving system effective input noi se temperature at LNA input, reference location 2, K. Note that T_f for calibrations a, b, and c are not identical due to different configurations

$Y_{ah} = P_{\text{h}} / P_{\text{a}} = 17.79099$ (12.50200 dB) = hot (ambient) load and antenna *Y*-factor ratio

The primary requirements to be verified for a system in the field is the noise contribution of the Antenna Microwave system, $T_{\text{AMW}}(T_{\text{AMW}} = T_{\text{ant}} + T_{\text{UWV}})$, microwave front end feed assemble (feed and LNA), $T_{\text{UWV}}(= T_{e1})$, and the antenna, T_{ant} all defined at the feedhorn aperture reference location 1. Figure 2-8 shows the feedhorn aperture input of the XTR front end assemble as installed on the Goldstone 70-m antenna. The following analysis determines all these parameters.

The microwave noise contribution, T_{UWV}, using i tems 12 a nd 14 of Table 2-3 and the values given above can be determined by

$$T_{\text{UWV}} = T_{e1} = L_{\text{feed}}T_{e2} + (L_{\text{feed}} - 1)T_p = (1.0092296)(4.66399) + (0.0092296)(297.1500) = 7.45 \text{ K} \tag{2.5-20}$$

where

$T_{\text{UWV}} = T_{e1} =$ microwave receiver effective input noise temperature defined at the feedhorn aperture reference location 1, K

It is important to determine an equivalent system noise temperature T_{AMW} not accounting for the external contributions of T_{sky} for use in the DSN 810-005 mission interface document [24] as required for the operational missions. T_{AMW} is determined for the XTR feedcone using the values given above

$$T_{\text{AMW}} = T_{\text{op1}} - T_{\text{sky1}} = 17.12 - 4.80 = 12.32 \text{ K} \qquad (2.5\text{-}21)$$

where

T_{AMW} = antenna-microwave receiver input noise temperature defined at the feed aperture location 1, K

The zenith antenna noise temperature contribution defined at the feedhorn aperture, system reference location 1, is given by

$$T_{\text{ant1}} = T_{\text{AMW}} - T_{e1} - T_{\text{dichroic1}} = 12.32 - 7.45 - 1.10 = 3.77 \text{ K} \qquad (2.5\text{-}22)$$

For the Goldstone 70-m antenna pointed at zenith, the system noise temperature is defined at the XTR feedhorn aperture input system reference 1 as $T_{\text{op1}} = 17.12$ K. These results are for a Goldstone "clear weather" day whereas the DSN 810-005 document [24] provides averaged weather performance. Weather statistics are useful for mission planning. The atmospheric contribution of noise temperature and loss as a function of elevation angle and cumulative distribution (CD) value are added to the vacuum zenith noise temperature for the overall system performance needed for operational applications.

It is useful to perform a noise temperature calibration sequence of five or six sequential independent measurements with the antenna pointed toward zenith. This provides statistical results for the system operational noise temperature as well as an analysis of the receiving system nonlinearity (as described in Section 2.6) and the calibration noise diodes.

The calibration noise diodes (not shown in Fig. 2-13) are installed in a separate "noise box" module and connected to the side arm of the "CAL COUPLER," usually a waveguide directional coupler. With a 35-dB coupler, injecting 100-K noise requires a noise diode capable of generating more than 100 K × 3162.3 = 316,230 K, accounting for cabling and other losses. With the noise diode turned off and a physical temperature of 300 K, the noise temperature coupled into the system is the physical temperature reduced by the 35 dB, or 300/3162.3 = 0.09 K. Low-noise receiving systems require large coupling factors to reduce the noise temperature contribution. It is not necessary to account for this coupling factor beyond using a, b, and c calibrations as long as the coupler termination is at the same physical temperature as the feed. Ignoring resistive losses, the equivalent coupler "main line" loss L, due to the coupler side arm coupling factor L_c is given by $L_c/(L_c - 1)$. For a 35-dB coupler, with L_c = 3162.3, L = 1.000316 or 0.00137 dB. A resistive attenuator with the same 0.00137 dB loss at 300 K would also contribute 0.09 K.

2.5.2.5 Operational System Noise Temperature Calibrations. The above series of measurements for calibrations a, b, and c serve the engineering purpose of determining the various noise temperature parameters of a new or modified antenna or front-end assembly. In addition, some of these calibrations are repeated routinely to monitor the system performance or to requalify a modified system after system changes have been made. The most likely recalibration would be for the LNA replacement on the antenna or the recalibration of a repaired LNA on the ground to verify the input noise temperature, T_{LNA2}, requiring the use of the gain standard feedhorn and the associated ground instrumentation. In addition to these tests it is necessary to perform system linearity performance verification of the system and confirm that the measurement instrumentation is accurate.

The primary product of the above calibrations is T_{AMW} for calculating the system operating noise temperature with application to tracking spacecraft under varying weather conditions.

$$T_{\mathrm{op}} = T_{\mathrm{sky}} + T_{\mathrm{AMW}} \tag{2.5-23}$$

where

$$T_{\mathrm{sky}} = T_{\mathrm{atm}} + \frac{T_{\mathrm{CMB}}}{L_{\mathrm{atm}}} \tag{2.5-24}$$

and

$$T_{\mathrm{AMW}} = T_{\mathrm{ant}} + T_{\mathrm{feed}} + T_{\mathrm{LNA}} + T_f \tag{2.5-25}$$

For application to mission operations, these terms are all defined at the feedhorn aperture. Missions can estimate T_{op} for DSN antenna systems using documented values of T_{AMW} [24 (module 101B, Eq. A2, p. 36 and Table A-3, p. 40)] and either measured T_{sky} or statistical data [24 (module 105 B, Eq. 1, p. 10 and Table 13, p. 53)].

For routine system noise temperature calibrations of T_{op}, it is customary to use the *Y*-factor method switching between the system ambient (hot) load and the antenna, using the known value of $T_{\mathrm{UWV}}(T_e)$ with Eq. (2.2-8) (item 25 in Table 2-3). T_{AMW} can also be determined from Eq. (2.5-23) ($T_{\mathrm{AMW}} = T_{\mathrm{op}} - T_{\mathrm{sky}}$), but the T_{op} measurement requires knowing T_e, which could change over time. Routine verification of T_{AMW} can be measured

directly, independent of knowing T_e using $T_c = T_{\text{sky}} + T_{\text{ant}}$, and $T_h = T_p$ with $T_e = (T_h - Y_{ch}T_c / Y_{ch} - 1)$, Eq. (2.2-7) (item 22 in Table 2-3) giving

$$T_{\text{AMW}} = \left(\frac{\left(T_p - T_{\text{ant}}\right) - Y_{ah}T_{\text{sky}}}{Y_{ah} - 1} \right) \tag{2.5-26}$$

where

Y_{ah} = Y-factor switching between the antenna (sky) and hot (ambient) load, ratio

T_{sky} data is available in the DSN from a water vapor radiometer. T_{ant} data is available from the initial calibrations on the antenna as shown in Eq. (2.5-22). T_{AMW} combined with T_{sky} as given by the weather statistics for each location determines the overall system noise temperature for each antenna. The T_{AMW} measurement also provides an updated value for $T_{\text{UWV}}(T_e)$ using $T_e = T_{\text{AMW}} - T_{\text{ant}}$.

2.5.2.6 Sources of Noise Temperature Calibration Errors. Table 2-5 shows a, b, and c calibration "1-sigma errors" (Cal a, b, and c in the table). Except for the statistical measurement error, the estimated peak error or 'limit of error' for the "systematic errors" are each divided by three for an estimate of an equivalent 1-sigma' error. This approach [34 (p. 35)] is commonly used for combining the effect of disparate errors for an overall total estimate. Each error is calculated by perturbing the input data, one parameter at a time by the estimated 1σ accuracy, for each type of calibration. The system nonlinearity (NL), voltage standing wave ratio (VSWR), and measurement errors are small compared to the error in T_{sky}; therefore, they are not included for the a and b calibrations. For calibration c, on the antenna, the DSN receiving system requirement for NL, 0.5-percent peak (0.167 percent 1 sigma) is used. For the 17.1-K T_{op} system noise temperature of the XTR feedcone, the resultant NL error is 0.03 K. The dominant error due to mismatch (VSWR) of the microwave components is given by [13 (p. 14, case 2, error 1)]

$$E_{\text{VSWR}} = \left[\frac{1 - 4S_eS_p}{(S_eS_p + 1)^2} \right] \frac{T_p}{Y_{ah}} \tag{2.5-27}$$

Table 2-5. Parameter errors evaluated for calibrations a, b, and c at the DSN Goldstone DSS 14, 70-m antenna XTR feedcone microwave X-band receive (8.42 GHz).

Parameters Evaluated	Input Parameter 1-sigma Errors (boxed) and Output Parameter 1-sigma Errors (not boxed)										RSS Errors
Cal a	T_p, C	P_{load}, dBm	P_{sky}, dBm	P_{off}, dBm	T_{sky1}, K		A_{STD}, dB				
	0.10	0.01	0.01	0.33	0.20		0.003				
T_{LNA2}, K = 4.3950	0.0031	0.0285	0.0292	0.0244	0.2063		0.2082				0.297
T_{LNA}, %	0.072	0.649	0.664	0.555	4.694		4.738				6.76
Cal b	T_p, C	P_{load}, dBm	P_{sky}, dBm	P_{off}, dBm	T_{sky1}, K	T_{LNA2}, K					
	0.10	0.01	0.01	0.33	0.20	0.297					
A_{feed}, dB = 0.0399	0.00005	0.00041	0.00042	0.00036	0.00297	0.00427					0.0053
A_{feed}, %	0.113	1.029	1.055	0.901	7.449	10.714					13.16
Cal c	T_p, C	P_{load}, dBm	P_{sky}, dBm	P_{off}, dBm	T_{sky1}, K	T_{LNA2}, K	A_{feed}, dB	NL, K	VSWR, K	Measurement, K	
	0.10	0.01	0.01	0.33	0.20	0.297	0.0053	0.0285	0.1058	0.0856	
					Engineering						
$T_{UWV}(T_{e1})$, K = 7.4496	0.0010	0.0006	0.0000	0.0215	0.0000	0.3000	0.3720	0.0285	0.1058	0.0856	0.491
$T_{UWV}(T_{e1})$, %	0.014	0.008	0.000	0.288	0.000	4.027	4.993	0.383	1.420	1.149	6.69
T_{ant1}, K = 3.7714	0.0047	0.0388	0.0395	0.0202	0.2000	0.2831	0.3510	0.0285	0.1058	0.0856	0.516
T_{ant1}, %	0.124	1.028	1.047	0.537	5.303	7.507	9.308	0.757	2.805	2.270	13.68
T_{f1}, K = 0.2715	0.0001	0.0006	0.0000	0.0215	0.0000	0.0003	0.0003				0.021
T_{f1}, %	0.033	0.231	0.000	7.902	0.000	0.098	0.121				7.91
					Operations						
T_{op1}, K = 17.1210	0.0057	0.0394	0.0395	0.0012	0.0000	0.0169	0.0209	0.0285	0.1058	0.0856	0.152
T_{op1}, %	0.033	0.230	0.231	0.007	0.000	0.098	0.122	0.167	0.618	0.500	0.89
T_{AMW1}, K = 12.3210	0.0057	0.0394	0.0395	0.0012	0.2000	0.0169	0.0209	0.0285	0.1058	0.0856	0.251
T_{AMW1}, %	0.046	0.320	0.320	0.010	1.623	0.137	0.170	0.232	0.859	0.695	2.04

NL = nonlinearity, RSS = root-sum-squared, VSWR = voltage standing wave ratio

NL error = 0.5% T_{op} divided by 3, VSWR error Eq. (2.5–27) divided by 3 with $S_p = 1.1$ and $S_e = 1.2$

where

E_{VSWR} = error in T_{op} due to mismatched microwave components, K

S_e = LNA input VSWR, ratio

S_p = calibration load input VSWR, ratio

T_p = calibration load physical temperature, K

$Y_{ah} = P_h / P_a$ = hot (ambient) load and antenna *Y*-factor, ratio

Only the VSWR mismatch between the calibration load and the receiver LNA is needed for calculating the error in T_{op} due to mismatch [13]. Although mismatch in the antenna/LNA receiving system microwave components modifies T_{op}, this is not a measurement error. The VSWR's S_p and S_e for the XTR system at band center of 8.4 GHz are estimated as 1.10 (return loss = –26.4 dB) and 1.20 (return loss = 20.8 dB) for the calibration load and the LNA. With T_p and Y_{ah} values of 297.15 K and 17.79 dB for the calibration load and system, the peak mismatch error is 0.317 K (other mismatch effects are less than 0.003 K and are neglected for this calculation) with 0.1058 K 1-sigma value.

The root sum square (RSS) [34 (p. 35)] of these individual error sources are calculated for each output parameter. The RSS error for the key deliverables is less than 1 percent for T_{op1} and 2 percent for T_{AMW} for the Goldstone XTR feedcone example. The biggest error sources for T_{op} and T_{AMW} are due to the VSWR mismatch and inaccuracy in the T_{sky} calibrations, respectively.

2.6 Measurements

2.6.1 *Y*-Factor Noise Temperature Calibrations

A widely used technique for the measurement of the system, receiver, and follow-up amplifier noise temperatures is the *Y*-factor method [4 (p. 26)]. Since thermal noise limits a receiving system's sensitivity performance, it is important to measure this noise source for communication systems. The equations for noise temperature calibrations by the *Y*-factor method follow.

2.6.1.1 Receiver. The *Y*-factor power ratio for the receiver effective input noise temperature measurement configuration shown in Fig. 2-19 by switching between the cold and hot loads, as measured at the receiver output, is given by

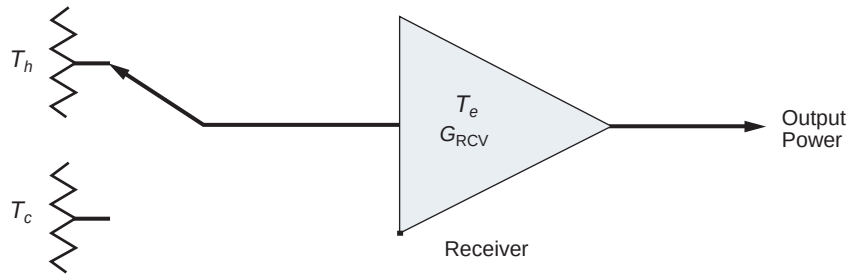


Fig. 2-19. Configuration for determining the receiver noise temperature, T_e, measuring the *Y*-factor, switching between hot and cold terminations.

$$Y_{ch} = \frac{T_h + T_e}{T_c + T_e} \tag{2.6-1}$$

where

$Y_{ch} = P_h / P_c =$ hot (ambient) and cold load *Y*-factor, ratio

$T_h =$ physical temperature of the hot (ambient) load, K

$T_c =$ physical temperature of cold load, K

$T_e =$ receiver effective input noise temperature, K

Although the *Y*-factor ratios are power measurements, both Boltzmann's constant *k*, and the noise bandwidths cancel, resulting in noise temperature ratios, as shown in Eq. (2.6-1), eliminating the need for their determination. In practice, the system bandwidth is usually restricted with a bandpass filter prior to the power-meter input. This reduces measurement errors due to RFI and variations in bandwidth during the measurement process.

Solving Eq. (2.6-1) for the receiver effective input noise temperature, T_e [5 (p. 11-1)]

$$T_e = \frac{T_h - Y_{ch}T_c}{Y_{ch} - 1} \tag{2.6-2}$$

2.6.1.2 System. With T_e known, a similar *Y*-factor ratio measurement technique switching the amplifier input between the antenna and the hot load, as shown in the system configuration Fig. 2-20 [37 (p. 41)], results in the *Y*-factor

$$Y_{ah} = \frac{T_h + T_e}{T_i + T_e} \tag{2.6-3}$$

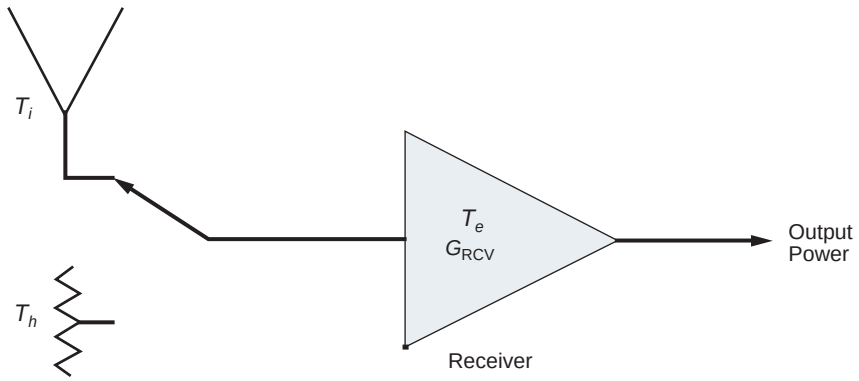


Fig. 2-20. Configuration for determining the system operating noise temperature, T_{op}, measuring the *Y*-factor, switching between the hot termination and the antenna.

where

T_i = antenna noise temperature, including all external noise inputs such as cosmic, atmosphere, ground radiation and microwave loss effects, K

Using $(T_i + T_e) = T_{op}$ and solving for T_{op} considering T_e (as known from previous measurement),

$$T_{op} = \frac{T_h + T_e}{Y_{ah}} \tag{2.6-4}$$

where

$Y_{ah} = P_h / P_a$ = hot (ambient) load and antenna *Y*-factor, ratio

Manual switching using an external aperture load for measuring system noise temperature is shown in Fig. 2-21. The aperture load is alternately placed over and removed from the horn aperture. Good results have been obtained using commercial resistive material designed for absorbing microwave energy.

2.6.1.3 Antenna. From the above two *Y*-factor measurements and using $T_{op} = T_i + T_e$ = system operating noise temperature

$$T_i = T_{op} - T_e \tag{2.6-5}$$

2.6.1.4 Follow-up amplifiers. The receiving system discussed above consists of an LNA and follow-up amplifiers. T_e is composed of the sum of T_{LNA} and T_f

Fig. 2-21. X-band (8.5-GHz) calibrated feed and LNA system demonstrating a manual aperture load noise temperature measurement technique.

$$T_e = T_{\text{LNA}} + T_f \tag{2.6-6}$$

where

T_{LNA} = LNA effective input noise temperature, K

T_f = follow-up amplifiers noise temperature, K

A convenient method to measure T_f is to perform a *Y*-factor measurement, turning the LNA on and off. When the LNA is turned off, the post amplifier

input is terminated in the cold load at the LNA cryogenic temperature, T_{cryo}. See Fig. 2-22.

$$Y_{oo} = \frac{(T_h + T_e)G_{\text{LNA}}}{T_{\text{cryo}} + G_{\text{LNA}}T_f} \tag{2.6-7}$$

where

$Y_{oo} = P_{\text{LNA on}} / P_{\text{LNA off}}$ = LNA on and off *Y*-factor, ratio

T_{cryo} = physical temperature of LNA (estimated input te rmination temperature for the follow-up amplifier when LNA turned off), K

G_{LNA} = LNA gain, ratio

T_e = receiver effective input noise temperature defined at the LNA input, K

T_h = physical hot load temperature (usually ambient load), K

T_f = follow-up amplifiers noise temperature contribution, K

$G_{\text{LNA}}T_f$ = LNA follow-up amplifier noise temperature contribution defined at LNA output, K

Solving Eq. (2.6-7) for T_f

$$T_f = \frac{T_h + T_e}{Y_{oo}} - \frac{T_{\text{cryo}}}{G_{\text{LNA}}} \tag{2.6-8}$$

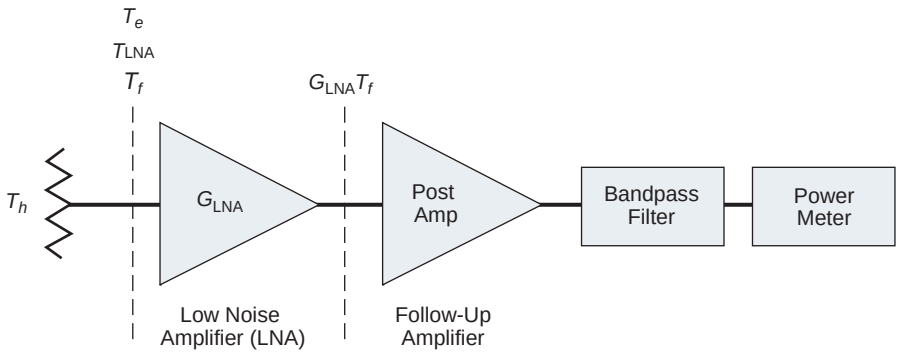


Fig. 2-22. Configuration for determining the follow-up amplifiers noise temperature contribution, T_f, measuring the *Y*-factor, switching the LNA on and off.

This can also be written using $T_e = T_{\mathrm{LNA}} + T_f$, so that, alternately

$$T_f = \frac{T_h + T_{\mathrm{LNA}} - \dfrac{Y_{oo} T_{\mathrm{cryo}}}{G_{\mathrm{LNA}}}}{\mathrm{Y_{oo}} - 1} \tag{2.6-9}$$

Eq. (2.6-8) is useful when T_e is known and Eq. (2.6-9) is useful when T_{LNA} is known.

2.6.2 Attenuation

As discussed in Section 2.2.5 even small attenuation values in microwave components contribute significant thermal noise in low-noise receiving systems. For low-noise receiving systems, 0.01 dB of attenuation at the receiver input increases the system temperature about 0.67 K (~0.67 K per 0.01 dB). This can reduce the system sensitivity (G/T, dB) by much more than the 0.01 dB loss.

It is important to measure the losses of microwave front-end components for calibration and design purposes. Precision calibrations of microwave components require such techniques as lapping waveguide flanges (Fig. 2-23) and the use of waveguide measurement equipment components [5 (p. 15-8)]. This specialized equipment can measure waveguide components to a precision better than 0.001 dB [38]. The LNA input waveguide components require

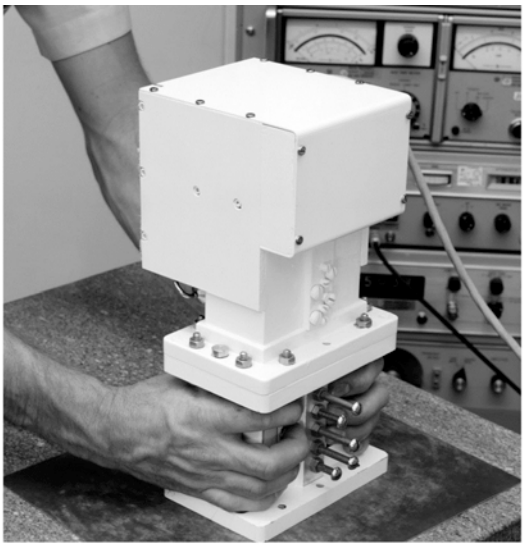


Fig. 2-23. WR430 S-band (2.295-MHz) waveguide flange lapping process for precision-insertion loss measurements.

precision waveguide flanges for best results. Assembled waveguide systems are maintained with clean, dry, pressurized gas systems. Swept-frequency insertion loss measurements over a bandwidth can be made with a commercial instrument such as the HP 8510. It is good practice to combine both measurement techniques (for both selected frequency or frequencies) with higher accuracy and a wider frequency band with less accuracy.

It is sometimes convenient to determine the loss of microwave components from measurements of a receiver effective noise temperature defined on each side of the loss. From Eq. (2.2-23), it can be shown that the insertion loss (ratio) is given by

$$L = \frac{T_p + T_{e1}}{T_p + T_{e2}} \tag{2.6-10}$$

This technique is used in Section 2.5 for calibration of the XTR feedhorn loss. Similarly, it can be shown that

$$L = \frac{T_p - T_{i1}}{T_p - T_{i2}} \tag{2.6-11}$$

2.6.3 Receiving System Nonlinearity

DSN microwave receiving systems are used for many types of noise measurements (including calibrations of antenna gain and system noise temperature T_{op}), as well as radio astronomy and radio science applications. The accuracy of the measurements requires, among other things, knowledge or verification of the receiver-system linearity. A quantitative and proven technique [39,40], for this calibration involves the use of an ambient (for the hot) load, a non-calibrated noise diode (ND) installed at the receiver front-end (Fig. 2-24), a power meter, and a data collection system at the system output. The analysis is applicable to all types of configurations (such as "total power," "noise adding," and "Dicke" radiometers (discussed in Section 2.7).

Figure 2-25 shows an exaggerated nonlinear system where the input noise temperature has been increased by 50 K (turning the ND on and off) with the receiving system switched sequentially between the antenna and an ambient load. The receiver output power is compressed (solid curve) at the higher output level when switched to the ambient load so that the output reading change is less when turning the ND on and off when switched to the ambient load as compared to the antenna.

An analysis of this nonlinear system, using the linear model (dashed line) would result in incorrect noise temperature results. For example, the reading, R on the antenna, with the ND turned on, of slightly over 0.4 W would imply a

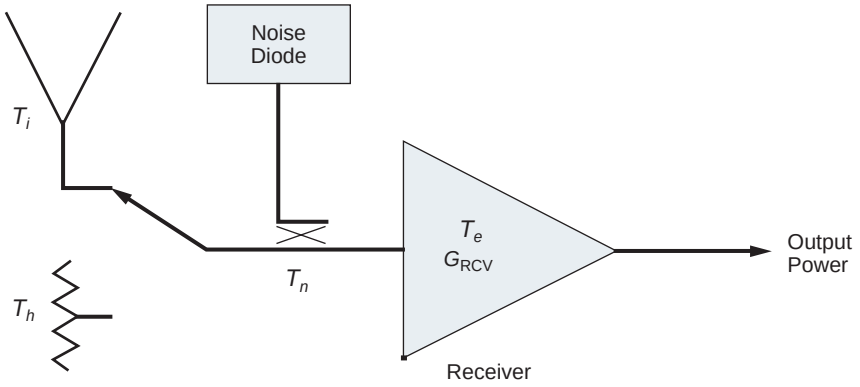


Fig. 2-24. Configuration for determining the system operating noise temperature, T_{op}, and nonlinearity by turning the noise diode on and off while switched to the antenna and the calibration load.

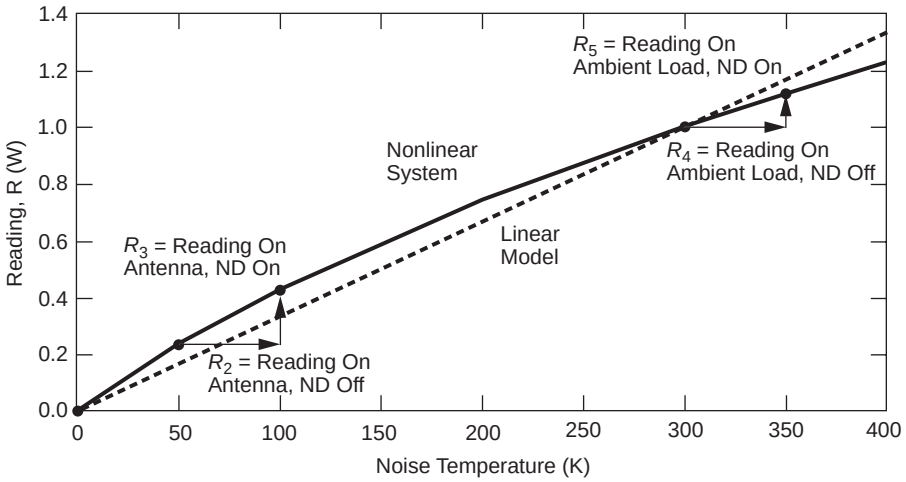


Fig. 2-25. An exaggerated receiving system nonlinearity due to gain compression (solid line) showing 50-K noise diode increase in system noise temperature with the receiver input switched to the antenna and ambient load.

noise temperature of about 125 K, an error of about 25 K (125–100). The following analysis provides the equations to quantitatively measure the degree of receiver nonlinearity and hence allow determination and maintenance of acceptable system linearity limits.

For both the linear and nonlinear model analysis, the receiving system output power readings are defined as

R_1 = power meter input terminated, W

R_2 = receiver input switched to the antenna with noise diode (ND) off, W

R_3 = same as R_2 except ND on, W

R_4 = receiver input switched to ambient load with ND off, W

R_5 = same as R_4 except ND on, W

The R_1 reading is subtracted from all subsequent readings to eliminate the power meter reading zero bias. The four remaining readings are proportional to system noise temperatures T_2, T_3, T_4, and T_5, respectively.

The linear model analysis using the bias corrected readings R_2 and R_4 has two equations of the form

$$T = BR \tag{2.6-12}$$

where

$T = T_{\text{op}}$ = receiving system operating noise temperature, K

B = receiving system linear model scale factor = T_4/R_4, K/W

R = receiving system output power meter readings with input connected to the designated source defined for T, W

For calibration of the receiving system as a Total Power Radiometer (TPR) the LNA is switched to the hot load and the system scale factor determined from Eq. (2.6-12) (see Section 2.7)

$$B = \frac{T_4}{R_4} \tag{2.6-13}$$

where

$T_4 = T_h + T_e$ = receiving system noise temperature, with input switched to the calibration load, K

T_h = calibration load standard (usually an ambient load, with physical temperature monitored), K

$T_e = T_{\text{LNA}} + T_f$ = receiving system front-end amplifier effective input noise temperature, K

R_4 = system output power meter reading, input switched to calibration load, W

Following Eq. (2.6-12) and with Eq. (2.6-13), the system noise temperature with the receiving system input connected to the antenna is given by

$$T_{\text{op}} = T_2 = BR_2 \tag{2.6-14}$$

For this linear analysis, the evaluated noise diode noise temperatures determined by turning the noise diode on and off with input switched to the antenna or the calibration load

$$T_{n2} = T_3 - T_2 \tag{2.6-15}$$

and

$$Tn_4 = T_5 - T_4 \tag{2.6-16}$$

where

T_{n2} = noise diode noise temperature determined from linear analysis with receiver input switched to the antenna, K

T_{n4} = noise diode noise temperature determined from linear analysis with receiver input switched to the load, K

T_2, T_3, T_4, and T_5 = the system noise temperatures corresponding to readings R_2, R_3, R_4, and R_5, K

If the system is nonlinear, the measured values for T_{n2} and T_{n4} will not be equal. A quadratic "corrected" solution (model) with constants B_C and C_C to be determined for the system noise temperature in terms of the linear solution T is given by

$$T_C = B_C T + C_C T^2 \tag{2.6-17}$$

where

B_C = Coefficient of the linear term for T_C, ratio ($B_C = 1$ for a perfectly linear system)

C_C = Coefficient of the quadratic term for T_C, K^{-1} ($B_C = 0$ for a perfectly linear system)

The system response to the four different states (i.e., configured to the antenna and the antenna and ambient load while the noise diode is switched on and off generates four power levels. These four power levels can be used to calculate the coefficient for the nonlinear model.

The resulting four equations are

$$T_{2C} = B_C T_2 + C_C T_2^2 \tag{2.6-18}$$

$$T_{3C} = B_C T_3 + C_C T_3^2 \tag{2.6-19}$$

$$T_4 = B_C T_4 + C_C T_4^{\,2} \text{ (known calibration)} \tag{2.6-20}$$

$$T_{5C} = B_C T_5 + C_C T_5^{\,2} \tag{2.6-21}$$

These four receiving system equations contain the nonlinear characteristics to be determined. The increase in system noise due to the injected noise of the noise diode should be equal for the antenna and ambient load configurations. The differences between T_{3C} and T_{2C} Eqs. (2.6-19 and 2.6-18) should equal the difference between T_{5C} and T_{4C} Eqs. (2.6-21 and 2.6-20). It is also known that the calibration noise temperature represented by T_4 and T_{4C} are equal and identical to the system noise temperature with the receiver input switched to the calibration ambient load $(T_{\text{op AMB}} = T_p + T_{\text{LNA}} + T_f)$.

The actual differences, T_{nC}, in the two cases are equal.

$$T_{nC} = T_{3C} - T_{2C} \tag{2.6-22}$$

$$T_{nC} = T_{5C} - T_{4C} \tag{2.6-23}$$

where

$T_{nC} = \text{ND}$ noise temperature contribution to T_{op} corrected for the receiving system nonlinearity, K

Solving the above equations, the constants B_C and C_C are given by

$$C_C = \frac{T_5 - T_4 - T_3 + T_2}{T_4(T_5 - T_4 - T_3 + T_2) - (T_5^{\,2} - T_4^{\,2} - T_3^{\,2} + T_2^{\,2})} \tag{2.6-24}$$

$$B_C = 1 - C_C T_4 \tag{2.6-25}$$

For a nearly linear receiving system, C_C approaches zero, and B_C approaches one. A receiving system "linearity factor" (FL) is defined with the receiving system connected to the antenna

$$\text{FL} = \frac{T_{2C}}{T_2} = \frac{T_{\text{op}},\ \text{corrected value}}{T_{\text{op}},\ \text{uncorrected value}},\ \text{ratio} \tag{2.6-26}$$

And the nonlinearity (NL) value as

$$\text{NL} = 100(\text{FL} - 1),\ \text{percent} \tag{2.6-27}$$

or

$$\mathrm{NL} = 100\left[\frac{T_{2C}}{T_2} - 1\right], \text{ percent} \qquad (2.6\text{-}28)$$

For the ideal linear receiving system, FL = 1, and NL, percent = 0. A DSN receiving system/radiometer system should have a measured nonlinearity magnitude of less than 0.5 percent; a goal of less than 0.2 percent is desirable. A negative NL value is due to the receiving system compressing at higher noise levels; this is the most common case. A positive NL value is explained by the receiving system gain increasing at higher noise levels; it is theorized that this could be explained by a mixer with insufficient local oscillator (LO) drive level.

Some effort has been expended looking for other curve fits than the quadratic presently used. However, the quadratic fit is well understood, it is simple to evaluate, and no advantage has been found for other models of this application with small system nonlinearity [41 (R. Unglaub, p. 15)]. The analysis is successfully performed routinely with the Goldstone antennas and is planned for all DSN antennas to monitor their noise temperature and linearity performance.

2.6.4 Receiving System Mini-cals

The equations of Section 2.6.3 are used to analyze receiving system noise temperature, linearity, and the noise diode (ND). The mini-cal data sets for this analysis consist of measurements of the receiver output power with the input switched sequentially to the antenna and calibration load (usually ambient temperature) with the ND on and off for each condition.

For example, in the Goldstone DSS-13, research 34-m antenna S-band (2.295 GHz) low noise system data set taken on 2004, day-of-year 357, the averaged system noise temperature (linear analysis) at zenith was 35.8 K (5 data points, measurement 1 sigma = 0.054 K) defined at the feedhorn aperture input. The TPR averaged linear analysis gain scale factor constant B (determined to be 3.893×10^8 K/W (measurement 1 sigma = 3.58×10^5) is used with Eq. (2.6-14) to convert the receiver output power meter readings to system noise temperature. The averaged nonlinearity was measured to be –0.44 percent (measurement 1 sigma = 0.1 percent), indicating only slight receiving system gain compression (slight since magnitude is small and compression indicated by the minus sign). The nonlinearity analysis is used to monitor the receiving system linearity performance and provides the information needed to verify, maintain, and report system performance and provide for error budget analysis. The analysis is as an aid used to modify the instrumentation and verify and report system linearity performance rather than to correct the data results for

system nonlinearity. For this example, the noise diode input defined at the feedhorn aperture was measured to be 55.95 K (measurement 1 sigma = 0.037 K).

2.7 Radiometers in the DSN

2.7.1 Introduction

DSN antennas and receivers are frequently configured for use as radiometers. Radiometers are used for DSN applications such as calibrating antenna efficiency needed to support tracking applications and for scientific applications such as planetary blackbody and "radio star" flux measurements. The antenna calibrations are performed during scheduled maintenance periods. The scientific measurements are dependent on availability of DSN antennas when they are not used for normal spacecraft tracking. The requirements for both types of measurements are similar; large stable antennas capable of operating at microwave frequencies with low system noise temperatures. Receiver linearity and amplitude stability are key issues for radiometers, although of less importance for spacecraft tracking. Well calibrated antenna and receiver systems are important for spacecraft tracking to reduce link margins and hence increase data rates and important for the scientific applications for precise results with known and reportable errors. In addition to using the DSN operational antennas for these purposes, a DSN outreach program has made an older Goldstone 34-m antenna available to the Goldstone Apple Valley Radio Telescope (GAVRT) [42] program for full-time radio astronomy use. It is expected that this program will expand to include another 34-m antenna not currently used by the DSN.

2.7.2 Total Power Radiometers

A total power radiometer (TPR) is presently available for all the DSN antenna gain measurements and system noise temperature calibrations. This is the simplest type of radiometer (Fig. 2-26) and is the standard by which other radiometers are compared [3,25]. For the TPR, from Eq. (2.6-12), the total system noise temperature is given by

$$T = BR \tag{2.7-1}$$

where

$T = T_{op}$ = system operating noise temperature, K

B = radiometer scale factor $= T_4 / R_4$, K/W

R = system output reading, W

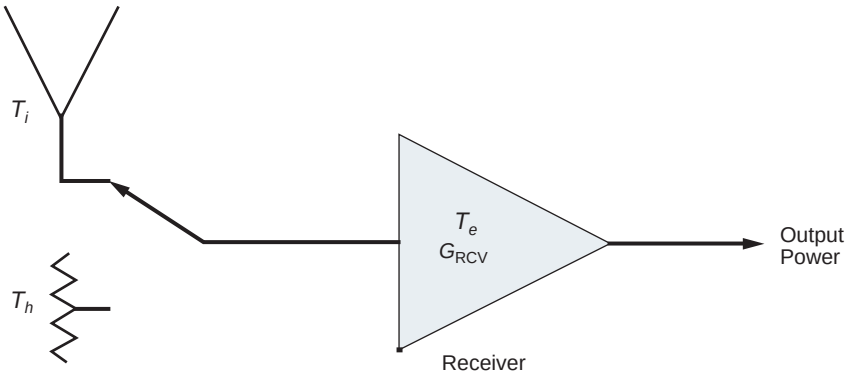


Fig. 2-26. Simplified diagram of a total power radiometer (TPR).

$T_4 = T_h + T_e =$ receiving system noise temperature, with input switched to the calibration load, K

$T_h = T_{\text{load}} =$ calibration of hot load physical temperature, K

$T_e = T_{\text{LNA}} + T_f =$ receiver effective input noise temperature, K

$R_4 =$ system output reading, input switched to calibration hot load, W

The TPR is calibrated by switching the receiver input to the calibration hot load for the determination of the radiometer scale factor *B*. The antenna system with a TPR can perform a variety of radio astronomy measurements. For example, the "absolute" noise measurement of an external natural radio source can be determined by pointing the antenna on and off source. The noise temperature measurement of an unknown source can be determined relative to a known source temperature. This is accomplished by measuring the increase in the maximum on-source noise temperature for both the known and unknown sources. The measurement of antenna gain and other related parameters are determined by observing a radio source with a known temperature [36]. In addition, this radiometer capability is used with natural radio sources to determine antenna pointing model performance.

The minimum detectable noise of the TPR assuming perfect gain stability, is given by [3 (p. 244)]

$$\Delta T_{\text{min}} = \frac{T_{\text{op}}}{\sqrt{\tau B}} \tag{2.7-2}$$

where

$\Delta T_{\text{min}} =$ minimum detectable noise, K

$T_{\text{op}} =$ system noise temperature, K

τ = radiometer integration time, seconds

B = radiometer receiving system bandwidth, Hz

Gain instability can be a serious problem for TPR measurements; more complicated radiometers circumvent this but with some degradation of the minimum $\Delta T_{\min}$. Gain instability [3 (p. 248)] degrades the radiometer sensitivity to

$$\Delta T_{\min} = T_{\text{op}} \sqrt{\left(\frac{1}{\tau B}\right) + \left(\frac{\Delta G}{G}\right)^2} \tag{2.7-3}$$

where

$\Delta G / G$ = radiometer system gain variations, ratio

Instabilities in system noise temperature or bandwidth further decrease the radiometer sensitivity in the same manner as gain variations. Some or most of these can be improved by controlling the physical temperature of the system components. Sometimes this is best accomplished passively using thermal insulation and mass to generate a long time constant compared to the measurement time.

For characterizing the radiometer sensitivities, it is important to measure $\Delta G / G$ with respect to the environment such as the physical temperature and the amplifier's power supply voltages. This is accomplished by changing these parameters one at a time and measuring ($\Delta G / G$). This can be accomplished with the radiometer input switched to a known calibration load. In some cases it is convenient to change the physical temperature with the temperature regulator. This should be done several times noting both the temperature change and ($\Delta G / G$). Similarly, in some cases the line voltage can be changed in a controlled manner with a commercial "variac" transformer. With this information, the radiometer stability performance can be estimated for these parameters.

Radio frequency interference (RFI) within the radiometer operating bandwidth degrades radiometer performance. A site survey for RFI, radiometer equipment shielding design and fabrication and time of observation are important. Strong RFI signals outside the normal operating bandwidth can also saturate the normal power handling capabilities of LNAs and impact radiometer performance. RFI monitoring, detection, and management are important in minimizing the effect of RFI. This includes deletion of corrupted data and (if necessary) rescheduling observations.

2.7.3 Dicke Radiometers

Dicke [3 (p. 248)] radiometers (Fig. 2-27) are used in the radio astronomy community to reduce the effect of receiver gain instability by continuously

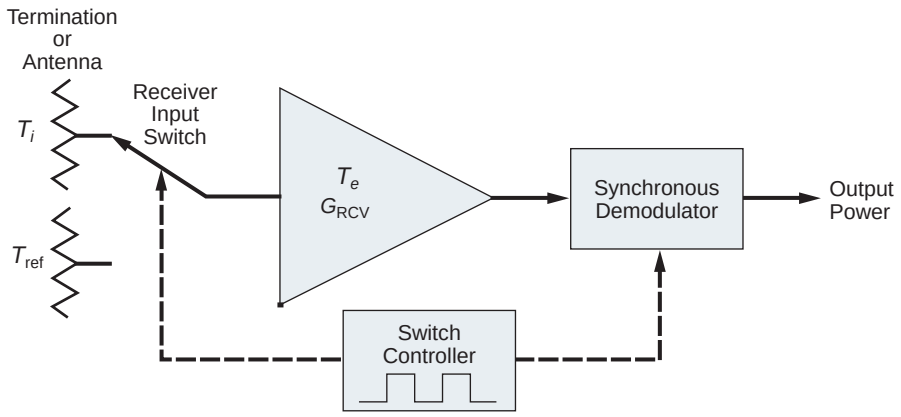


Fig. 2-27. Simplified diagram of a Dicke radiometer.

switching the receiver input between the antenna and a reference load. These are not used in the operational DSN due to compromising the receiving system when switched to the reference load. In addition, the typically higher resistive loss of a fast waveguide switch compared to the specially designed DSN low-loss waveguide switches used in the microwave front-end is incompatible with a very low system noise temperature (0.01 dB loss contributes ~0.67 K noise temperature, see Section 2.6.2).

The sensitivity of a basic "balanced" (reference load (T_{ref}) and antenna noise (T_i) temperatures equalized, square wave multiplication and modulation) Dicke radiometer optimum performance is given by twice the value for a TPR [3 (pp. 248, 258)]

$$\Delta T_{\text{min}} = 2\frac{T_{\text{op}}}{\sqrt{\tau B}} \tag{2.7-4}$$

The DSN Goldstone 70-m antenna research feedcone is equipped with a K-band (22-GHz) beam-switching radiometer configuration using two feedhorns useful for radio astronomy applications. This has the advantage of the Dicke concept, using the second feed as the reference load, thus obtaining the performance of a balanced Dicke radiometer. An added advantage is that atmospheric instabilities are largely canceled since the two feedhorns "see" nearly identical regions of the sky.

2.7.4 Noise-Adding Radiometers

The DSN uses "noise-adding" radiometers (NARs) operationally for monitoring system noise temperature during spacecraft tracking (Fig. 2-28). The injected noise from a noise diode is small compared to the system noise

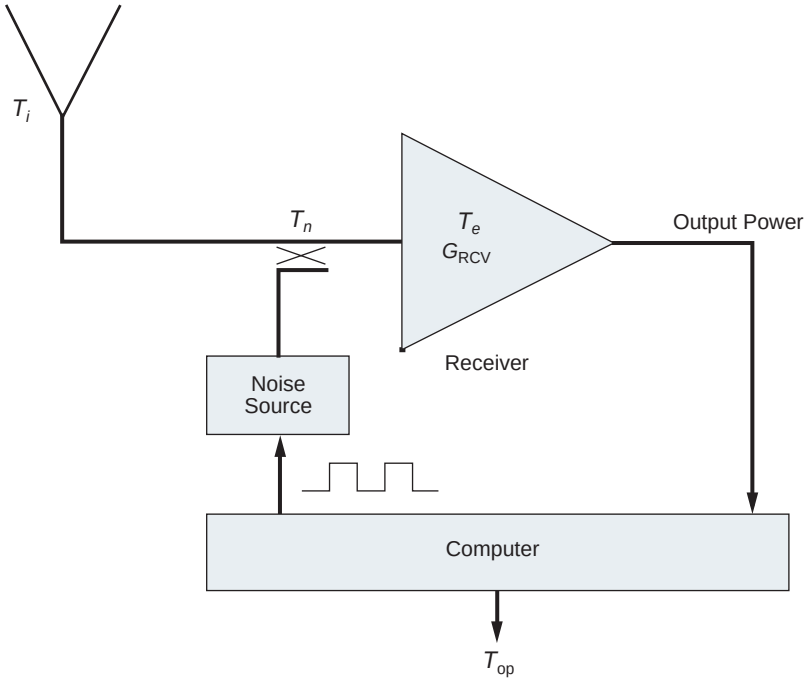


Fig. 2-28. Simplified diagram of a noise adding radiometer (NAR).

temperature (T_{op}) to minimize the impact on the system noise temperature and the modulated telemetry signals.

For the application of radio astronomy observations, antenna calibrations, and system performance measurements the injected noise from a noise diode needs to be large compared to the system noise temperature (T_{op}) to improve the measurement resolution. To satisfy all applications of the NAR, multiple noise diode levels are required.

The NAR *Y*-factor (with the receiving system switched to the antenna) with the noise diode pulsing on and off is given by

$$Y_n = \frac{T_{op} + T_n}{T_{op}} \tag{2.7-5}$$

where

$Y_n = P_{non} / P_{noff} =$ noise diode on and off *Y*-factor, ratio

$T_n =$ noise diode temperature contribution at the input of the receiving system, K

The technique for calibrating the noise diode as required for measuring the system noise temperature depends on the calibration load; if T_e is known, switching the receiver to the calibration hot (usually ambient) load provides a known system temperature, T_{opload}; solving Eq. (2.7-5) for T_n with this condition and measuring Y_n

$$T_n = T_{\text{opload}}(Y_n - 1) \tag{2.7-6}$$

where

$T_{\text{opload}} = (T_{p\text{load}} + T_e) =$ system noise temperature with the LNA connected to the hot (ambient) calibration load, K

$T_h = T_{p\text{load}} =$ physical temperature of the ambient calibration hot load, K

$T_e =$ receiving system noise temperature, K

An alternate method is to perform "mini-cals" (Section 2.6.4) in the TPR mode with the antenna at zenith, measuring system linearity and other parameters as well as the calibration of the noise diode, T_n. Then use the noise diode for the NAR. The noise diode is best evaluated relative to the antenna (T_{2n}) instead of relative to the hot load (T_{4n}) due to the higher measurement scatter on the hot load.

With T_n calibrated and continuously measuring Y_n, the NAR provides a measurement of system noise temperatures

$$T_{\text{op}} = \frac{T_n}{Y_n - 1} \tag{2.7-7}$$

The NAR radiometer noise temperature measurement resolution for equal on and off periods is given by [35; 5 (p. 19-1)]

$$\Delta T_{\text{min}} = \frac{2T_{\text{op}}}{\sqrt{\tau B}}\left(1 + \frac{T_{\text{op}}}{T_n}\right) \tag{2.7-8}$$

Inspection of Eq. (2.7-8) reveals that ΔT_{min} is optimized with large values for T_n. This is not an option for use during tracking spacecraft due to the increase in T_{op}. However, for radio astronomy and other applications, it is desirable to use a large value for T_n. With high values of T_n, $(T_{\text{op}} / T_n) << 1$, the NAR sensitivity performance approaches that of a Dicke radiometer.

However, there is a limit for high values of T_n due to saturation of the receiver amplifier(s) (Section 2.6.4). Eq. (2.7-8) neglects some sources of instability such as the noise diode itself. The noise diode should be temperature and current stabilized [17]. Low-loss microwave couplers are used for the noise diode coupling to the receiving system input, minimizing increased system noise temperature associated with this radiometer capability. The advantage of the NAR for the DSN is that the effects of gain instability are largely eliminated as compared with the TPR.

The external influences for the NAR are mostly transferred from the stability of the amplifiers to the stability of the noise diode. Expanding Eq. (2.7-8) to account for this,

$$\Delta T_{\min} = T_{\text{op}} \sqrt{\left(\frac{4}{\tau B}\right)\left(1+\frac{T_{\text{op}}}{T_n}\right)^2 + \left(\frac{\Delta T_n}{T_n}\right)^2} \tag{2.7-9}$$

where

$\Delta T_n / T_n =$ noise diode instability, ratio

The stability performance for the NAR depends on the noise diode stability instead of the system gain stability.

The system noise temperature for a typical DSN antenna configuration is monitored with the NAR during normal operational tracking of a spacecraft. In this case T_n is usually made quite small, on the order of 0.25 to 1 K to minimize the increase of system noise temperature. This requires a larger number of samples, with a longer total integration time to obtain a suitable measurement resolution. However, longer integration time is limited due to dynamic changes in noise temperature with elevation angle and other effects.

2.7.4.1 Noise Diode Duty Cycle. The NAR performance is analyzed using unequal noise diode on-and-off periods. Treating the noise diode on-and-off periods separately in the NAR Eq. (2.7-7) each with an independent delta noise temperature appropriate for a TPR

$$\begin{aligned}(\Delta T_{\text{op}})^2 = &\left(\frac{\partial T_{\text{op}}}{\partial P_{n\text{off}}}\right)^2 (\Delta P_{n\text{off}})^2 \\ &+ \left(\frac{\partial T_{\text{op}}}{\partial P_{n\text{on}}}\right)^2 (\Delta P_{n\text{on}})^2\end{aligned} \tag{2.7-10}$$

where

$P_{n\text{off}}$ = radiometer output power with noise diode off, W

$P_{n\text{on}}$ = radiometer output power with noise diode on, W

For a single switching cycle

$$(\Delta T_{\text{op}})^2 = \left(\frac{T_{\text{op}} P_{n\text{on}}}{P_{n\text{off}} \left(P_{n\text{on}} - P_{n\text{off}} \right)} \right)^2 \frac{{P_{n\text{off}}}^2}{\tau_1 B} + \left(\frac{T_{\text{op}}}{\left(P_{n\text{on}} - P_{n\text{off}} \right)} \right)^2 \frac{{P_{n\text{on}}}^2}{\tau_2 B} \tag{2.7-11}$$

where

τ_1 = time during the NAR cycle period that noise diode off, s

τ_2 = time during the NAR cycle period that the noise diode on, s

Collecting terms and using

$$\frac{P_{n\text{on}}}{P_{n\text{on}} - P_{n\text{off}}} = \left(1 + \frac{T_{\text{op}}}{T_n} \right) \tag{2.7-12}$$

For 1 cycle

$$\Delta T_{\text{min}} = T_{\text{op}} \left(1 + \frac{T_{\text{op}}}{T_n} \right) \frac{\sqrt{\left[\frac{1}{\tau_1} + \frac{1}{\tau_2} \right]}}{\sqrt{B}} \tag{2.7-13}$$

For multiple cycles with total integration time τ

$$\Delta T_{\text{min}} = m T_{\text{op}} \frac{\left(1 + \frac{T_{\text{op}}}{T_n} \right)}{\sqrt{\tau B}} \tag{2.7-14}$$

where

m = NAR ΔT_{min} multiplier

$= \Delta T_{\text{min}}(F) / \Delta T_{\text{min}}(F = 0.5)$

$= \sqrt{1/(F(1-F))}$ (obtained by se tting Eq. (2.7-13) = Eq. (2.7-14) with $\tau = p$), ratio.

F = fraction of the cycle period time with noise diode on, ratio

$\tau_1 = (1-F)p =$ time during the NAR cycle period with noise diode off, s

$\tau_2 = Fp =$ time during the NAR cycle period with the noise diode on, s

$p = (\tau_1 + \tau_2) =$ cycle period, s

$\tau =$ total integration time allowing for multiple cycles, s

Figure 2-29 shows a plot of *m*, the NAR $\Delta T_{\min}$ multiplier as a function of *F*. The optimum performance (minimum value for $\Delta T_{\min}$) is obtained for *F* = 50 percent, equal noise diode on-and-off times, in agreement with Eq. (2.7-8). For 10-percent on-time, *m* is increased from 2 to 3.33 (67 percent).

2.7.5 Radiometer Stability Performance

System gain stability is one of the most important parameters for a radiometer; a sequence of output power readings from the receiver provides

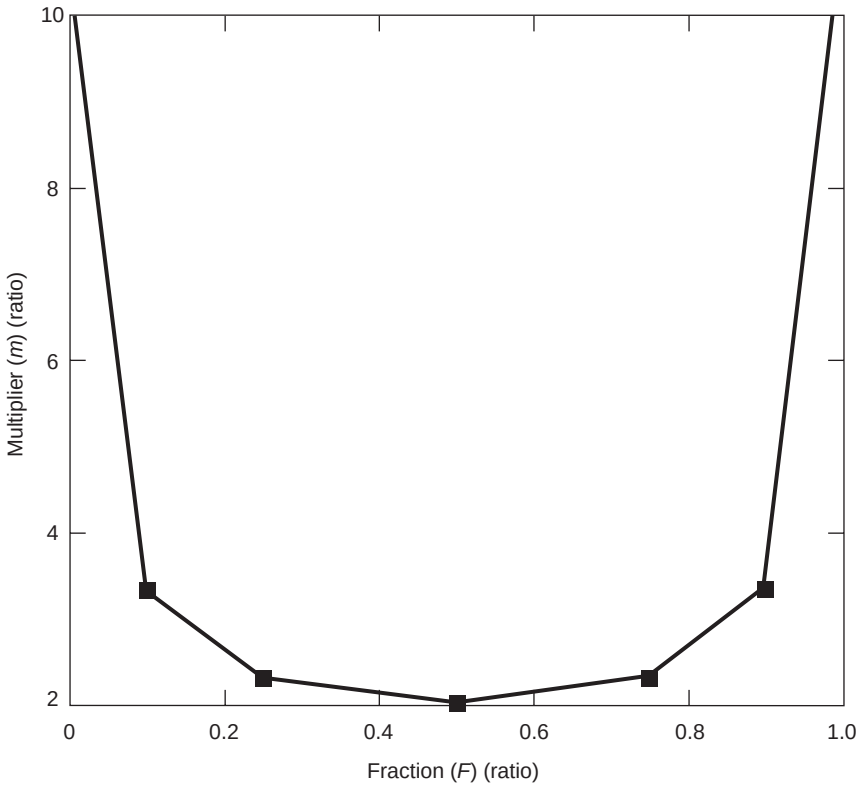


Fig. 2-29. Computation of the noise adding radiometer (NAR) $\Delta T_{\min}$ multiplier *m* as a fraction of the time the noise diode is turned on during an on-off cycle; 50 percent is optimal.

data for this estimate. Since the thermal noise power available from a load is given by $P = kTB$, the system gain is given by

$$G = \frac{R_4}{kT_4B} \tag{2.7-15}$$

where

G = system gain, ratio

R_4 = system output power reading, W

T_4 = system noise temperature when the receiver amplifier input switched to the calibration load, K

k = Boltzmann's constant = 1.38065×10^{-23} J/K

B = BW = system bandwidth, Hz

The physical temperature of the calibration load combined with the amplifier output power readings provides a useful technique for measuring system gain stability. Figure 2-30 shows analysis results of measurements with the Goldstone DSS 12 (GAVRT) 34-m antenna receiving system operating at X-band (8.420 GHz) on 2006, DOY 270 (September 27, 2006). Plot (a) shows a plot of the input thermal noise power $(P = kT_4B)$ relative to the first reading (ratio) and output measured power (R_4) relative to the first reading (ratio) over time combined to calculate the receiving system gain using Eq. (2.7-15) over the same time. This calculation accounts for the changing input power with the measured physical temperature of the calibration load. The gain is reduced by ~0.5 dB over ~5 hours or about 0.1 dB/hr. This was during a physical temperature change of about 9.4 C. It is presumed that the gain change will be reduced during periods of more stable environmental temperatures. This can be achieved with improved temperature control of the system components.

To accommodate gain changes, either frequent mini-calibrations (mini-cals) are helpful, or another configuration is required (such as the NAR). In cases where the noise diode (ND) is more stable than the receiver gain (G), the NAR solution will compensate and could provide the "better" system performance depending on the application requirements and system parameters.

2.8 Status and Future

The science and art of calibrating low-noise receiving systems for the DSN large antennas has been maturing for many years, and will continue to improve in the future. Over the years of the DSN's existence, problems and uncertainties with precision calibrations have been overcome. For example, the DSN has developed, tested, and implemented convenient quantitative methods for

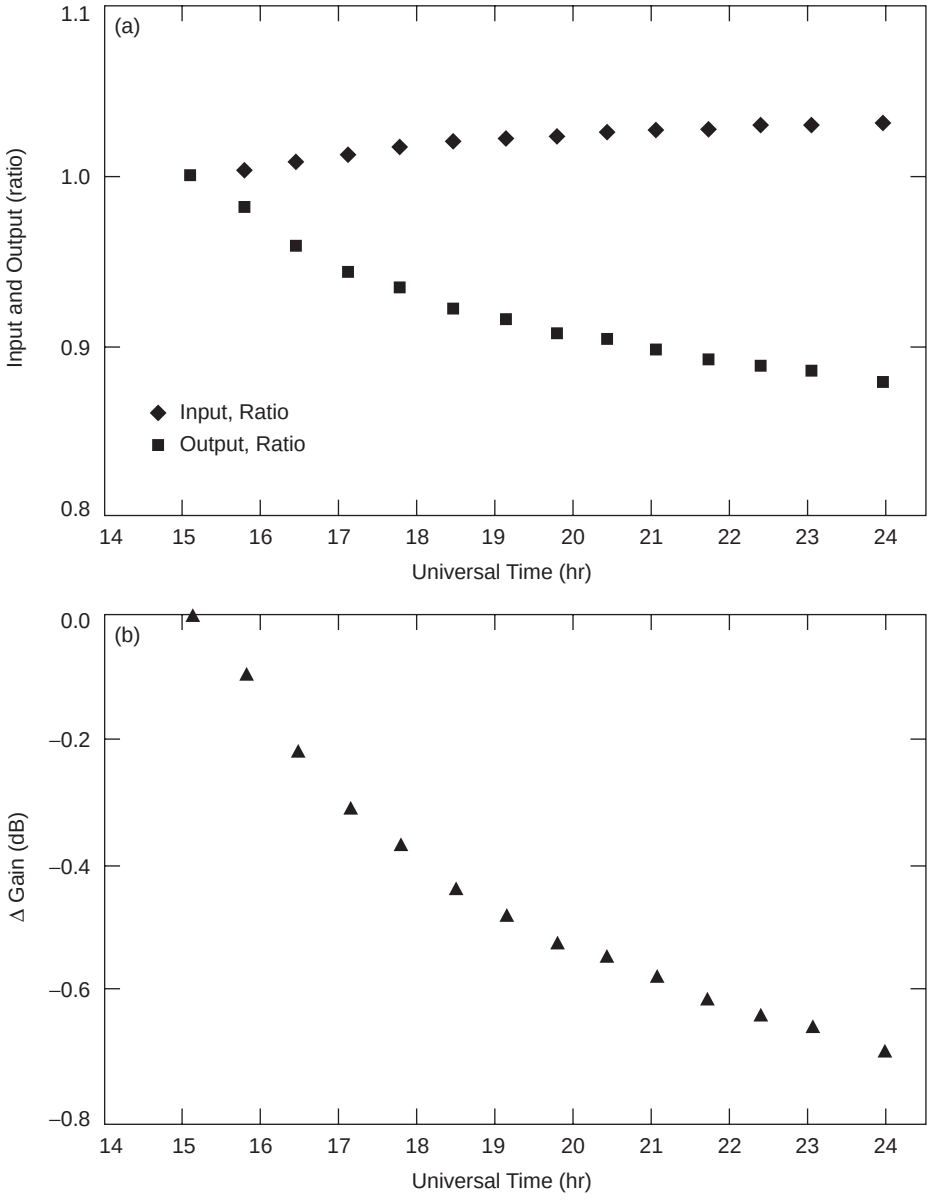


Fig. 2-30. Analysis of Goldstone DSS 12 (GAVRT) 34-m antenna receiving system operating at X-band: (a) input (kT_4B) and output (R_4) changes over time and (b) gain (R_4/kT_4B) change over time.

verifying system linearity. New receiving system feedcones are now calibrated using a tested and agreed-upon sequence of measurements both on the ground and on the antenna. Each antenna system noise temperature is measured and verified periodically using a simple ambient calibration load and power meter.

The theory for these techniques is understood, shared, and coordinated throughout the DSN, often by microwave workshops for engineers and technicians at all DSN locations.

The DSN will continue to make improvements in calibrations and standardization, instrumentation hardware, error analysis, training, and documentation and reporting. The large number of remotely located antennas in the DSN requires diligence in providing future missions with timely and accurate ground station performance data. These are needed for system design, future mission planning, and commitments to optimize data transfer rates between the spacecraft and the worldwide multiple frequency ground antenna systems. This suggests improved automation for calibrating system noise temperature in the DSN antenna systems. These calibrations are essential for validating requirements and also for reducing costs.

In order to meet the higher data rate requirements for deep space tracking in the future, the DSN will develop, implement, and operate large arrays of antennas. These will require sophisticated calibration and operation procedures. The arrays of antennas with lower implementation, operation, and maintenance costs per antenna will result in different constraints and techniques not reflected in this chapter.

Possible improvements for DSN low-noise receiving systems include:

1) Improve the accuracy and calculation convenience for estimating the increase in DSN antennas system noise temperature due to the contributions of the Sun, Moon, planets, and radio sources both near and within each DSN antenna's main beam. This should include sources with solid angles from small to large compared with the antennas solid-angle beamwidth.
2) Improve the accuracy for DSN antennas system noise temperature specifications of T_{AMW} at low elevation angles (6 deg$< T_{AMW} <$15 deg) for all DSN antennas at all DSN communications frequencies.
3) Improve and simplify calibrating techniques for the DSN LNAs at JPL and at the DSN antenna sites, including
   a) Provide LNA noise temperature calibration techniques on the antenna
   b) Provide simplified and accurate noise temperature calibration techniques for the antenna accounting for the atmosphere independent of a water vapor radiometer (WVR).

Challenges for the DSN for future improvements in noise temperature calibrations and standardization include

1) Continuing the worldwide coordination and improvements of noise temperature calibration techniques and standards for the DSN high G/T antenna systems
2) More and improved automation

3) New concepts and methods for future arrayed antenna systems
4) Improve and expand the radio source list for antenna calibrations [36]

Notation and Terms

2005-320 = year-day of year (DOY) date notation example

A = attenuation, dB

A_{atm} = propagation attenuation through the atmosphere, dB

$A_{\text{atm1,2}}$ = propagation attenuation through the atmosphere at elevation angles 1 or 2, dB

$A_{\text{atm}Z}$ = propagation attenuation through the atmosphere at zenith, dB

A_e = antenna effective area, m^2

A_{feed} = feed assembly attenuation, dB = $10^{(L_{\text{feed}},\ \text{ratio}/10)}$

AM = Air Mass = equivalent to 1 atmosphere, ratio

AMW = combined antenna and microwave system

A_p = antenna physical area, m^2

$A_{\text{std}} = 10 \log L_{\text{std}}$, dB

B = BW = noise bandwidth, Hz

B = TPR linear model scale factor, K/W

B_C = coefficient (= 1 for a perfect linear system) of the linear term for T_C, ratio

C_C = coefficient (= 0 for a perfect linear system) of the quadradic term for T_C, K^{-1}

CMB = Cosmic Microwave Background

CMF = Cosmic Microwave Foreground

D_e = antenna effective diameter, m

del = Δ

D_p = antenna physical diameter, m

$\Delta G / G$ = receiving system gain instability, ratio

ΔT_{min} = minimum detectable noise level, K

$\Delta T_n / T_n$ = noise diode instability, ratio

EL = antenna elevation angle, deg

$EL_{1,2}$ = antenna elevation angle 1 or 2, deg

ε = antenna gain efficiency (less than 1), ratio

f = operating frequency, Hz

F = function of the cycle period with noise diode on, ratio

feed assembly = feedhorn + ambient waveguide components

front-end assembly = feed assembly + LNA assembly

FL = T_{2C} / T_2 = receiving system linearity factor (= 1 f or a perfect linear system), ratio

G = available power gain, ratio

G_i = antenna gain relative to isotropic radiator, ratio (G_i , dB = 10 log G_i)

G_m = maximum available power gain, ratio

G_n = gain of amplifier n of cascaded amplifiers, ratio

h = Planck's constant = 6.626069×10^{-34}, Js

HPBW = half-power beam width of antenna pattern main beam, deg

J = joule

k = Boltzmann's constant = 1.38065×10^{-23}, J/K

K = kelvin

L = loss, ratio

λ = wavelength, m

L_{atm} = atmospheric loss, ratio

$L_{atm1,2}$ = atmospheric loss, at elevation angles 1 or 2, ratio

L_{atmZ} = atmospheric loss at zenith, ratio

L_{ATT} = adjustable attenuator located between the post amplifier output and the downconverter input for system gain "level set," ratio

L_{feed} = feed assembly loss, ratio

LNA = low-noise amplifier

LNA Assembly = LNA + post amplifier + gain set attenuator

L_{std} = loss of calibrated standard (std) feedhorn used with the LNA noise temperature calibration defined from the feedhorn aperture to the LNA input, ratio

m = NAR ΔT_{min} multiplier, ratio

$M = G / T$ = receiving system performance figure of merit in terms of antenna gain (relative to isotr opic radiator and receiving system operating noise temperature (relative to 1 K), ratio (M, dB = 10 log M)

N_a = noise power density delivered by the antenna into a matched termination, W/Hz

ND = noise diode

NL= 100 (FL – 1) = receiving system nonlinearity, percent

N_o = amplifier output noise power, W

$p = (\tau_1 + \tau_2)$ = NAR ND cycle period, seconds (τ_1 = ND off time, τ_2 = ND on time)

P_h / P_a = hot (ambient) load and antenna Y-factor, ratio

P_h / P_c = hot (ambient) and cold loads Y-factor, ratio

P_{LNAoff} = receiver output power with LNA off, W

P_{LNAon} = receiver output power with LNA on, W

$P_{n\text{off}}$ = receiver output power with ND off, W

$P_{n\text{on}}$ = receiver output power with ND on, W

P_o = receiver output noise power, W

$Q = (\Delta T_{\text{op}} - \Delta T_{\text{ant}}) / (T_{p\text{atm}} - T_{\text{CMB}})$ = constant used in the quadratic solution of atmospheric loss, ratio

R_1, R_2, R_3, R_4, and R_5 = power meter readings associated with receiving system T_1, T_2, T_3, T_4, and T_5 temperatures, W

Rayleigh-Jeans (R-J) = source noise temperature approximated by the source physical temperature, K

s = second

S_e = LNA input VSWR, ratio

SEP = Sun–Earth–Probe offset angle

S_p = calibration load input VSWR, ratio

SPD = S-band polarization diversity (feedcone)

std = standard

τ = NAR total integration time normally consisting of numerous cycles, seconds

$T = T_2 = T_{\text{op}}$, K

$T_1, T_2, T_3, T_4,$ and T_5 = radiometer system operating noise temperatures using linear equations, K

T_1 = power meter input terminated in ambient termination, K

$T_2 = T_{\text{op}}$ = system operating noise temperature with LNA connected to antenna and ND off, K

T_3 = same as T_2 except ND on, K

T_4 = system operating noise temperature with LNA connected to calibration load with ND off, K

T_5 = same as T_4 = except ND on, K

$T_{2C} = T_{\text{op}}$ corrected for nonlinearity, K

$T_a = T_{\text{ant}} = N_a / k =$ antenna noise temperature, K

$T_{\text{AMW}} = T_{\text{ant}} + T_{\text{feed}} + T_{\text{LNA}} + T_f =$ antenna microwave system noise temperature, K

T_{ant} = antenna noise temperature, K

T_{atm} = noise temperature of the atmosphere, K

T_c = physical temperature of the cold load, K

T_C = corrected receiving system noise temperature in terms of the linear solution, K

T_{CMB} = Cosmic Microwave Background noise temperature, K

T_{cryo} = LNA cryo system physical temperature, K

T_{DC} = down converter input noise temperature, K

T_e = receiver effective input noise temperature, K

$T_{e1,2}$ = receiver effective input noise temperature, defined at reference locations 1 or 2, K

T_{en} = effective input noise temperature of amplifier n of cascaded amplifiers, K

T_f = follow-up amplifier noise temperature, K

T_F = follow-up amplifier noise temperature defined at its input, K

T_{f1} = follow-up amplifier noise temperature defined feedhorn aperture, K

T_{f2} = follow-up amplifier noise temperature defined LNA input, K

T_{feed} = noise temperature of the feed assembly, K

T_h = physical temperature of the hot (ambient) load, K

T_i = receiving system input noise temperature, K

$T_{i1,2}$ = receiving system input noise temperature at reference locations 1 or 2, K

T_k = blackbody disk temperature of the planet, K

$T_{L1} = (L-1)T_p$ = loss noise temperature at reference location 1, K

$T_{L2} = (1-1/L)T_p$ = loss noise temperature at reference location 2, K

T_{LNA} = LNA noise temperature, K

$T_{\text{LNA1,2}}$ = LNA noise temperature at reference locations 1 and 2, K

$T_{\text{load}} = T_h = T_{\text{amb}}$ = physical temperature of the hot (ambient) load, K

T_n = ND noise temperature contribution to T_{op} when turned on, K

T_{n2} = ND noise temperature contribution to T_{op} (LNA input connected to antenna) when turned on, K

T_{n4} = ND noise temperature contribution to T_{op} (LNA input connected to ambient load) when turned on, K

$T_{n\text{C}}$ = ND noise temperature contribution to T_{op} corrected for system nonlinearity, K

$T_{\text{op}} = T = T_{\text{sky}} + T_{\text{AMW}}$ = system operating noise temperature, K

* $T_{\text{op1,2}}$ = system operating noise temperatures at antenna elevation angles 1 or 2, K

* $T_{\text{op1,2,3}}$ = system operating noise temperatures at reference locations 1, 2, or 3, K

$T_{\text{op90,30}}$ = system noise temperature at elevation angles 90 and 30 degrees, K

$T_{\text{opamb}} = T_{\text{opload}} = (T_{p\text{amb}} + T_e) = T_{p\text{load}} + T_e)$ = system noise temperature with the LNA connected to the hot (ambient) calibration load, K

$T_p = T_{\text{phy}}$ = physical temperature of the calibration load and the feed assembly, K

---

* $T_{\text{op1,2}}$ = can be either elevation or reference location depending on the context.

T_{patm} = equivalent physical temperature of the atmosphere, K

T_{pl} = increased system noise temperature due to a planet in the antenna beam, K

T_{P1} = source noise temperature with Planck's radiation law correction, K

T_{pload} = physical temperature of the load, K

T_{ref} = Dicke radiometer reference termination, K

$T_s = P_s / kB$ = signal source equivalent noise temperature contribution to the receiver input, K

$T_{sky} = T_{atm} + T_{CMB} / L_{atm}$ = sky noise temperature due to the atmosphere and CMB, K

$T_{UWV} = T_{e1}$ = microwave receiver effective input noise temperature, defined at feedhorn aperture, K

θ / θ_0 = angle between planet and antenna beam centers relative to the antenna HPBW, ratio

VSWR = voltage standing wave ratio, ratio

WVR = water vapor radiometer

$x = hf/kT$, ratio

XKR = X-band/K-band radar (feedcone

XTR = X-band transmit/receive (feedcone)

$Y_{ah} = P_h / P_a$ = hot (ambient) load and antenna *Y*-factor, ratio

$Y_{ch} = P_h / P_c$ = hot (ambient) and cold load *Y*-factor, ratio

$Y_n = P_{non} / P_{noff}$ = noise diode on and off *Y*-factor, ratio

$Y_{oo} = P_{LNAon} / P_{LNAoff}$ = LNA on and off *Y*-factor, ratio

Z = zenith

References

[1] CCIR, "Worldwide Minimum External Noise Levels, 0.1 Hz to 100 GHz," Report 670, Comité Consultatif International des Radio Communications (French: International Radio Consultative Committee) (CCIR) XIV Plenary Assembly, Kyoto, Japan, Vol. I, International Teelecommunications Union (ITU), Geneva, Switzerland, 1978, pp. 422–427.

[2] W. Flock and E. Smith, "Natural radio noise—A mini-review," *IEEE Transactions on Antennas and Propagation*, vol. 32, issue 7, pp. 762–767, July 1984.

[3] J. D. Kraus, *Radio Astronomy*, McGraw-Hill Book Company, New York, New York, 1966.

[4] W. W. Mumford and E. H. Scheibe, *Noise; Performance Factors in Communication Systems*, Horizon House–Microwave, Inc., Dedham, Massachusetts, 1968.

[5] C. T. Stelzried, *The Deep Space Network – Noise Temperature Concepts, Measurement, and Performance*, JPL Publication 82-33, Jet Propulsion Laboratory, Pasadena, California, September 1982.

[6] A. A. Penzias and R. W. Wilson, "A Measurement of Excess Antenna Temperature at 4080 Mc/s," *The Astrophysical Journal*, vol. 142, no. 1, July 1965, pp. 419–421.

[7] J. C. Mather, D. J. Fixsen, R. A. Shafer, C. Mosier, and D. T. Wilkinson, "Calibrator Design for the COBE Far Infrared Absolute Spectrophotometer (FIRAS)," *The Astrophysical Journal*, vol. 512, pp. 511–520, February 20, 1999.

[8] A. Olivieria-Costa and M. Tegmark, editors, *Microwave Foregrounds: Astronomical Society of the Pacific Series*, Vol. 181, San Francisco, California, 1999.

[9] J. D. Johnson, "Thermal Agitation of Electricity in Conductors," *Physical Review*, vol. 32, no.1, pp. 97–109, July 1928.

[10] H. Nyquist, "Thermal Agitation of Electric Charge in Conductors," *Physical Review*, vol. 32, no. 1, pp. 110–113, July 1928.

[11] *IEEE 100: The Authoritative Dictionary of IEEE Standards Terms*, 7th ed., The Institute of Electrical and Electronics Engineers, Inc., New York, New York, December 2000.

[12] "IRE Standards on Electron Tubes: Definitions of Terms, 1962 (62 IRE 7.S2)," *Proceedings of the IEEE*, vol. 51, pp. 434–435, March 1963.

[13] T. Y. Otoshi, "Antenna System Noise-Temperature Calibration Mismatch Errors Revisited," *The Interplanetary Network Progress Report 42-148, October–December 2001*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–31, February 15, 2002.
http://ipnpr.jpl.nasa.gov/progress_report/

[14] J. Layland and L. Rauch, "The Evolution of Technology in the Deep Space Network: A History of the Advanced Systems Program," *The Telecommunications and Data Progress Report 42-130, April–June 1997*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–44, August

15, 1997, (also available as JPL publication 95-20 at http://tmot.jpl.nasa.gov/, click on Program Overview Information). http://ipnpr.jpl.nasa.gov/progress_report/

[15] W. A. Imbriale, *Large Antennas of the Deep Space Network*, John Wiley and Sons, Hoboken, New Jersey, 2003.

[16] C. T. Stelzried, G. S. Levy, and M. S. Katow, "Multi-feed Cone Cassegrain Antenna," U.S. Patent 3,534,375, October 13, 1970.

[17] K. B. Wallace, "Noise Diode Evaluation," JPL Technical Report, 22-1526, Vol. 3, p. 121, April 15, 1973.

[18] C. T. Stelzried, R. C. Clauss, and S. M. Petty, "Deep Space Network Receiving Systems' Operating Noise Temperature Measurements," *The Interplanetary Network Progress Report 42-154, April–June 2003,* Jet Propulsion Laboratory, Pasadena, California, pp. 1–7, August 15, 2003. http://ipnpr.jpl.nasa.gov/progress_report/

[19] H. Dwight, *Tables of Integrals and Other Mathematical Data*, 4th ed, 9th printing, The Macmillan Company, New York, New York, 1968.

[20] W. V. T. Rusch and P. D. Potter, *Analysis of Reflector Antennas*, Academic Press, New York, New York, 1970.

[21] T. Y. Otoshi, "Noise Temperature Due to Reflector Surface Resistivity," *The Interplanetary Network Progress Report 42-154*, April–June 2003, pp. 1–13, August 15, 2003. http://ipnpr.jpl.nasa.gov/progress_report/

[22] T. Y. Otoshi, Y. Rahmat-Samii, R. Cirillo, Jr., and J. Sosnowski, "Noise Temperature and Gain Loss due to Paints and Primers on DSN Antenna Reflector Surfaces," *Telecommunications and Mission Operations Progress Report 42-140*, October–December 1999, Jet Propulsion Laboratory, Pasadena, California, pp. 1–26, February 15, 2000. http://ipnpr.jpl.nasa.gov/progress_report/

[23] T. Y. Otoshi and M. M. Franco, "Radiometric Tests on Wet and Dry Antenna Reflector Surface Panels," *The Telecommunications and Data Progress Report 42-100, October–December 1989*, Jet Propulsion Laboratory, Pasadena, California, pp. 111–130, February 15, 1990. http://ipnpr.jpl.nasa.gov/progress_report/

[24] *DSMS Telecommunications Link Design Handbook*, 810-005, Rev. E, JPL D-19379 (JPL internal document that has been cleared for external release), Jet Propulsion Laboratory, Pasadena, California, modular document with initial release of Rev. E, January 15, 2001 with ongoing updates. http://eis.jpl.nasa.gov/deepspace/dsndocs/810-005/

[25] R. S. Colvin, *A Study of Radio-Astronomy Receivers*, Publication No. 18A, Stanford Radio Astronomy Institute, Stanford, California, October 31, 1961.

[26] J. R. Pierce and E. C. Posner, *Introduction to Communication Science and Systems*, Plenum Press, New York, New York, 1980.

[27] D. Bathker and D. Brown, "Large Ground Antenna Performance with Solar Noise Jamming," *Proceedings of the IEEE*, vol. 54, issue 12, pp. 1949–1951, December 1966.

[28] I. de Pater, "Radio Images of the Planets," *Annual Review of Astronomy and Astrophysics*, vol. 28 (A91-28201 10-90) Palo Alto, California, Annual Reviews, Inc., pp. 347–399, 1990.
http://www.annualreviews.org/

[29] A. V. Kantak and S. D. Slobin, *System Noise Temperature Increase from the Sun, Moon, or Planet Blackbody Disk Temperature*, JPL D-33697 (JPL internal document), Jet Propulsion Laboratory, Pasadena, California, December 2005.

[30] J. Bautista, R. Clauss, S. Petty, and J. Shell, "DSN Low Noise Amplifiers in the New Millennium," *TMOD Technology and Science Program News*, Issue 13, Rev. 1, Jet Propulsion Laboratory, Pasadena, California, January 2001.
http://tmot.jpl.nasa.gov/Program_Overview_Information/IND_Program_News/ind_program_news1.html

[31] D. F. Wait, "Satellite Earth Terminal G/T Measurements," *Microwave Journal*, vol. 20, no. 4, pp. 49, 51, 58, April 1977.

[32] N. A. Renzetti, *The Deep Space Network—A Radio Communications Instrument for Deep Space Exploration*, JPL Publication 82-104, Jet Propulsion Laboratory, Pasadena, California, July 1983.

[33] P. H. Stanton, D. J. Hoppe, and H. Reilly, "Development of a 7.2-, 8.4-, and 32-Gigahertz (X-/X-/Ka-Band) Three-Frequency Feed for the Deep Space Network," *The Telecommunications and Mission Operations Progress Report 42-145, January–March 2001*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–20, May 15, 2001. http://ipnpr.jpl.nasa.gov/progress_report/

[34] Yardley Beers, *Introduction the Theory of Error*, Second Edition, Addison-Wesley Publishing Company, Menlo Park, California, June 1962.

[35] P. D. Batelaan, R. M. Goldstein and C. T. Stelzried, "A Noise-Adding Radiometer for Use in the DSN," *DSN Space Programs Summary 37-65*, Vol. 2, pp. 66–69, Jet Propulsion Laboratory, Pasadena, California,, September /30, 1970.

[36] P. Richter, "Radio Source List for Antenna Calibration," DSN No. 890-269, JPL D-3801 (JPL internal document), Jet Propulsion Laboratory, Pasadena, California, October 15, 1994.

[37] C. T. Stelzried, "Operating Noise-Temperature Calibrations of Low-Noise Receiving Systems," *Microwave Journal*, vol. 14, pp.41–48, June 1971.

[38] C. T. Stelzried, M. S. Reid, and S. M. Petty, "A Precision DC Potentiometer Microwave Insertion Loss Test Set," *IEEE Transactions on Instrumentation and Measurement*, IM-15, pp. 98–104, March 1966.

[39] C. T. Stelzried and M. J. Klein, "Precision DSN Radiometer Systems: Impact on Microwave Calibrations," *Proceedings of the IEEE*, vol. 82, no. 5, pp. 636–645, May 1994.

[40] C. T. Stelzried and M. J. Klein, Corrections to " Precision DSN Radiometer Systems: Impact on Microwave Calibrations," *Proceedings of the IEEE*, vol. 84, no. 8, p. 1187, August 1996.

[41] A. Freiley, C. Stelzried, R. Unglaub, *Microwave Radiometer System Performance including Linearity Verification*, JPL D-31547 (JPL internal document), Jet Propulsion Laboratory, Pasadena, California, February 28, 2005.

[42] D. F. Miller, *Basics of Radio Astronomy for the Goldstone-Apple Valley Radio Telescope*, JPL Document D-13835 (JPL internal document), Jet Propulsion Laboratory, Pasadena, California, April 1998.
http://www.jpl.nasa.gov/radioastronomy

[43] D. A. Bathker, W. Veruttipong, T. Y. Otoshi, and P. W. Cramer, Jr., "Beam-Waveguide Antenna Performance Predictions with Comparisons to Experimental Results," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 40, no. 6, p. 1279, June 1992.

Chapter 3
Ruby Masers

Robert C. Clauss and James S. Shell

3.1 Introduction

Ruby masers are very low-noise pre-amplifiers used in the microwave receiving systems of the Deep Space Instrumentation Facility (DSIF) and the Deep Space Network (DSN) since 1960. Ruby masers are cryogenically cooled to temperatures below 5 kelvin (K) to achieve extremely low noise levels, and they are used near a focal point of large antennas for receiving signals from spacecraft exploring the Solar System.

Microwave amplification by stimulated emission of radiation (maser) is very different than the process in amplifying devices using a flow of electrons in a crystal (such as transistors) or using a beam of electrons in a vacuum (such as traveling wave tubes, klystrons, or triodes). The maser process may be quantified in terms of photons. A photon at microwave frequencies might be thought of as a particle or an amount of electromagnetic energy equal to hf where h is Planck's constant, 6.626×10^{-34} joule-seconds (J-s), and f is the frequency in hertz (Hz). For example, at 8.42 gigahertz (GHz), a frequency used for deep-space-to-Earth telecommunications, receiving one photon per second provides an amount of power equal to 5.58×10^{-24} watts (W) or –202.5 decibels referenced to milliwatts (dBm). The power level of one photon per second at 8.42 GHz is also equal to the noise power available from a 0.4041-K source in a 1-Hz bandwidth. 0.4041 K is the hf/k quantum noise at 8.42 GHz, where k is Boltzmann's constant (1.38×10^{-23} joule/kelvin (J/K)). The ultimate operating noise temperature lower limit of any linear receiving system is accepted as the quantum noise.

Ruby masers are the most sensitive and lowest noise microwave amplifiers used in the field, yet they are rugged and are not susceptible to the microscopic failures that sometimes occur in sub- micron junction devices. These characteristics of ruby masers result from the stimulated emission process used to amplify a signal. Stimulated emission is a photons-to-photons amplification process. The paramagnetic ruby crystal contains chromium ions, and these ions give the ruby a pink or red color. These ions each have three unpaired electrons. The intrinsic spin angular momentum of each of these electrons gives them the property of a magnetic moment. These three magnetic moments add together to form a larger effective permanent magnetic dipole moment. This permanent magnetic dipole moment associated with each chromium ion is referred to in the following text as a "spin." Spins in particular energy levels are excited by a process called pumping, the application of microwave energy to the ruby. Incoming signal photons stimulate the excited spins, which then drop from a higher energy level to a lower energy level, thereby emitting photons in-phase with, and in numbers proportional to, the stimulating signal photons. This photons-to-photons amplification process enables linear low-noise amplification while being quite immune to the generation of inter-modulation products when strong signals interfere with the reception of weak signals. Ruby masers can withstand input signal power levels of many watts without damage and are not subject to burnout or damage by voltage transients.

Cavity masers using ruby were developed at the Jet Propulsion Laboratory (JPL) at 960 megahertz (MHz) and at 2388 MHz. JPL's first cavity maser installed near the prime focal point on a 26-meter (m) diameter DSIF antenna at a Goldstone, California Deep Space Station (DSS) is shown in Fig. 3-1. The September 1960 occasion was JPL's first liquid helium transfer into an antenna-mounted maser at the DSS-11 "Pioneer" Site. The 960-MHz cavity maser was cooled to 4.2 K by liquid helium in an open-cycle dewar. The maser and the dewar were designed by Dr. Walter H. Higa and built by the group that he supervised at JPL. This first experimental field installation and testing period was followed by the installation of a 2388-MHz cavity maser in February 1961. The 2388-MHz maser was used to receive and amplify microwave radar echoes from the planet Venus [1]. Liquid helium was transferred on a daily basis at a height of about 24 m above the ground, from the fiber-glass bucket of the "cherry-picker" (High Ranger) seen in Fig. 3-1. Servicing the maser was exciting work, especially when the wind speed reached 18 meters per second (m/s) and buffeted the cherry-picker's bucket.

Several single-cavity 960-MHz (L-band) masers were built, some using open-cycle cooling and one cooled by a closed-cycle helium refrigerator [2]. A dual-cavity 2388-MHz maser was developed and used at the Goldstone DSS-13 "Venus" Site to receive radar echoes from Venus and Mars [3]. Traveling-wave maser (TWM) systems were developed for use at various S-band frequencies

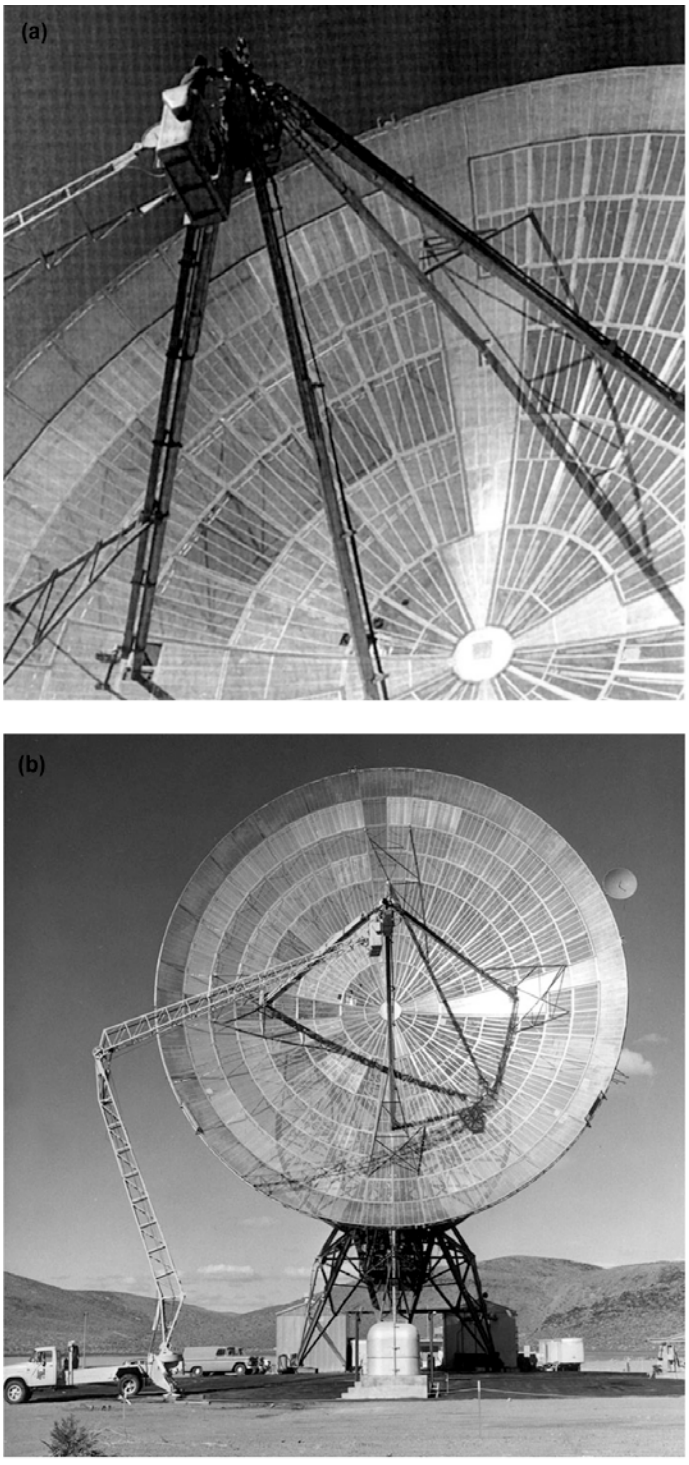


Fig. 3-1. Liquid helium transfer. Including (a) manual transfer of liquid helium at DSS-11 and (b) zoom-out view of cherry picker (high ranger) in use for liquid-helium transfer.

between 2200 MHz and 2400 MHz, X-band frequencies between 7600 MHz and 8900 MHz, and at Ku-band frequencies between 14.3 GHz and 16.3 GHz [4]. K-band reflected-wave maser (RWM) systems were developed for radio astronomy applications at frequencies between 19 GHz and 26.5 GHz. The TWMs and RWMs all operated in closed-cycle refrigerators (CCRs) at temperatures near 4.5 K. An RWM was developed to cover the 31.8-GHz to 32.3-GHz deep-space-to-Earth frequency allocation at Ka-band. A 33.7-GHz two-cavity maser was developed and used at DSS-13 for the Ka-band link experiment with the Mars Observer spacecraft [5].

Table 3-1 lists the ruby masers built for and used in the DSN. Twenty-seven Block III S-band TWMs were implemented. 90 ruby masers were used in the DSN since 1960. CCRs and other key parts were reused as they became available, after replacement by later model masers. R&D masers were built and used in the field, usually on the research antenna at DSS-13 and in the research (radar) cone at DSS-14, to evaluate the long-term performance prior to operational spacecraft tracking commitments. Occasionally R&D masers were installed and used to support and enhance missions following spacecraft problems that affected the communications link.

3.2 Ruby Properties

The pink ruby crystal used in maser amplifiers is about 99.95 percent aluminum oxide (Al_2O_3) with approximately 0.05 percent chromium oxide (Cr_2O_3). The paramagnetic Cr^{3+} ions occupy some of the sites normally occupied by the Al^{3+} ions. Ruby used in the early cavity masers and TWMs was grown by the flame-fusion process and usually contained multiple crystal orientations and variations in the chromium concentration (doping gradients). Inspection and selection of the ruby to find pieces of adequate size and quality were needed. This inspection process was done in polarized light, using a cross-polarized lens to view the flaws in the ruby. Superior quality ruby was developed primarily for laser applications by the Union Carbide Corporation using the Czochralski process. Czochralski ruby first became available for maser applications in 1966.

Ruby is a hard, stable, rugged low-microwave-loss, and high-dielectric-strength crystalline material that can be cut or ground to precise dimensions with diamond tools. The dielectric constant is anisotropic, varying from 11.54 in the direction parallel to the axis of symmetry (c-axis) to 9.34 in the direction perpendicular to the c-axis. The dielectric loss tangent is very low (< 0.0001) and not measurable in ruby-filled cavities or in the slow-wave structures of TWMs. Ruby survives repeated thermal cycling from ambient temperatures above 300 K to cryogenic temperatures below 5 K. The thermal conductivity of ruby is about 1 watt per centimeter-kelvin (W/cm-K) at 4 K and is adequate for

Table 3-1. DSN ruby maser history.

Time in Use[3]	Frequency (GHz)	Maser Type	Bath Temp. (K)	Gain (dB)	Bandwidth (MHz)	Noise Temp. (K)	Quantity[3]
1960–65	0.96	Cavity	4.2	20	0.75	17–30[1]	5
1961	2.388	Cavity	4.2	20	2.5	25	1
1962–63	2.388	Dual cavity	4.2	34	2.5	18	1
1963–66	2.388	R&D TWM	4.5	40	12	8	1
1964–66	8.45	Multiple cavity	4.2	33	17	18	1
1964–71	2.295	Block I S-band TWM	4.5	33	17	9–15[1]	6
1965–71	2.27–2.30	Block II S-band TWM	4.5	35	17–30	9–15[1]	8
1966–68	8.37–8.52	R&D TWM	4.5	30–45	17	18–23[1]	1
1966–74	2.24–2.42	R&D TWM	4.5	27–50	16	4–6[1]	3
1970–89	2.285	Block III S-band TWM	4.5	45	40	4–6[1]	27
1970–72	7.6–8.9	R&D TWM	4.5	30–42	17	7–13[1]	1
1971–80	14.3–16.3	R&D TWM	4.5	30–48	17	8–13[1]	1
1973–89	7.8–8.7	R&D TWM	4.5	45	17–20	7–11[1]	1
1974–now	2.25–2.4	R&D TWM	4.5	30–50	15–30	2–4[1]	1
1975–89	8.42	Block I X-band TWM	4.5	45	40	5–10[1]	12
1979–now	2.285	Block IV S-band TWM	4.5	45	40	2	4
1980–now	8.45	Block II X-band TWM	4.5	40	100	3–4.5[1]	9
1981–95	18–25	RWM	4.5	30	100–300	12	3
1986–95	2.21–2.32	Block V S-band TWM	4.5	35–45	70	3–5	2
1992–now	8.475	R&D TWM	1.6	34	100 1.	9[2]	1
1992–94	33.3–34.0	Dual cavity	1.5	25	85	4–6[1,2]	1

[1] Range (varies across tuning range, from unit to unit, or due to measurement uncertainty)

[2] At the feedhorn aperture

[3] Quantities and time in use dates were obtained via personal communications with D. Hofhine of the Goldstone DSCC, and M. Loria of DSMS Operations.

maser applications. For comparison, thermal conductivity of 50–50 lead-tin soft solder is about 0.15 W/cm-K at 4 K, and the thermal conductivity of electrolytic-tough-pitch copper is about 4 W/cm-K at 4 K.

3.3 Spin Resonance, the Applied Magnetic Field, Ruby Orientation, the Low-Temperature Requirement, and Excitation

The energy of the spins in ruby varies with the intensity and orientation of a biasing magnetic field with respect to the c-axis of the ruby. Two different orientations of ruby are used in JPL masers, depending upon the signal frequency. The "90-degree" orientation used at L-band, S-band, X-band, and Ku-band is discussed here. The "54.7-degree" orientation is discussed later in the sections describing masers at X-band and higher frequencies. The direction of the c-axis can be determined in polarized light.

The four ground-state microwave spin levels that occur are called paramagnetic levels or Zeeman levels. The magnetic field strength and orientation used for an S-band maser at a frequency near 2.4 GHz is given here, for example. A magnetic field strength of 0.25 tesla (T) (2500 gauss (G)) is applied to the ruby in a direction that is perpendicular to the c-axis. This is called the 90-degree orientation. The energy spacings (hf) between the four ground-state energy levels in ruby under these conditions were given in terms of frequency (f). The spacing between these levels is: 1–2 = 2.398 GHz, 1–3 = 12.887 GHz, and 1–4 = 24.444 GHz. Slightly more accurate values are available today, but these values used during the development of masers at JPL between 1959 and 1990 are sufficiently accurate.

The example above uses values from the Appendix, "Ruby energy levels and transition-probability matrix elements," in *Microwave Solid-State Masers* [6]. Chang and Siegman published "Characteristics of ruby for maser applications," in 1958 [7]. Professor Siegman's thorough history and explanation of masers, together with his acknowledgment and descriptions of the research work and publications of many in the maser field are not duplicated here. His book contains a large volume of material about masers, including much that was produced by many workers who shared their knowledge generously. The extensive material published about maser theory and techniques between 1956 and 1964 was most helpful, aiding in the timely development of ruby masers at JPL for the Deep Space Network. Professor Siegman's book is recommended to those interested in the history of masers, and for the detailed theory that leads to a better understanding of masers.

Spin resonance absorption in ruby-filled cavity or waveguide may be observed with a microwave spectrometer as a function of frequency and an applied magnetic field. Absorption of power occurs when signals are applied at frequencies corresponding to the difference in frequency between energy levels.

The resonance absorption line-width of single crystal ruby with about 0.05-percent Cr^{3+} in a uniform magnetic field is about 55 MHz. This value, plus or minus 10 percent, is independent of temperature and has been measured at many different frequencies between 2 GHz and 40 GHz. The S-band example above gives the three frequencies between the 1–2, 1–3, and 1–4 levels. Spin resonance absorption also occurs at frequencies corresponding to the energy differences between levels 2–3 (10.489 GHz), 2–4 (22.046 GHz), and 3–4 (11.557 GHz). The magnitude of the absorption depends upon the ratio of spins in the various levels and is very weak at room temperature. At thermal equilibrium, the ratio of spins in an upper state (N_i) with respect to the lower state (N_j) is an exponential function of energy difference and temperature. The Boltzmann expression gives this ratio and the results show the need for physically cooling the maser material to low cryogenic temperatures.

$$\frac{N_i}{N_j} = e^{-\frac{\mathrm{h} f_{ij}}{\mathrm{k} T}} \tag{3.3-1}$$

where h is Planck's constant, f_{ij} is the frequency difference between levels i and j in hertz, k is Boltzmann's constant, and T is the thermodynamic (bath) temperature in kelvins.

Consider the difference in the ratios of spins in levels 2 and 1 (N_2 / N_1) in the S-band example above at various temperatures. When $T = 300$ K, $\mathrm{h}f_{12} / \mathrm{k}T = 3.8362 \times 10^{-4}$ and the ratio is 0.99962. The two spin populations are almost equal, and the absorption measured with a microwave spectrometer is very weak. As the temperature is lowered from 300 K to 100 K, 20 K, 10 K, 4.5 K, 2.5 K and 1.5 K, the ratios (N_2 / N_1) decrease from 0.99962 to 0.99885, 0.99426, 0.98856, 0.97475, 0.95501, and 0.92615, respectively.

Whether large or small, when the ratio of spins in the upper level to the lower level is less than 1, signal absorption occurs. Amplification by stimulated emission depends on a population inversion where the ratio is greater than 1. The low temperature advantage occurs in both the absorption case and the emission case.

Inversion of the spin population between energy levels 1 and 2 requires the use of an additional level or levels. The Boltzmann expression shows that, in thermal equilibrium, the number of spins in each level decreases from the lowest to the highest level. The example of the S-band maser at a temperature of 4.5 K is used here. The ratio of spins in level 2 to level 1 is 0.97475. The ratio of spins in level 3 to level 1 is 0.87159. The ratio of spins in level 4 to level 1 is 0. 77051. Application of a sufficiently strong pump signal at the frequencies of either 12.887 GHz or 24.444 GHz will equalize the spin

populations of levels 1 and 3, or 1 and 4. Either of these frequencies can be used as the "pump" frequency to excite the spin system, to create a population inversion.

The total number of spins in ruby with slightly more than 0.05 percent Cr^{3+} is about 2.5×10^{19} per cubic centimeter (cc). The S-band cavity maser used as an example has a ruby volume exceeding 1 cc. The actual dimensions are not important for this excitation example. The ratios of the spin densities (in spins per cubic centimeter) in one level to another level are important. Consider x number of spins in level 1 for the thermal equilibrium un-pumped case. The spins in levels 1, 2, 3, and 4 are x, 0.97475 x, 0.87159 x, and 0.77051 x, for a total spin population of 3.61685 x. The fraction of spins in each level are, level 1 = 0.27648, level 2 = 0.26950, level 3 = 0.24098, and level 4 = 0.21304. The spin density values for each of the levels are level 1 = 6.9120×10^{18}, level 2 = 6.7375×10^{18}, level 3 = 6.0245×10^{18}, and level 4 = 5.3260×10^{18}. The level 2 spin density value is 1.745×10^{17} less than level 1. Equalizing the spins in levels 1 and 4 reduces the spin density in level 1 from 6.9120×10^{18} to 6.1190×10^{18} [$(6.9120 \times 10^{18} + 5.3260 \times 10^{18})/2$]. The number of spins in level 2 and level 3 remain unchanged. The ratio of spins in level 2 to level 1 is now 1.1011, a ratio greater than one. The population inversion between levels 2 and 1, resulting from pumping between levels 1 and 4, provides 6.185×10^{17} more spins per cubic centimeter (cc) in level 2 than in level 1.

Emission of the 6.185×10^{17} excess spins in level 2 in a 0.02-microsecond (μs) time period suggests that a pulse from a 1-cc ruby crystal at 2398 MHz could reach 9.8×10^{-7} J (49 W for 2×10^{-8} seconds). This pulse is capable of damaging a transistor amplifier following the maser, as demonstrated in the laboratory during an unfortunate pulse-amplification experiment. A spin-relaxation time of 50 ms for ruby at 4.2 K suggests this pulse amplification process could be repeated at a rate of 20 times per second, resulting with an equivalent continuous power level of about 20 microwatts (–17 dBm). This is about twice the maximum emission level of a –20-dBm signal observed to be available from an S-band TWM. Such a TWM, with low-level signal net gain of 45 dB shows about 1 dB of gain compression when the input signal reaches –84 dBm (the maser output signal is –40 dBm). The maser gain decreases gradually as the input signal strength is increased until the input reaches –20 dBm. At this point, the maser is a unity-gain amplifier. Increasing the input signal above –20 dBm causes the TWM to become a passive attenuator.

Pumping the S-band maser between levels 1 and 4 in the example above appears to give much better maser performance than would be achieved by pumping between levels 1 and 3. Measurements show that this is not the case. This simple example does not include other considerations that affect the performance of masers. Various parameters, including the spin-lattice relaxation time, inversion ratios, transition probabilities, and the filling factor must be considered.

3.4 Spin-Lattice Relaxation Time, Inversion Ratios, Transition Probabilities, the Filling Factor, and Magnetic Q

It is expedient to use the spin-lattice relaxation times measured and published by others. For ruby, in the microwave region at a temperature of 4.2 K, these times are of the order of 0.050 seconds. Below 5 K, these times are roughly inversely proportional to temperature [6, p. 225]. It seems reasonable to expect the spin relaxation times for the various levels to be roughly equal. Optimum spin relaxation times provide higher inversion ratios than equal spin relaxation times. Inversion ratios also vary in accordance with the pumping scheme used. Siegman, [6, p. 292 and 293] mentions four multiple pumped schemes and gives inversion ratio equations for the single-pumping three-level scheme and the multiple-pumping push-pull scheme. The signal frequency (f_s) and the pump frequency (f_p) in the three-level scheme used at S-band results in inversion ratios of $(f_p / f_s) - 1$ for the case of optimum spin-relaxation times and $(f_p / 2 f_s) - 1$ for the case of equal spin relaxation times.

The actual inversion ratio is determined by measuring the signal frequency amplification (the electronic gain in decibels) in the pumped case and dividing it by the signal frequency spin system absorption (in decibels) in the un-pumped case. Spin system absorption and inversion ratios were measured experimentally for many ruby orientations at many different frequencies and for other potential maser materials as well. The measurements at JPL used either a small dielectric resonator in waveguide, in the reflection mode, or the slow-wave structure of a TWM operating in the transmission mode. The use of a dielectric resonator in the reflection mode gives reliable results when the absorption values are a few decibels and amplification values are modest, between a few decibels and 10 dB. Regenerative gain can give gain values higher than those caused by the actual inversion ratio and must be avoided during the inversion ratio measurements.

Our most accurate inversion ratio measurements for S-band used a TWM at 2295 MHz. The pump frequencies used were 12.710 GHz (1–3 transition) and 24.086 GHz (1–4 transition). These calculated inversion ratios are shown in Table 3-2, along with the measured inversion ratios.

The absolute accuracy of the inversion ratio measurement could not resolve the ratio to better than ±0.1, but the difference between the two measurements observed by alternately turning the pump sources on and off was more precise, and it clearly showed the slight advantage for the higher frequency pump. Pump power was adequate for complete maser saturation in both cases.

The results of these inversion ratio tests at 2295 MHz suggest one of three possibilities: (1) The measured inversion ratio of 4.5 is close to the calculated

Table 3-2. Inversion ratios for a 2295-MHz TWM.

Pump Frequency (GHz)	Calculated Inversion Ratios for		Measured Inversion Ratios
	Equal Spin Relaxation Times	Optimum Spin Relaxation Times	
12.710	1.77	4.54	4.5
24.086	4.25	9.5	4.6

value of 4.54 for the 12.710-GHz pump case with optimum spin relaxation times; (2) The measured ratio of 4.6 is less than half of the value of 9.5 calculated for the 24.086-GHz pump case with optimum spin relaxation times, but is not far above the calculated value of 4.25 for the 24.086 pump case with equal spin relaxation times. One might assume that the spin relaxation time from level 3 to level 2 is much shorter than the spin relaxation time from level 2 to level 1, the optimum case. One might also assume that the spin relaxation time from level 4 to level 2 is slightly shorter than the spin relaxation time from level 2 to level 1. Both assumptions could explain the results; (3) There might also be a third explanation for the apparent almost ideal inversion ratio obtained with the 12.710-GHz pump, one that includes nearly equal spin relaxation times. The spacing between level 3 and level 4 is 11.376 GHz, the difference in frequency between the 24.086 GHz pump frequency and the 12.710 frequency. This 3–4 transition has an extremely strong transition probability. The action of the 12.710-GHz pump, that pumps spins from level 1 up to level 3, may also elevate spins from level 3 to level 4, achieving an effect similar to pumping from level 1 to level 4 directly with 24.086 GHz. This process, called push-push pumping, will receive more attention in the X-band maser section of this chapter.

Transition probabilities for the various ruby orientations and for various magnetic field strengths are given in the appendix to Professor Siegman's book [6] that was mentioned earlier. Transition probabilities are directional, having a maximum-probability direction and a zero-probability direction called the null direction. In the case of the 90-degree (deg) orientation where the applied magnetic field is perpendicular to the c-axis, we find the null direction for the signal frequency transition is in the direction of the applied magnetic field. Any microwave magnetic field components in this direction will not interact with the spin system. The c-axis direction is the direction of maximum signal transition probability for the 90-deg orientation in ruby and the radio frequency (RF) magnetic field components in this direction will have the maximum interaction with the spin system at frequencies below 15 GHz. A third direction, perpendicular to the plane determined by the c-axis direction and the applied magnetic field direction, has a transition probability less than the maximum transition probability at frequencies below 15 GHz, being about one-tenth the maximum transition probability at 2295 MHz.

A ruby-filled coaxial cavity, similar to the one used at the 2388-MHz maser for the Venus radar in 1961, is used to give a visual example of the RF magnetic field interaction with the ruby spin system. Figure 3-2 shows the cavity in which the center conductor is shorted at the bottom and open ¼ wavelength from the bottom. The ruby cylinder has a 1.9-cm outside diameter with a 0.636-cm hole through the center, and it is approximately 1 cm in length. The outside diameter of the ruby fits tightly into a copper outer conductor, and the center conductor of the ¼-wavelength resonator is a close fit in the 0.636-cm hole. The center conductor extends slightly beyond the ruby, and this length is adjusted to obtain resonance at the signal frequency. The signal is coupled through the capacitor formed by a sapphire sphere located between the center conductor of the ¼-wavelength cavity and the open-ended center conductor of the coaxial transmission line shown in part above the maser. The orientation of the sapphire sphere affects the capacitive coupling to the cavity

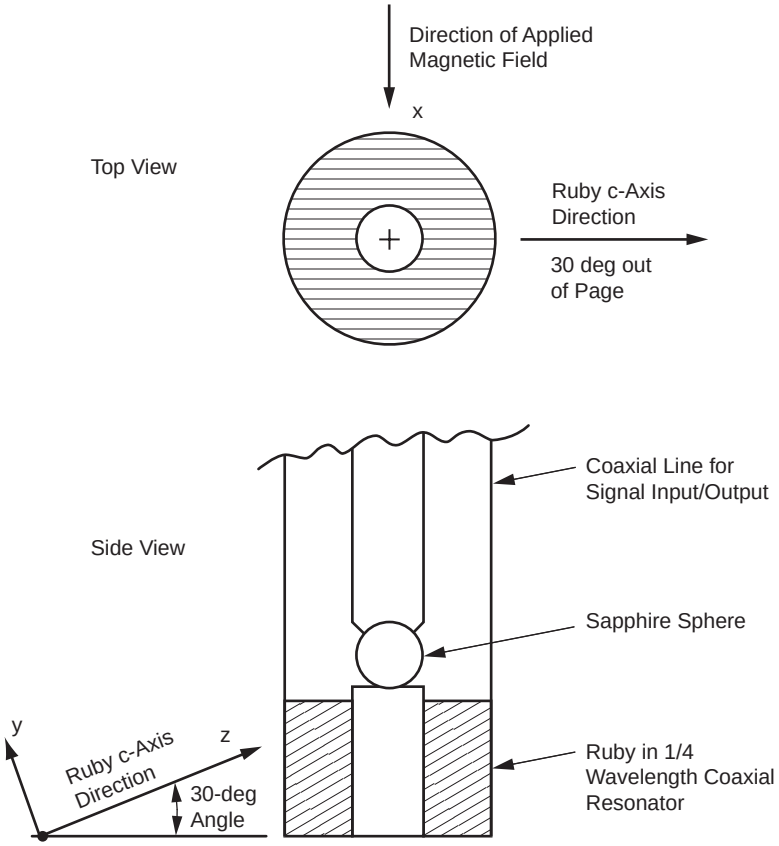


Fig. 3-2. Ruby-filled 2388-MHz coaxial cavity maser. (The ruby c-axis is also the z-axis of maximum transition probability.)

and the resulting cavity resonant frequency; this is caused by the anisotropy of the sapphire's dielectric constant. Figure 3-3 shows a cross-sectional view of the coaxial cavity maser and dewar system.

The ruby cavity is located inside of a liquid helium dewar for cooling to 4.2 K. The ruby c-axis direction in the cavity forms a 60-deg angle with the direction of the center conductor as shown in Fig. 3-2. The applied magnetic field of about 0.25 T comes from a magnet located at room temperature, outside the dewar. The coaxial transmission line and ruby-filled cavity is rotated until the ruby c-axis direction is perpendicular to the direction of the magnetic field.

The directional transition probabilities in " ruby energy levels and transition-probability matrix elements" [6] are given in terms of alpha, beta, and gamma, corresponding to the x, y, and z directions in Fig. 3-2. In this case the x-direction, being the same direction as the applied magnetic field direction, is the null-direction. The c-axis direction is the z-direction of maximum transition probability and is the direction of greatest importance. The projection of the z-direction into the plane of the circular RF magnetic field determines the resultant efficiency or filling factor of about 0.43 for the signal frequency interaction with the spin system. The projection of the y-direction into the plane of the circular RF magnetic field is such that the interaction between the spin system and the RF magnetic field in that part of the cavity is very low, having little affect on amplification. The transition probability in the y-direction is about 1/10 of the transition probability in the z direction. If the c-axis of the ruby had been perpendicular to the direction of the center conductor, in the plane of the signal frequency RF magnetic field, the filling factor would have been near the maximum value of 0.5. The combination of the circular RF magnetic field pattern and the linear c-axis direction results in a 0.5 maximum filling factor when the entire cavity is filled with ruby. The important considerations in determining the filling factor are the placement of the active material in the regions where the RF magnetic field is strong, and a parallel alignment of the RF magnetic field direction with respect to the direction of maximum transition probability.

The filling factor, inversion ratio, and transition probability values all affect the maser's amplifying ability. These, combined with the physical temperature, the maser material's spin density, the number of energy levels, the frequency, and the maser material 55-MHz line-width, determine the magnetic Q value (Q_m). Q_m is used to calculate the gain of a maser, be it a cavity maser, a traveling-wave maser, or a reflected wave maser. A simplified approximation for the Q_m of ruby with 0.053-percent Cr^{3+} (a spin density of 2.5×10^{19} spins/cm^3) at frequencies below 10 GHz and temperatures above 4 K is

$$Q_m \approx \frac{(141) \cdot T}{f I \sigma^2 \eta} \tag{3.4-1}$$

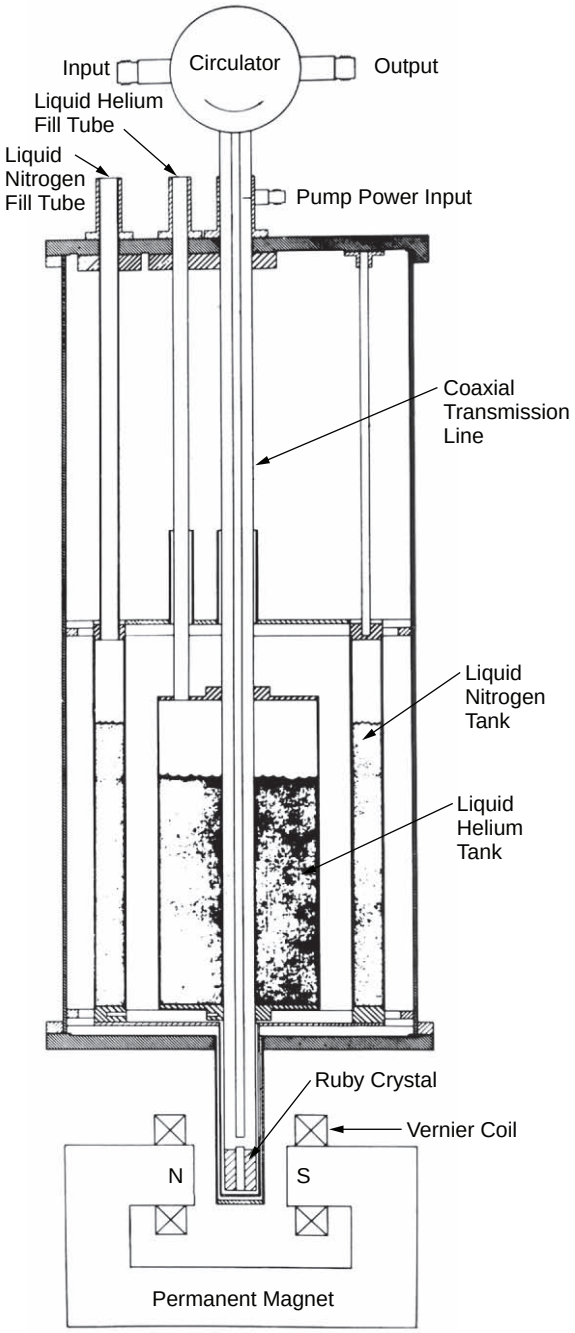


Fig. 3-3. Cross-section of a coaxial cavity maser and dewar system.

where

T = the thermodynamic temperature, K

f = the frequency, GHz

I = the inversion ratio

σ^2 = the maximum value of the transition probability matrix elements, and

η = the filling factor

A Q_m of 71.6 is calculated using $I = 4.3$, $\sigma^2 = 1.6$, and $\eta = 0.5$ for the 2.398-GHz cavity maser at 0.25 T (2500 G) and at 4.2 K. Siegman indicates [6, p. 255] that the approximation $hf / kT << 1$ holds for this estimate of magnetic Q.

The use of the actual spin density difference (ΔN) between the two levels adjacent to the signal transition results in a magnetic Q calculation that is not dependent upon $hf / kT << 1$.

$$Q_m = \frac{4.22 \times 10^{19}}{(\Delta N) I \sigma^2 \eta} \tag{3.4-2}$$

The thermal equilibrium (un-pumped) spin density values for the 2.398-GHz cavity maser example at 4.2 K are level 1 = 6.95925×10^{18}, level 2 = 6.77125×10^{18}, level 3 = 6.00625×10^{18}, level 4 = 5.26325×10^{18}. $\Delta N = 1.88 \times 10^{17}$.

$$Q_m = \frac{4.22 \times 10^{19}}{1.88 \cdot 10^{17} \cdot (4.3) \cdot (1.6) \cdot (0.5)} = 65.2 \tag{3.4-3}$$

The Q_m values of 71.6 for the approximation and 65.2 for the more exact calculation of the S-band example are within 10 percent of each other, suggesting that the low-frequency high-temperature approximation ($hf / kT << 1$) is useful in some cases.

An example of a maser at a higher frequency and a lower temperature uses the 33.7-GHz maser at 1.5 K. Here $I = 2$, $\sigma^2 = 1$, and $\eta = 0.5$. Using the approximation from Eq. (3.4-1) for this case, we calculate a Q_m of 6.26. The energy level distribution for the 33.7-GHz maser is calculated using the equation for the Boltzmann distribution. The frequencies corresponding to these levels are 1–2 = 35.92 GHz, 1–3 = 69.62 GHz, and 1–4 = 105.54 GHz. The signal frequency transition is the 2–3 transition at 33.70 GHz. The thermal equilibrium (un-pumped) spin density values for the 33.70-GHz cavity maser at 1.5 K are: level 1 = 1.7137×10^{19}, level 2 = 5.4303×10^{18}, level

$3 = 1.8473 \times 10^{18}$, and level $4 = 5.8850 \times 10^{17}$. ΔN, the difference between level 2 and level 3, is 3.5830×10^{18}. The Q_m given by Eq. 3.4-2 is 11.77, a factor of 1.88 higher than the Eq. 3.4-1 approximation. Use of the approximation is not acceptable for the higher frequency and lower temperature Q_m calculations.

The gain of a maser depends upon Q_m and upon the interaction time between the RF magnetic field and the spin system. The interaction time in a cavity maser is dependent on Q_1, the loaded Q of the cavity. The cavity maser electronic gain equation in decibels is

$$G_e^{dB} = 10 \cdot \log\left[\frac{(Q_1 + Q_m)}{(Q_1 - Q_m)}\right]^2 \tag{3.4-4}$$

The interaction time in a TWM is dependent upon the TWM's electrical length, the slowing factor (S) multiplied by the physical length in wavelengths (N). The TWM electronic gain equation in decibels is

$$G_e{}^{dB} = \frac{(27.3) \cdot S \cdot N}{Q_m} \tag{3.4-5}$$

These equations are simplified but accurate forms, derived from those found in [6, pp. 268 and 309]. Ruby masers developed at JPL from 1959 through 1992 were based on the material described above. These designs were analyzed by using performance measurement data of the amplifying structures.

Future waveguide-cavity maser designs and TWMs using waveguide resonators as slow-wave structures can be analyzed using computer programs. Various 32-GHz ruby maser designs have been created and analyzed using a combination of commercial design software and a specialized JPL-developed program based on a waveguide mode-matching algorithm. The new analysis of maser designs will be described in a later part of this chapter.

3.5 Ruby Maser Noise Temperatures

A distinct advantage of the ruby maser is its low effective input noise temperature (T_{ampl}). It is also an advantage that the maser noise temperature can be calculated accurately and reliably from the maser gain-and-loss characteristics. The maser net gain, the structure loss, the ruby "pump off" spin-system loss, and the thermodynamic or bath temperature of the maser (T_{b}) are needed for the noise temperature calculation.

The equation used for an accurate calculation of a ruby maser's equivalent input noise does not require that hf/kT be much less than one. It is accurate at all frequencies and temperatures, and it includes the quantum noise contribution. The maser's noise temperature is predicted with calculations broken into two parts. The first part accounts for the noise temperature of the ruby spin system. The second part accounts for noise contributed by the dissipative losses in the slow-wave structure of a TWM, RWM, or in the resonators of a cavity maser. Noise contributions from dissipative losses in components preceding the maser must be accounted for separately.

Siegman has described and discussed noise, in masers and in general [6, p. 398], and Siegman derives the noise power from a maser to be

$$P_{n(\text{ampl})} = (G-1)\cdot\left[\frac{\alpha_m}{\alpha_m-\alpha_0}P_n(-\alpha_m,-T_m)+\frac{\alpha_o}{\alpha_m-\alpha_o}P_n(\alpha_o,T_o)\right] \qquad (3.5\text{-}1)$$

where Siegman's symbols, α_m and α_o, are the electronic gain and loss coefficients per unit length, $(\alpha_m-\alpha_o)$ is the net gain coefficient per unit length, and T_s, the spin temperature, equals $-T_m$. This expression is valid for a traveling-wave maser (TWM), a reflected-wave maser (RWM), and the reflection-type cavity maser.

Numerically, the gain and loss coefficient ratios can be replaced by ratios of the maser's electronic gain $G_e^{(dB)}$, net gain $G^{(dB)}$, and loss $L_o^{(dB)}$ in decibels. Therefore, $\alpha_m/(\alpha_m-\alpha_o)$ equals $G_e^{(dB)}/G^{(dB)}$ and $\alpha_o/(\alpha_m-\alpha_o)$ equals $L_o^{(dB)}/G^{(dB)}$, where the electronic gain is provided by the inverted ruby spin system and the loss is the signal attenuation due to dissipation. The dissipation is caused by the resistance of the copper or other metals used in the structure, by dielectric material loss, and by ferrite or garnet isolator forward loss in the case of a TWM. This loss term, $L_o^{(dB)}$, does not include the spin-system signal absorption that occurs when the maser is not pumped. The net gain in decibels is the electronic gain in decibels minus the loss in decibels.

Using the measured gain and loss values in decibels, and replacing $P_n(T)$ with the Planck radiation law

$$P_n(T) = \frac{hfB}{e^{\frac{hf}{kT}}-1} \qquad (3.5\text{-}2)$$

The equation for the noise power from the maser becomes

$$P_{n(\text{ampl})} = (G-1)\left\{\left(\frac{G_e{}^{dB}}{G^{dB}}\right)\left(\frac{\text{h}\,fB}{1-e^{\frac{-\text{h}\,f}{\text{k}T_m}}}\right)+\left(\frac{L_o{}^{dB}}{G^{dB}}\right)\left(\frac{\text{h}\,fB}{e^{\frac{\text{h}\,f}{\text{k}T_o}}-1}\right)\right\} \tag{3.5-3}$$

and the noise temperature of the maser becomes

$$T_{\text{ampl}} = \left(\frac{G-1}{G}\right)\left(\frac{\text{h}\,f}{\text{k}}\right)\left\{\left(\frac{G_e{}^{dB}}{G^{dB}}\right)\left(\frac{1}{1-e^{\frac{-\text{h}\,f}{\text{k}T_m}}}\right)+\left(\frac{L_o{}^{dB}}{G^{dB}}\right)\left(\frac{1}{e^{\frac{\text{h}\,f}{\text{k}T_o}}-1}\right)\right\} \tag{3.5-4}$$

The spin temperature is determined by the ratio of the spin population densities in the two levels associated with the signal transition.

$$r = \frac{n_i}{n_j} = e^{\frac{-\text{h}\,f}{\text{k}T_s}} = e^{\frac{\text{h}\,f}{\text{k}T_m}} \tag{3.5-5}$$

The ratio r is substituted in Eq. (3.5-4), resulting in the simplified form

$$T_{\text{ampl}} = \left(\frac{G-1}{G}\right)\left(\frac{\text{h}\,f}{\text{k}}\right)\left\{\left(\frac{G_e{}^{dB}}{G^{dB}}\right)\left(\frac{r}{r-1}\right)+\left(\frac{L_o{}^{dB}}{G^{dB}}\right)\left(\frac{1}{e^{\frac{\text{h}\,f}{\text{k}T_o}}-1}\right)\right\} \tag{3.5-6}$$

"Pump saturation" is a requirement meaning the pump power is sufficient to equalize the spins in the energy levels spaced properly for pump excitation. Verification of this pump-saturated condition is determined by increasing the pump power until no further increase in maser gain is observed. The pump-saturated inverted spin population density ratio between levels i and j, denoted by $r = n_i / n_j$, is determined by the difference between the thermal equilibrium level spin densities and the inversion ratio I.

$$I = \frac{n_i - n_j}{N_j - N_i} = \frac{G_e{}^{dB}}{L_r{}^{dB}} \tag{3.5-7}$$

N_i is the spin population of level i when the spin system is in thermal equilibrium, and n_i is the spin population in level i when the ruby spin system is saturated with pump power. The inversion ratio is numerically equal to the ratio of the electronic gain in decibels, with the microwave pump source on, to the ruby absorption in decibels, $L_r^{(dB)}$, with the microwave pump source off.

Depending on the pumping scheme, the expression for r is different. For the three-level system used at S-band, at the 90-deg orientation, the signal transition is between levels 1 and 2, and the pump transition is between levels 1 and 3. The 4th level is not involved, and the number of spins in it is unaffected by the pumping. Pumping equalizes the number of spins in levels 1 and 3, and two equations for the populations in each level are:

$$n_2 - n_1 = I(N_1 - N_2) \tag{3.5-8}$$

and

$$n_2 + 2n_1 = N \tag{3.5-9}$$

where N is the total number of spins in levels 1, 2, and 3; and

$$N = N_1 + N_2 + N_3 = n_1 + n_2 + n_3$$

Multiplying Eq. (3.5-8) by 2 and adding the resulting product to Eq. (3.5-9) results in

$$3n_2 = N + 2I(N_1 - N_2) \tag{3.5-10}$$

Subtracting Eq. (3.5-8) from Eq. (3.5-9) results in

$$3n_1 = N - I(N_1 - N_2) \tag{3.5-11}$$

Dividing Eq. (3.5-10) by Eq. (3.5-11) results in

$$r_s = \frac{1 + 2I(N_{1'} - N_{2'})}{1 - I(N_{1'} - N_{2'})} \tag{3.5-12}$$

where $N_{i'} = N_i / N$

A four-level push-push pumping scheme is used at X-band and Ku-band, and it may also be used at S-band. The signal transition is between levels 1 and 2, and the pump transitions are between levels 1 and 3 and also between levels 3 and 4. Calculating the spin densities in the four levels, using the knowledge that levels 1, 3, and 4 are equalized by pumping, and using the measured inversion ratio I, the ratio of spins in level 2 divided by level 1 is

$$r_x = \frac{1 + 3I(N_{1'} - N_{2'})}{1 - I(N_{1'} - N_{2'})} \tag{3.5-13}$$

where $N_{i'} = N_i / N$ and N is the total number of spins in levels 1, 2, 3, and 4.

A four-level push-pull pumping scheme is used with JPL masers at K-band and Ka-band. The signal transition is between levels 2 and 3, and the pump transitions are between levels 1 and 3 and also between levels 2 and 4. The 54.7-deg ruby orientation produces symmetrical energy levels, so that the 1–3 transition is equal to the 2–4 transition. Calculating the spin densities in the four levels is similar to the processes above. Using the knowledge that the spin density in level 1 equals the spin density in level 3, and that the spin density in level 2 equals the spin density in level 4, and the using the measured inversion ratio *I*, the ratio of spins in level 3 divided by level 2 is

$$r_k = \frac{1 + 2I(N_{1'} - N_{2'})}{1 - 2I(N_{1'} - N_{2'})} \tag{3.5-14}$$

where $N_{i'} = N_i / N$

In contrast with the noise temperature calculations described above, a very simple equation can be used to calculate the approximate noise temperature of a maser (T_{maprx}) when $\mathrm{h}f / \mathrm{k}T$ is much less than one and when the maser gain is very high [5].

$$(T_{\text{maprx}}) = T_b \frac{L_t^{(dB)}}{G^{(dB)}} \tag{3.5-15}$$

where $L_t^{(dB)}$ is the forward loss through the maser structure in decibels with the pump turned off. This loss includes all dissipative losses, such as resistive losses in copper, forward isolator losses, dielectric losses and the ruby (spin-system) signal absorption. $G^{(dB)}$ is the net gain of the maser in decibels. T_b is the thermodynamic temperature or bath temperature of the ruby. A comparison of the results is shown in Table 3-3.

Table 3-3. T_{maprx} values compared with accurate T_{ampl} calculation results.

Parameter	Value			
Frequency (GHz)	2.295		8.430	33.7
Bath temperature (T_b) (K)	4.5	1.6	4.5	1.5
T_{maprx} (K)	1.61	0.89	2.15	0.83
T_{ampl} (K)	1.64	1.02	2.23	2.33
Difference (K)	–0.03	–0.13	–0.08	–1.50

Table 3-3 shows reasonable T_{maprx} values for S-band masers at 2295 MHz and X-band masers at 8430 MHz operating at a bath temperature of 4.5 K. The value of hf/kT is 0.0245 at 2295 MHz and 4.5 K, and 0.0899 at 8430 MHz and 4.5 K. At 8430 MHz and a bath temperature of 1.6 K, the value of hf/kT is 0.2529, and the error of the T_{maprx} calculation is about <15 percent. It seems best to use the more accurate noise temperature calculation when hf/kT is greater than 0.1. Maser performance values used to calculate the maser noise temperatures are shown in Table 3-4.

Variations of maser noise temperature caused by changes in loss, inversion ratio, and bath temperature are quantified to help determine the accuracy of the maser noise temperature. For example, the TWM at 2295 MHz in Table 3-3 is based on a slow-wave structure forward loss of 5 dB. An increase of 1 dB in the structure forward loss value causes a noise temperature increase of 0.12 K. An inversion ratio reduction from 4.5 to 4.0 causes a noise temperature increase of 0.14 K. An increase in the bath temperature from 4.5 K to 4.7 K causes a gain reduction of 2 dB and a noise temperature increase of 0.10 K. Maser performance characteristics at 8430 MHz and 33.7 GHz can be measured to accuracies that are little different than those at 2295 MHz. For example, at 33.7 GHz and a bath temperature of 1.5 K, a bath temperature increase to 1.7 K increases the maser noise temperature by less than 0.1 K. Similar percentage changes of loss and inversion ratio cause maser noise temperature changes of less than 0.1 K. Based on these sensitivity values and our ability to measure maser performance gain and loss characteristics to a fraction of a dB, and the bath temperature to better than 0.1 K, all calculated maser noise temperatures are expected to be accurate to within ± 0.1 K.

These T_{ampl} maser noise temperature values are for the input to the ruby maser at the bath temperature. In the case of a cavity maser, the noise contribution of the circulator preceding the cavity amplifier must be added. In all cases, any loss preceding the maser adds noise to the maser and the

Table 3-4. Maser performance values for noise temperature calculations.

Parameter	Value			
Signal frequency (GHz)	2.295	8.430		33.7
Bath temperature (T_b) (K)	4.5	4.5	1.6	1.5
Net gain (dB)	45.0	45.0	34.0	27.0
Electronic gain (dB)	50.0	49.0	39.0	28.0
Structure loss (dB)	5.0	4.0	5.0	1.0
Inversion ratio	4.5	2.8	2.8	2.0

appropriate correction must be used to determine the effective input temperature at the ambient interface.

The input transmission line from the maser to the ambient interface and the vacuum seal add about 0.5 K ±0.1 K for the Block IV S-band maser described. The effective input noise temperature (T_{m}) of the S-band Block IV maser is 2.1 K ±0.2 K at the ambient interface.

The waveguide input system from the maser to the ambient interface and the vacuum window add about 1.0 K ±0.4 K for the Block II X-band maser described. The Block II X-band maser operates in a closed cycle refrigerator at a bath temperature of 4.5 K. The effective input noise temperature of the maser is 3.5 K ±0.5 K at the ambient interface. The X-band maser cooled to 1.6 K has an effective input noise temperature of 1.8 K ±0.4 K.

The 33.7-GHz cavity maser operates in super-fluid helium at a bath temperature of 1.5 K. The loss of a circulator and isolator preceding the first stage cavity, cooled waveguide, and a partially cooled feed horn with a vacuum window at the ambient interface add about 2.7 ±1 K. The effective input noise temperature of this maser is 5.0 K ±1 K at the ambient interface.

Measurement of a maser's effective input noise temperature using hot and cold loads at 300 K and 80 K is often done in the laboratory at JPL and in the DSN. Using the maser's calculated effective input noise temperatures above (T_m without the tolerances shown), and using the Planck Radiation Law correction for the noise power from the loads, the amplifier output power ratio (Y-factor) obtained while switching from the hot load to the cold load can be calculated. Using this calculated power ratio and the Rayleigh-Jeans approximation for hypothetical hot and cold load measurements gives slightly different maser effective input noise temperatures (T_{m2}). The results are shown in Table 3-5.

The use of T_m for the maser's effective input noise temperature here is the same as the use of T_{lna} for the low-noise amplifier effective input noise

Table 3-5. T_m calculated results compared with hypothetical measurement results (T_{m2}) from the use of the Rayleigh–Jeans approximation with reference loads at 300 K and 80 K.

Parameter	Value			
Frequency (GHz)	2.295	8.430		33.7
T_b (K)	4.5	1.6	4.5	1.5
T_m (K)	2.1	1.8	3.5	5.0
T_{m2} (K)	2.045	1.598	3.298	4.195
Difference (K)	0.055	0.202	0.202	0.805
h*f*/2k (K)	0.055	0.202	0.202	0.809

temperature in other chapters of this book. Table 3-5 shows the differences in the effective input noise temperature of masers caused by the differences in the reference load noise powers that result from the use of the Rayleigh–Jeans approximation. These differences are very close to one-half of the quantum noise ($hf/2k$) for the cases analyzed. A more detailed explanation of this result was published previously [8,9]. System operating noise temperature measurements in the DSN currently use the Rayleigh-Jeans approximation.

3.6 Ruby Masers as Noise Temperature Standards

The DSN ruby masers used as pre-amplifiers to achieve the lowest practical receiving system noise temperatures are also used as noise standards. The maser's effective input noise temperature (T_m) determines the receiver's effective input noise temperature (T_e) and enables the accurate measurement of the total receiving system's operating noise temperature (T_{op}). T_{op} is the sum of T_e and the antenna output temperature (T_i). T_{op} here is defined at the ambient input terminal of the maser, which is the reference 2 location defined in Chapter 2. In this case, T_i includes the sky brightness temperature, antenna pickup from ambient surroundings, noise contributed by reflector losses, and noise contributed by feed system loss when the feed system components are at ambient temperature. Use of masers as noise standards eliminates the need for cryogenically cooled reference terminations in DSN receiving systems.

DSN receiving systems use a waveguide switch and a room temperature load, or an ambient temperature absorber (aperture load) that can be positioned in front of the antenna feed-horn. The aperture load is used to terminate the receiver with a resistive source of known (ambient) temperature. The power ratio measured when the receiver's input is switched from the ambient load to the antenna is used to determine T_{op}.

DSN ruby masers at 2295 MHz, 8430 MHz, and 33.70 GHz achieved effective input noise temperatures of 2.1 K ±0.2 K, 1.8 K ±0.5 K, 3.5 K ±0.5 K, and 5.0 K ±1.0 K respectively, as explained in the previous paragraphs. The measured values agreed with the calculated maser effective input noise temperatures (T_m) at the maser's ambient interface.

The receiver's effective input noise temperature (T_e) is the sum of the maser noise temperature (T_m) and the follow-up receiver noise contribution. High DSN maser gain (typically 45 dB) reduces the follow-up receiver noise contribution to a value that can be measured in two different ways. The follow-up receiver noise temperature can be measured using hot and cold loads. The measured receiver noise temperature is then divided by the maser gain to determine the receiver's contribution to T_{op}. For example, when the noise

temperature of the receiver following the maser is 1000 K and the maser gain is 45 dB (31,623), the follow-up receiver contribution (T_f) to T_a is 1000/31623 = 0.0316 K.

Turning the maser pump source off and on and measuring the resulting receiver power ratio is another method to determine (T_f). This technique, suggested by Dr. R. W. DeGrasse in 1962[1], has been used since in the DSN. In this case, the maser with a 2.1-K noise temperature is connected to an ambient termination near 290 K and receiver output power ratio, Y_{pump}, is measured while turning the maser pump source off and on.

Without the knowledge of the follow-up receiver noise temperature, as used in the first example,

$$\frac{T_{\text{amb}} + T_a + T_f}{T_f} = Y_{\text{pump}} \tag{3.6-1}$$

and

$$T_f = \frac{T_{\text{amb}} + T_a}{Y_{\text{pump}} - 1} \tag{3.6-2}$$

Substituting the known values for the ambient load (290 K for example) and the maser we find

$$T_f = \frac{290\ \text{K} + 2.1\ \text{K}}{Y_{\text{pump}} - 1} \tag{3.6-3}$$

A large value of Y_{pump} is observed when the maser gain is 45 dB. The 1000-K follow-up receiver temperature mentioned above results in a 39.66-dB Y_{pump} measurement. A 10,000-K follow-up receiver temperature would result in a 29.66-dB Y_{pump} measurement, indicating a 0.316-K follow-up receiver contribution. A phone call to a receiver repair person would be in order following such a receiver noise temperature measurement.

A small error in T_f occurs when using the pump on–off method if the maser's pump-off loss is low, and a low-noise follow-up amplifier is used with a maser having modest gain. This error occurs when the noise from the ambient load, attenuated by the maser's pump-off loss, is a significant fraction of the follow-up amplifier noise temperature. Performance values for the 33.68-GHz

[1] Personal communication from Dr. R. W. DeGrasse to R. Clauss in 1962.

dual cavity maser are used here for example. The pump-off loss of 17 dB (14 dB ruby absorption and 3 dB microwave component loss) allows 6 K from a 300-K ambient load to r each the follow-up amplifier. The cooled high electron mobility transistor (HEMT) follow-up amplifier had an effective input noise temperature of 40 K. At 25-dB maser gain, the follow-up noise temperature was 0.1265 K. Using the pump on/off technique results in a follow-up calculation of 0.1455 K which is high by 0.019 K.

When the maser in the examples above is switched from the ambient load to the antenna, the measured power ratio ($Y_{\text{amb-ant}}$) is used to determine T_{op}.

$$T_{\text{op}} = \frac{T_{\text{amb}} + T_m + T_f}{Y_{\text{amb-ant}}} \tag{3.6-4}$$

For example, when ($Y_{\text{amb-ant}}$ is equal to 25.12 (14.00 dB), T_{op} = 11.624 K. This S-band example is used because it is similar to the results achieved with the DSN's S-band receiving system on the 70-m antenna in Australia (DSS-43) used to support the Galileo Mission during recent years. This S-band maser was originally developed to suppor t the Mariner-10 Mercury encounters in 1974 [10]. T_{op} measurements at X-band and Ka-band in the DSN use the same techniques described above.

The accuracy of the ruby maser's effective input noise temperature affects the accuracy of the system operating noise temperature measurement when the maser is used as a noise standard. The S-band example can be used to show that a 0.2-K error in maser noise temperature results in a 0.008-K error in T_{op}. The error is very small because the ($Y_{\text{amb-ant}}$) ratio is large. T_{op} values measured with the antenna at the zenith, looking thr ough one atmosphere in c lear dry weather, range from the S-band low va lue to typic al values of near 20 K at X-band and 30 K for the 33.70-GHz maser system used at DSS-13. The T_{op} error caused by a T_m error is proportional to T_{op}. For example, a 1-K error in T_m causes a 0.07-K error in T_{op} when T_{op} is 21 K, a 0.1-K error when T_{op} is 30 K, and a 0.2-K error when T_{op} is 60 K.

Directional couplers are used in a DSN receiving system to inje ct signals before and after the maser preamplifier to m easure the maser gain. The directional coupler between the antenna feed and the maser is also used to inject a small amount of noise from a noise source (gas tube or noise diode). An advantage of using a small amount of injected noise, that can be turned on and off, is to monitor T_{op} without disconnecting the receiver from the antenna. The injected noise pulse becomes a secondary measurement standard; the noise

pulse excess temperature is determined by comparison with the total system noise power when the maser is connected to the ambient load.

Another important advantage occurs when injecting a signal, or a small amount of noise, while the receiver is alternately connected to the antenna, and the ambient temperature termination. This technique enables the detection of receiver gain changes that might occur when the receiver is switched from the antenna to the ambient termination. Measurement of any receiver gain change that occurs when switching between the antenna and the ambient termination can be used to correct the error that results from the receiver gain change.

Loads or terminations at different temperatures, often at cryogenic and ambient temperatures, can be used to measure receiver noise temperatures. The effective noise temperature of loads at temperatures other than ambient depend upon accurate knowledge of the dissipative surface of the load, and the temperature gradient and loss of the transmission line used to thermally isolate the load from the switch between the receiver and the load. Errors in the effective temperature of cryogenically cooled loads can be caused by moisture condensing upon portions of the interconnecting transmission lines and windows (or gas barriers) required to separate the cryogenic system from the ambient environment.

The maser noise temperature, based on calculations and other performance measurements, is known to the accuracy described above. The input waveguides and transmission lines of the TWMs in CCRs are in a vacuum environment, eliminating the dangers of moisture getting into the lines. The maser's gain affects the maser's noise temperature and is easily measured. A maser gain reduction of 3 dB (out of 45 dB) occurs when the cryogenic temperature of the maser increases from 4.5 K to 4.8 K. The resulting noise temperature increase is 0.165 K for the S-band Block IV TWM and 0.33 K for the X-band Block II TWM. The maser gain is specified to be 45 dB ±1 dB, and noise temperature fluctuations will be accordingly less than 0.06 K at S-band and 0.11 K at X-band.

The effective noise temperature of an ambient load, combined with the use of the noise temperature of a low-noise amplifier with an accurately known noise temperature, eliminates the need for cryogenically cooled reference terminations in deep-space receiving systems. This measurement technique with ruby masers has been used for forty years at the operational DSN frequencies. The values of the total system operating noise temperatures needed for telemetry link evaluations can be measured to an absolute accuracy of better than 1 percent by using the maser as a noise standard.

3.7 Immunity from Radio Frequency Interference (RFI)

Ruby masers use cavities or a slow-wave structure to increase the interaction time between the ruby spin system and the microwave signal. The

cavity, or slow-wave structure, behaves like a filter. The filter properties of the microwave structure provide out-of-band rejection to unwanted signals (RFI). Unwanted, interfering signals within the maser structure bandwidth do not cause measurable inter-modulation products. The measured conversion loss exceeds 100 dB when a ruby maser is subjected to two unwanted input signals within the maser's amplifying bandwidth, at levels as great as –80 dBm. The maser is not a good mixer. In-band signals sufficiently large to cause an amplified maser continuous output level of greater than 0.3 μW reduce the spin population difference associated with the signal transition and reduce the maser gain.

3.8 Early DSN Cavity Masers

Cavity masers mentioned earlier in this chapter were used on DSN antennas at 960 MHz, experimentally in 1960, and later to receive signals from spacecraft going to Venus and to the moon [1]. Cavity masers at 2388 MHz were used to receive radar echoes from Venus and Mars.

The 1961 Venus radar experiment used a 13-kilowatt (kW) transmitter on a 26-m antenna at the Echo Site, and a single-cavity 2388-MHz ruby maser radar receiver preamplifier at the Pioneer Site. The radar receiver achieved a system operating noise temperature of 64 K. William Corliss, in *A History of the Deep Space Network* [1] wrote:

> "… This radar was operated in a two-way, phase coherent mode. Over 200 hours of good data were obtained while Venus was between 50 and 75 million miles from earth.
>
> A scientific result that was also of immense practical importance was the radio determination of the Astronomical Unit (A.U.) as 149, 598, 500 ±500 km, an improvement in accuracy of nearly two orders of magnitude. If the old optical value of the A.U. had been used in the Mariner-2 trajectory computations, its flyby of Venus might have been at a much greater distance with a resultant loss of scientific data."

Experience and information obtained through the use of the cavity masers on DSN antennas at Goldstone have been of great value to the designers and builders of DSN ruby masers. The most important lesson learned by those developing these masers (which were new types of equipment to many and destined for use in the field on large, fully steerable antennas in an operational environment) was the following.

New types of equipment, such as these ruby masers going into the field, should be accompanied by the developers. The developers of new masers should operate, service, and maintain the new equipment in the operational environment for at least several weeks. A quick, one-day, or one-week check-out is not adequate.

Additional lessons learned include the following:

1) Laboratory testing of masers planned for use in the DSN is not sufficient until the field environment is understood well enough to develop all of the needed tests.
2) Support personnel in the field operating, servicing, and maintaining masers will often be blamed for equipment failures beyond their control.
3) Support personnel who will operate, service, or maintain masers should be given appropriate training and documentation that includes installation, operations, and maintenance procedures.
4) Maser developers must understand the electrical and mechanical interfaces to the adjacent subsystems and design the interfaces to withstand mechanical and electrical strain and stress. An example of the needed interface understanding follows. A massive S-band waveguide section from the antenna feed system was attached to the type N connector on the circulator at the input to the maser. Antenna motion caused movement that broke the circulator. Circulators were new and expensive devices in February 1961, and we did not have a spare. The circulator was repaired on site by two very nervous technicians.
5) Maser support personnel working in the DSN should understand the entire receiving system sufficiently well to identify failures that occur in other subsystems. Otherwise any receiving system failure will be identified as a maser failure by the other subsystem engineers.
6) Instrumentation to monitor the cryogenic system's condition and a simple receiver used to measure the maser's performance were essential.
7) The effectiveness of vacuum insulation in a liquid helium dewar system, in a liquid helium transfer line, or in the vacuum housing of a closed-cycle helium refrigerator, is destroyed by the introduction of a very small amount of helium gas. Helium diffusion through rubber o-ring seals, vacuum windows in waveguides and connectors that are made from materials that are not impervious to helium allows helium gas to enter the vacuum jacket.
8) There is no substitute for long-term experience in the field environment. Systems should be field-tested on antennas having an environment as close as possible to the operational environment planned for mission support. Planetary radar tasks provided ideal field experience with systems planned for lunar and interplanetary mission support.
9) Adding a few "improvements" will sometimes disable a system that has worked well previously.
10) The best engineers sometimes make mistakes. This list could be much longer.

A dual-cavity maser at 2388 MHz was used on the 26-m antenna at DSS-13, the "Venus" site in the fall of 1962 for planetary radar to study Venus,

and in 1963 to receive radar echoes from Mars. The environment of an elevation-over-azimuth-drive (Az-El) with a cassegrain feed system was preferred over the prime-focal-point location of the 26-m polar mount antennas by those designing, building, and servicing ruby masers. Liquid helium and liquid nitrogen transfers were made from the antenna's main reflector surface through the wall of the cassegranian feed cone. The thrill of "high altitude" liquid helium transfers from the cherry-picker-bucket being buffeted by high winds was gone and not missed. The antenna elevation motion was from 10 deg above the horizon to the zenith (90-deg elevation). The maser package was mounted at an angle such that the dewar containing the maser and cryogens did not tilt more than 50 deg from vertical, thereby avoiding liquid cryogen spills. A cross-sectional view of one of the two identical units is shown in Fig. 3-4. The first circulator was an S-band waveguide configuration, providing a waveguide interface to the antenna feed system.

The noise temperature of the dual-cavity maser was determined by the loss of the ambient temperature circulator, the loss of a waveguide-to-coaxial-line transition, the loss of the 7/8-inch (~2-cm) diameter coaxial line transition, and the maser in the 4.2-K bath. These losses contributed about 16 K, bringing the maser's effective input noise temperature to 18 K ±2 K. A 35-K total system noise temperature was achieved for the radar receiving system, a significant improvement over the 64-K system at the Pioneer site.

3.9 Comb-Type Traveling-Wave Masers

Bell Telephone Laboratories (BTL) headquartered at Murray Hill, New Jersey with greatest concentration of facilities in northern New Jersey and Airborne Instruments Laboratories (AIL) in the Long Island area had demonstrated and published the characteristics of S-band traveling-wave masers (TWMs) in 1959 and 1960 [11,12]. The TWMs used a comb-type slow-wave-structure (SWS), ruby, and polycrystalline yttrium-iron-garnet (YIG) resonance isolators. The SWS is like a resonant comb-type band-pass filter. Incoming microwave signals to be amplified were coupled into the SWS from a coaxial line and traveled through the SWS at a speed as low as 1/100 the speed of light (the group velocity). The slowing-factor of the SWS is the reciprocal of group velocity divided by the speed of light. Ruby bars were located on both sides of the comb, and the spin system in the ruby provided the signal-frequency amplification. The signal's interaction with the spin system is proportional to the slowing factor. Resonance isolators were located in regions of circular polarization on one side of the comb. These isolators provided high attenuation in the reverse direction to enable stable, regeneration-free amplification of microwave signals traveling in the forward direction.

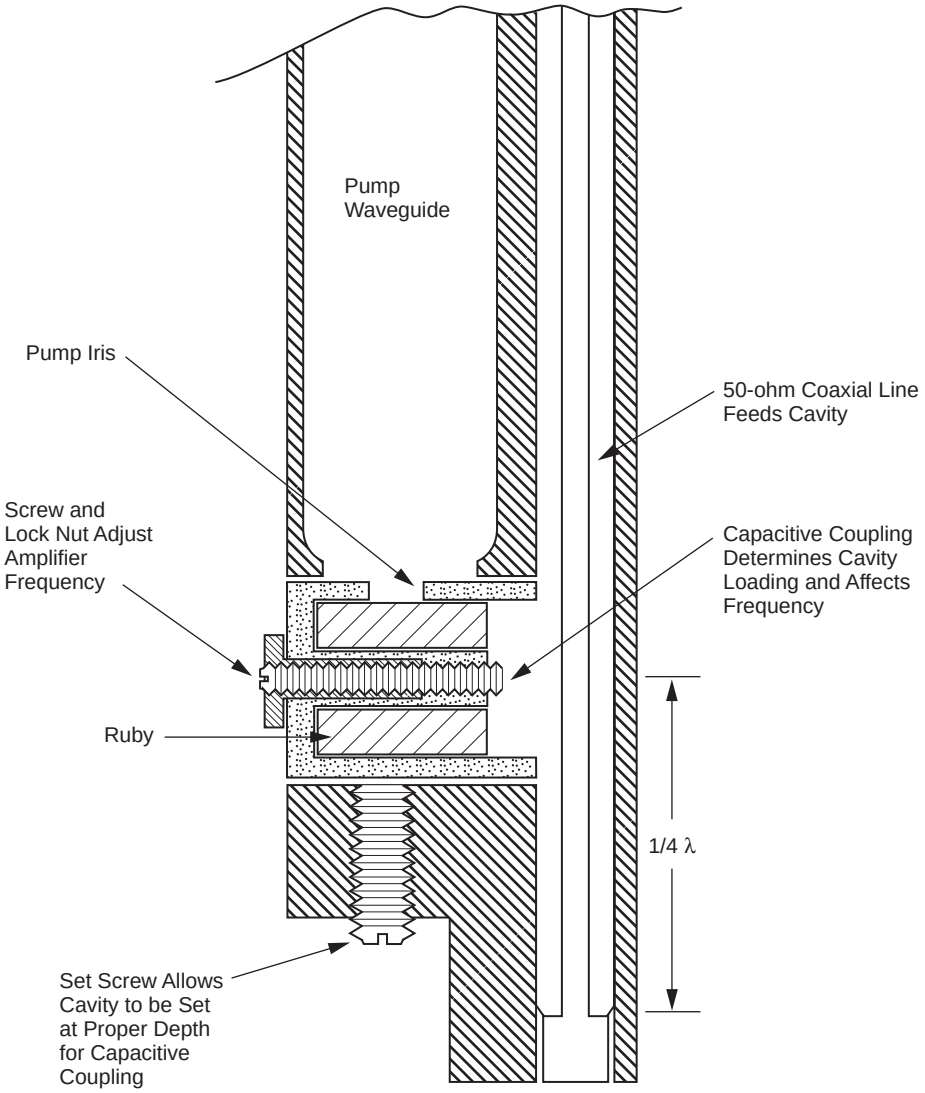


Fig. 3-4. Cross-section of one cavity of the two-cavity maser.

Dr. Walter Higa's group at JPL had concentrated on the development of the cavity masers needed for planetary radar and early mission support prior to 1963. The need to support future missions requiring more bandwidth than the cavity masers were able to provide resulted in the decision to purchase traveling-wave masers from those who were TWM experts. Dr. Robert W. DeGrasse had moved from BTL to the Microwave Electronics Corporation (MEC), providing another commercial source for TWMs. MEC developed a TWM for JPL in 1962. The TWM covered the S-band 2290-MHz to 2300-MHz

allocation for deep-space-to-Earth communications, and it was tunable to the 2388-MHz planetary radar frequency. Signal input transmission line loss and instabilities, and a fragile vacuum seal in that transmission line, prevented use of the MEC TWM in the DSN.

JPL purchased six TWMs for the DSN from AIL in 1964. Six closed-cycle refrigerators (CCRs) were purchased from Arthur D. Little (ADL) to provide 4.5-K cryogenic cooling for the TWMs. The maser group at JPL developed low-loss transmission lines, produced packages that supported a rugged S-band waveguide input interface, and integrated the TWMs and CCRs into antenna-mountable units with the power supplies, controls, instrumentation, and support equipment needed for the complete maser subsystems. The TWM/CCR subsystems included antenna cabling and helium gas lines connecting the helium compressor located near the antenna base, to the TWM/CCR package located in the cassegrain cone of the 26-m polar-mount antennas. Figure 3-5 is a photograph of the Block I S-band TWM package, and Fig. 3-6 is a block diagram of the Block I S-band subsystem.

Operations of these Block I TWM subsystems at DSS-11 (Pioneer Site at Goldstone, CA), DSS-12 (Echo Site at Goldstone, CA), DSS-41 (Woomera, Australia), DSS-42 (near Canberra, Australia), and DSS-51 (near Johannesburg, South Africa) supported the successful Mariner IV Mars encounter in July 1965. The play-back sequence of our first close-up photographs of the Martian surface returned one picture every eight hours. The data rate was 8.33 bits per second. These five TWM/CCR subsystems all performed without failure during the two-month picture playback period. A sixth Block I TWM subsystem was installed at DSS-61 in Robledo (near Madrid), Spain.

Eight more TWM/CCR systems using AIL TWMs and ADL CCRs were added to the DSN between 1965 and 1968. The eight new systems were identified as Block II S-band TWMs. The S-band Block II TWMs were modified to support the planned Manned-Space-Flight-Network (MSFN) frequencies (2270 MHz to 2290 MHz) as well as the Deep-Space-to-Earth frequency allocation (2290 MHz to 2300 MHz). Future missions with additional requirements, and the limited performance and life-times of the ADL CCRs showed the need for further TWM and CCR development work.

The business environment of that time affected JPL's and NASA's TWM and CCR development approach. Future users of TWMs in the commercial sector seemed limited. Earth-orbiting satellite communications networks did not need very low noise receiving systems in the many Earth-based (ground) installations. Earth-orbiting satellites, sending information to ground stations, could put more transmitter power on the satellite and eliminate the need for a multitude of cryogenically-cooled receivers on the ground. Ambient temperature preamplifiers for the ground-based receivers would have adequate sensitivity, and they would still cost less, be less expensive to maintain, and be more reliable than cooled-receiver preamplifiers.

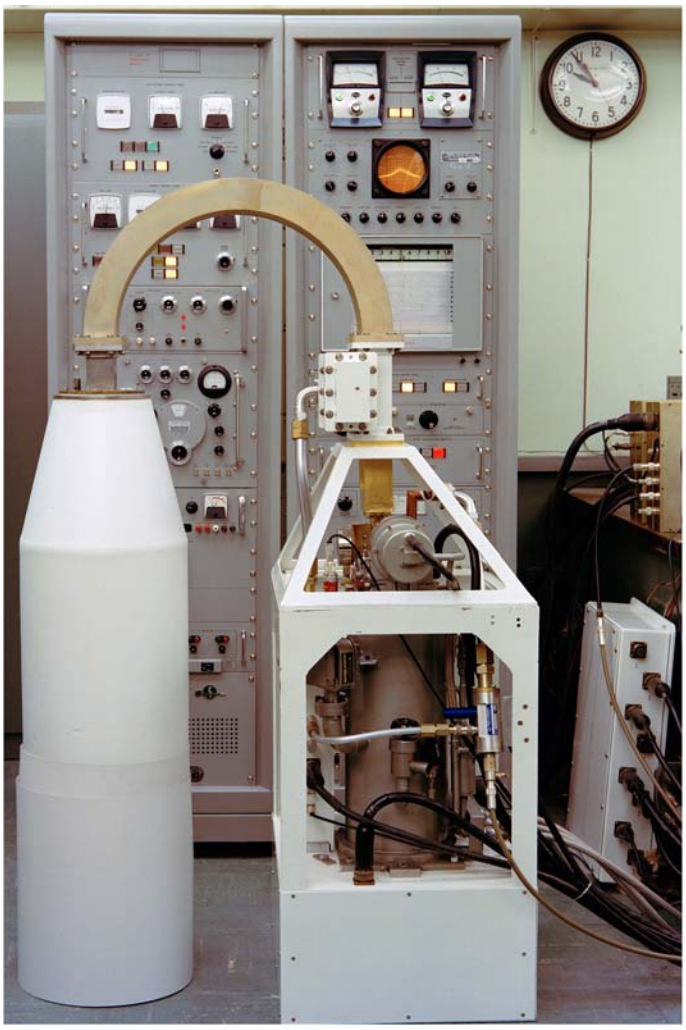


Fig. 3-5. Block I S-band TWM package during laboratory evaluation.

Most of the organizations and specialized personnel that had developed maser technology seemed anxious to move on to the optical spectrum. Light amplification by stimulated emission of radiation (laser) development work attracted many of the specialists who had worked on maser development. Those who continued developing masers after 1965 were, for the most part, concerned with deep-space communications or radio astronomy.

The success and experience with the Block I S-band masers during the Mariner IV mission to Mars showed the potential for future use of TWMs in the DSN. TWM/CCR subsystems with increased performance and reliability would

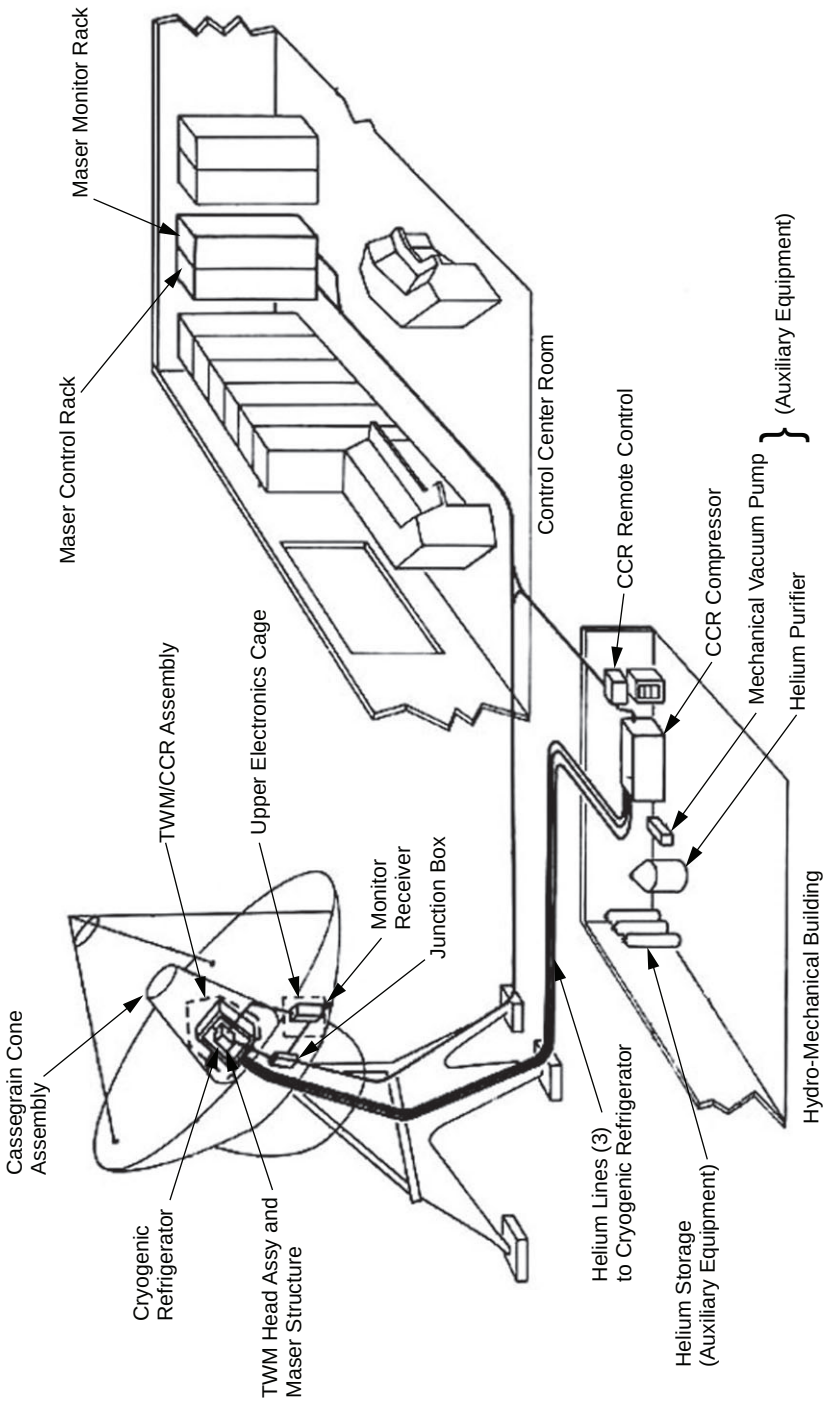


Fig. 3-6. Block I S-band subsystem block diagram.

be needed in the DSN. Future missions would utilize near-real-time video transmission from deep space to Earth. Data rates as high as 117,000 bits per second (kbps) would be used from a distance of about 5 A.U.

Dr. Walter Higa's and Ervin Wiebe's work on CCR development was aided by Professor William Gifford, a co-inventor of the Gifford McMahon (GM) cycle. Their work resulted in a JPL design for a CCR with a simplified cool-down procedure, a quicker cool-down time, and a mean time between failures (MTBF) that was increased from the previous 1500 hours to 3000 hours. Beginning in 1966, the first of these new CCRs was operated on Goldstone's 64-m antenna at DSS-14 for 20,000 hours without maintenance or failure. More about these CCRs will follow in the Cryogenic Refrigeration Systems chapter.

The development of new TWMs with significant performance improvement goals began in 1965. This work took advantage of the previously learned lessons and the knowledge obtained from the material published so generously by the TWM experts at BTL and AIL. New technology developed by the JPL maser team resulted in several patents [39–53]. S-band TWMs tunable from 2275 MHz to 2415 MHz were developed for use on the 26-m antenna at DSS-13 and on the 64-m antenna at DSS-14. S-band Block III and Block IV TWMs, covering 2265 MHz to 2305 MHz, X-band TWMs operating across a tuning range extending from 7600 MHz to 8900 MHz, and a Ku-band TWM tunable from 14.3 GHz to 16.3 GHz were also developed between 1965 and 1973.

Goals of the TWM development work at JPL were to maximize the TWM gain and bandwidth while minimizing the noise temperatures. Comb-type slow-wave structure loss (called forward loss) reduces the TWM gain and raises the noise temperature. High reverse loss, accomplished by resonance isolators in the slow-wave structure is needed for stable, non-regenerative amplification.

Initial attempts to produce copper comb-type slow-wave structures without joints in the regions of high-intensity RF magnetic fields were unsuccessful. The use of milling machines with various types of cutters (including saws and end-mills) did not yield sufficient precision. Experienced and knowledgeable machinists in the JPL machine shop developed techniques using a shaper and extremely hard carbide-alloy cutting tools to cut groves into copper with the needed precision. Electric discharge machining (EDM) was used to create a comb structure from the ribs of copper left between the grooves.

The first of a series of copper TWM comb-structures produced with this technique is shown in Fig. 3-7. Two side-by-side 6-inch (15.24-cm) long comb structures are separated by a center divider. A copper cover (not shown) was attached with machine screws and used an indium gasket to provide intimate contact with the center divider. Coaxial lines were used to couple the signal into and out of the comb structures. A wire loop was used at the end of the structures (opposite the flange) to couple the signal from one comb structure to the other. Pump energy entered from a waveguide through a rectangular hole in

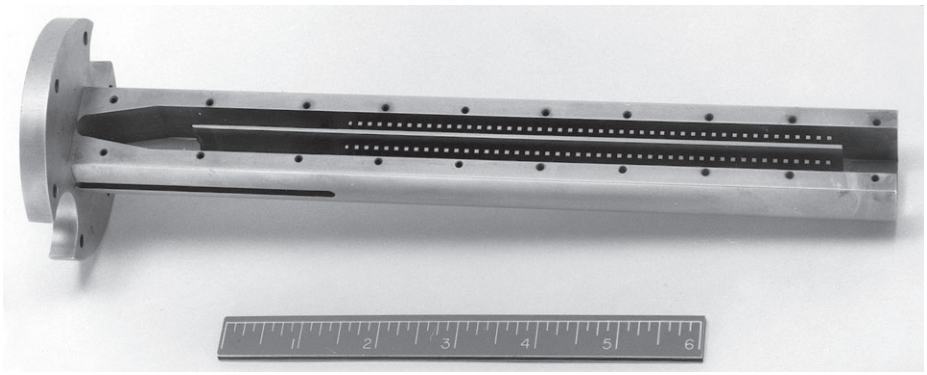


Fig. 3-7. One-piece S-band TWM comb structure (ruler for scale is in inches).

the center of the flange. The 7.26-cm (3-inch) diameter flange was bolted to the 4.5-K station of the CCR. An indium gasket between the TWM flange and the 4.5-K station provided adequate conduction.

These TWMs operated in the vacuum environment of CCRs, and the 50-mW to 100-mW heating caused by the maser pump energy was transferred from the TWM structure to the liquid helium in the 4.5-K station. The temperature difference caused by conduction through the 0.127-mm (0.005-inch) indium gasket was less than 0.01 K. The temperature difference caused by conduction through the copper maser body was about 0.1 K. The two rubies in the center of the structure were a snug fit between the center divider and the comb fingers. The two outer ruby bars were pressed against the comb structure fingers by spring tension achieved with contoured copper-plated beryllium-copper shims. The force of the outer rubies against the comb fingers caused the comb fingers to put pressure on the inner rubies, insuring intimate contact between all rubies and copper surfaces.

Properly oriented ruby of high quality was needed for maximum TWM gain. Inspection of the ruby bars was accomplished with polarized light. A ruby to be inspected was mounted between cross-polarized sheets. The crystalline structure flaws in a slab of ruby cut from a disc boule grown by the flame-fusion process are shown in Fig. 3-8. A ruby slab cut from a boule produced by the Czochralski process shows no crystalline structure flaws. The ruby was "pulled-from-melt" in the Czochralski process, and the crucible containing the molten ruby was made of iridium. The Crystal Products Division of the Union Carbide Corporation perfected the Czochralski process to grow large ruby boules of superb quality.

Early S-band TWMs and later X-band and Ku-band TWMs used polycrystalline yttrium-iron-garnet (YIG) resonance isolators. The resonant frequency of the YIG elements was determined by the magnetic field needed for ruby spin system resonance at the desired signal frequency. The shape of the YIG elements was adjusted to provide a demagnetizing factor that resulted in

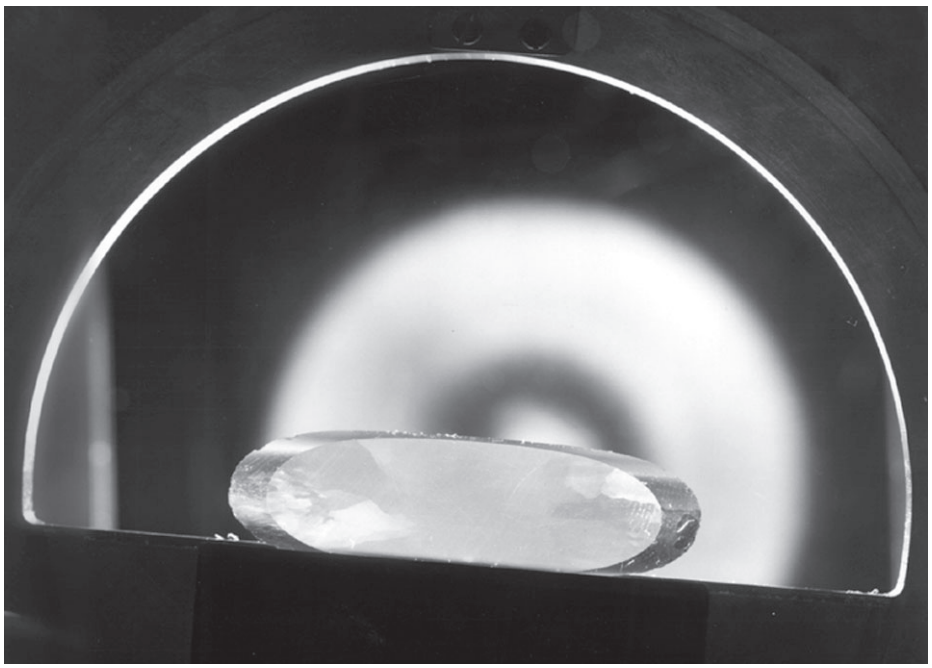


Fig. 3-8. Disc Boule ruby slab.

an internal magnetic field giving YIG resonance at the signal frequency. The internal field in a piece of YIG is determined by the saturation magnetization of the material, as well as the shape. The saturation magnetization of polycrystalline YIG is about 1750 G at room temperature and increases to a value between 2400 G and 2500 G at 4.5 K. Variations in the saturation magnetization occur, depending on the source of the YIG. The resonant frequency was also found to be dependent upon the material to which the YIG was mounted, and the adhesive used for mounting. YIG mounted directly on copper required a thickness-to-diameter ratio of 1/10 rather than 1/8 when mounted on alumina or ruby for the S-band TWMs. The difference in these thickness-to-diameter ratios shifts the YIG resonant frequency by about 400 MHz when the applied magnetic field is 2450 G.

The line width of polycrystalline YIG at 4.5 K is about 400 MHz, much wider than needed for a maser with less than 50 MHz of bandwidth. The YIG elements could be disks, or have rectangular shapes; either geometry could be made to work well. Single-crystal YIG isolators were developed for Block III and Block IV S-band masers where the regions of circular polarization in the SWS were relatively small. The volume of single-crystal YIG needed to provide adequate isolation was about 1/3 the volume of polycrystalline YIG. This reduced the forward isolator loss because the smaller pieces more closely matched the small regions of circular RF magnetic field polarization. The isolator performance improvement came at the cost of tighter tolerances. The

thickness tolerance for each piece of YIG was about ± 0.0008 mm (0.000031 inch); the measurement of dimensions was not suitable for the selection process. Doug Hofhine, a talented and patient technician, sorted and selected several thousand YIG pieces by measuring the resonant frequency of each piece in the appropriate magnetic field.

The TWM designs were optimized by using short test structures to measure and analyze the performance as a function of various changes. This cut-and-try optimization process was very efficient in terms of time and cost. Figure 3-9 shows an S-band TWM test-structure, isolators, and rubies. The comb-fingers were shaped to maximize the filling factor. Optimum isolator size and position were determined experimentally. The performance of the 7.62-cm (3-in.) test structure showed that a 30.48-cm (12-in.) slow-wave structure would give the desired gain and bandwidth. Two side-by-side structures were used to efficiently package the needed length.

The S-band TWM filling factor is at a maximum when the RF magnetic-field polarization is linear along the direction of the ruby c-axis, the direction of propagation in the case of these S-band TWMs. Large regions of linear polarization were created by the finger dimensions and spacing in the Block III S-band TWM. The optimum geometry for maximum gain leaves very small regions of circular polarization in the slow-wave structure.

The regions of circular polarization in comb-type slow-wave structures are located between the fingers, as shown in Fig. 3-10. The optimum distance from the centerline of the row of fingers towards the outside wall of the structure varies as a function of frequency.

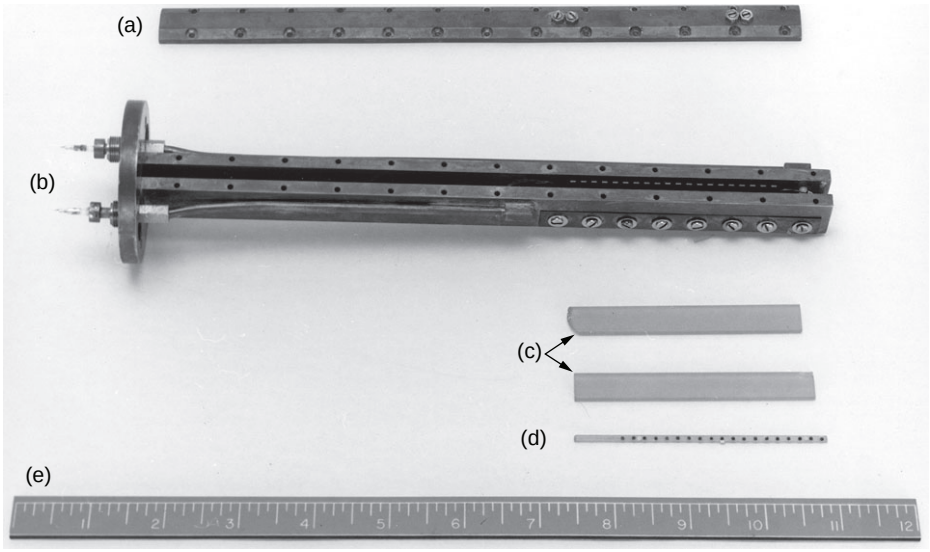


Fig. 3-9. S-band TWM test structure: (a) cover; (b) TWM test structure; (c) rubies; (d) isolator; (e) ruler for scale (inches).

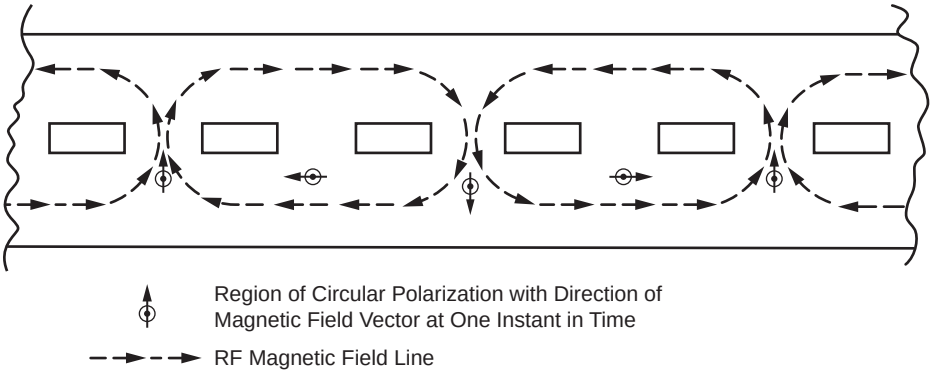


Fig. 3-10. Regions of circular polarization in a Block III S-band TWM. (Simplified RF magnetic field pattern as visualized around the base of the comb structure fingers).

At frequencies in the lower part of the slow-wave structure pass-band, the RF magnetic fields are most linear in the direction parallel to the direction of propagation. The signal phase shift from finger-to-finger approaches zero when approaching the comb-structure's low-frequency cut-off. At frequencies in the higher frequency part of the slow-wave structure pass-band, the RF magnetic fields are most linear in the direction perpendicular to the direction of propagation. The signal phase shift from finger-to-finger approaches 180 deg when approaching the comb-structure's high-frequency cut-off. The rotational sense of the circularly polarized field in the slow-wave structure is right-hand circularly polarized (RCP) on one side of the comb fingers and left-hand circularly polarized (LCP) on the other side of the comb fingers. The direction of propagation determines which side is RCP and which side is LCP.

These TWMs, using Czochralski ruby and single-crystal YIG isolators, were called Block III S-band TWMs, and 27 were built by 1973 for implementation into the DSN and NASA's S-band Tracking and Data Network (STDN) [13]. The Block III S-band TWM used a permanent magnet to supply a 2450-G (0.245-T) magnetic field. The magnet weighed about 82 kg (180 lb), and the complete maser package weighed about 204 kg (450 lb). The maser package was similar in size and shape to the Block I TWM package shown in Fig. 3-5.

A Block III S-band TWM structure was modified and used to produce a prototype of the Block IV S-band TWM. A superconducting magnet provided the 2450 G magnetic field and a short, cryogenically-cooled input transmission line reduced the noise temperature of the maser at the ambient interface to 2.1 K. The Block IV S-band TWM prototype is shown in Fig. 3-11 without the superconducting magnet. Part of the radiation shield and vacuum housing are removed to show the internal parts. This super-low-noise research and development (R&D) maser was used to (1) enable the return of Mariner 10 real-time pictures of Mercury's surface in 1974, (2) support the Inter-Cometary

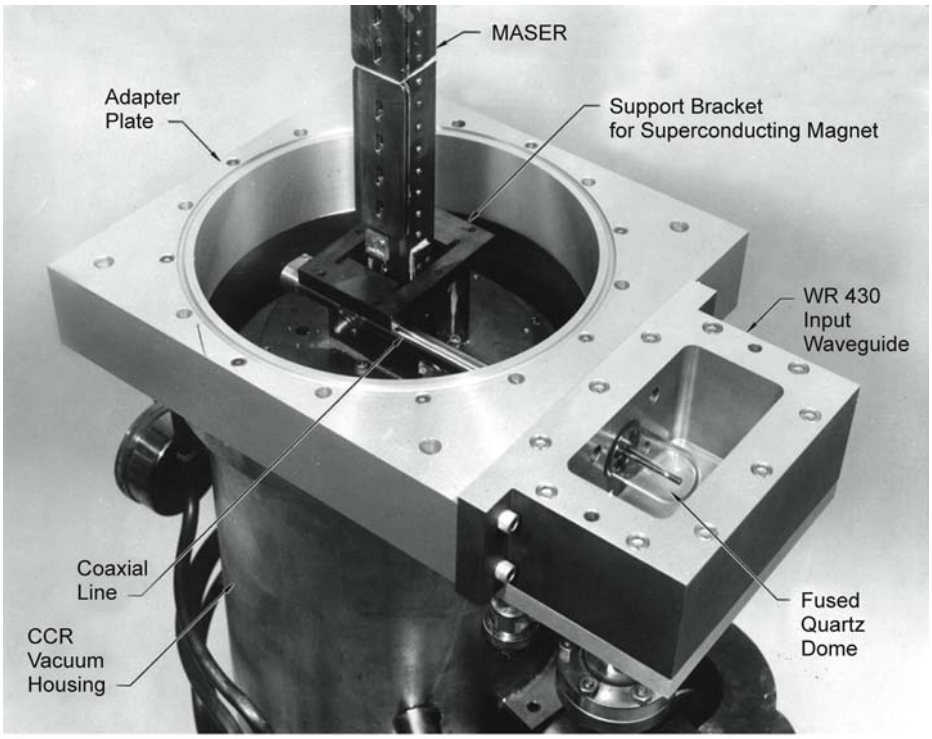


Fig. 3-11. Block IV S-band TWM prototype without superconducting magnet.

Explorer (ICE) mission at the Usuda beam-waveguide antenna in Japan [14], (3) provide radio-science support for Voyager at Usuda, and (4) support the Galileo mission at DSS-43 [15].

An S-band TWM using a half-wavelength comb structure and a box-type superconducting magnet was developed in 1985 to support the International Cometary Explorer (ICE) mission at DSS-14 (Goldstone) [14]. It had a wider instantaneous bandwidth and greater tunability than the Block III and Block IV S-band TWMs. The design was based on the development of X-band Block II TWMs between 1976 and 1980.

The development of X-band ruby masers for the DSN began long before Solar System exploration missions used the X-band deep-space-to-Earth allocation (8400 MHz to 8450 MHz). A cascaded-cavity type 8450-MHz maser was purchased from the Hughes Aircraft Company and used on the 10-m antenna at DSS-13 from 1964 to 1966 [16]. A comb-type TWM tunable from 8370 MHz to 8520 MHz was developed at JPL and used on the 64-m antenna at DSS-14 [17]. These early X-band TWMs used the 54.7-deg "push-pull" ruby orientation. The 3650-G magnetic field requirement for this orientation at X-band was provided by permanent magnets. The 5000-G field requirement for the 90-deg ruby orientation at X-band was met later with a superconducting

magnet design developed earlier for a Ku-band maser at 15.3 GHz [18]. The Ku-band maser, tunable from 14.3 GHz to 16.3 GHz was developed for use on the 64-m antenna at DSS-14 in anticipation of a deep-space-to-Earth allocation at 15.3 GHz [19].

The initial Ku-band maser plan included the possibility of using the 54.7-deg "push-pull" ruby orientation for the Ku-band TWM. Investigation of alternative maser materials was also in the plan. The investigation and tests included several materials including rutile, zinc tungstate, and emerald, with various doping agents. The results for ruby were better than the other materials [20]. Various orientations of ruby were tested with emphasis on the "push-pull" angle using the 2–3 transition for the signal and pumping the 1–3 and 2–4 transitions. These two-pump transitions are at the same frequency when the angle between the c-axis and the applied magnetic field is 54.7 deg (the energy levels are symmetrical). Tests of the push-pull pumping technique with separate pump frequencies at the 90-deg orientation showed performance inferior to the 54.7-deg orientation, but tests of push-push pumping at the 90-deg orientation (with the signal transition being the 1–2 level, and pumping the 1– 3, and 3– 4 levels) gave the best performance. The need for a higher magnetic field at this orientation was easily satisfied by use of the superconducting magnet. The superconducting magnet design for the Ku-band maser was scaled for use with X-band TWMs and the K-band (18 GHz to 26 GHz) reflected-wave masers (RWMs) developed for radio astronomy applications [21].

R&D X-band TWMs using the 90-deg ruby orientation were developed to cover tuning ranges of 7600 MHz to 8900 MHz and 7750 MHz to 8750 MHz. These TWMs were used on the 26-m antenna at DSS-13 and on the 64-m antenna at DSS-14. These TWMs served as prototypes for the Block I X-band TWM implemented to support the Viking missions in 1975 and the Voyager missions in 1977. The Block I X-band TWMs produced in excess of 45 dB gain, a noise temperature of 8 K, and an instantaneous bandwidth of more than 50 MHz, easily covering the 8400 MHz to 8440 MHz requirement [22]. The TWM with field shaping coils is shown mounted on a refrigerator in Fig. 3-12. The maser gain and bandwidth plots at five gain-control field-shaping coil settings are shown in Fig. 3-13.

These Block I X-band TWMs use push-push pumping at two frequencies near 19.2 GHz and 24.05 GHz. The Block I X-band pump klystron package with a power combiner and frequency modulator is shown in Fig. 3-14. The klystrons are frequency modulated at a rate near 100 kHz to cover the bandwidth needed. The ruby spin-relaxation times are sufficiently long to prevent the 100-kHz modulation from being transferred to the signal frequency. Each klystron is connected to a terminated circulator, used as an isolator. A waveguide that passes 24 GHz, but is cut-off at 19 GHz, passes the 24-GHz pump energy to a third three-port circulator used as a power combiner. The 19-GHz klystron is connected to the power combiner circulator at the previous

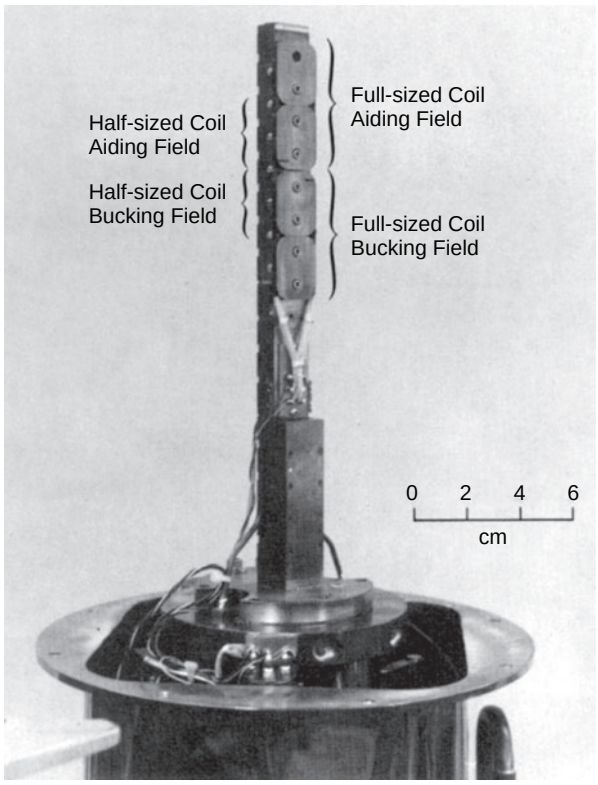


Fig. 3-12. Block-I X-band TWM with field-shaping coils mounted on refrigerator.

port, before the connection from the 24-GHz klystron. The 19-GHz pump power is reflected from the beyond-cutoff waveguide and sent on to the third port, along with the 24-GHz pump power.

The Voyager mission plans were for the two spacecraft to encounter Saturn in November 1980 and in August 1981, and for Voyager 2 to encounter Uranus in January 1986 and Neptune in August 1989. The distances from Earth would be 10 A.U. at Saturn, 19 A.U. at Uranus, and 30 A.U. at Neptune. The 115,200 bps data rate at a 5 A.U. from Jupiter translates into data rates of about 28,800 bps from Saturn, 8000 bps from Uranus, and 3200 bps from Neptune due to the square-of-the-distance penalty. Many enhancements were needed and implemented to enable the data rates of 44,800 bps from Saturn, 29,900 bps from Uranus, and 21,600 bps from Neptune that were achieved. One of the enhancements was a reduction in the system operating noise temperature of the DSN 64-m, and later 70-m, antennas. The Block II X-band masers were developed between 1976 and 1980 to contribute to this noise reduction.

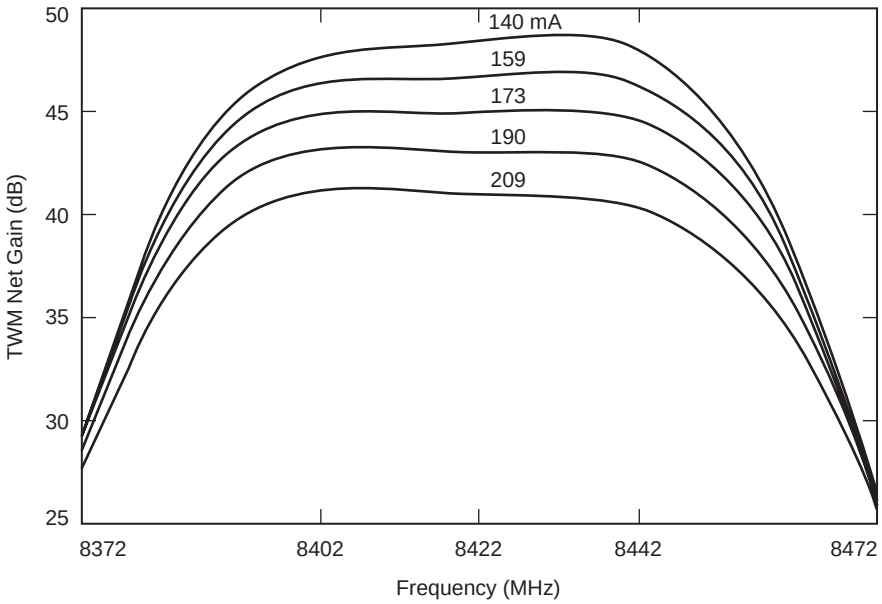


Fig. 3-13. Block-I X-band TWM gain versus frequency at five gain-control settings.

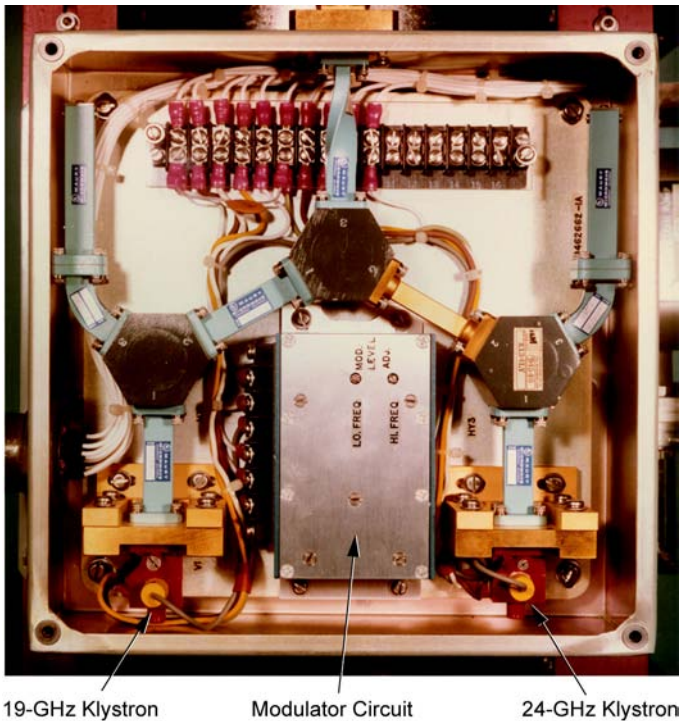


Fig. 3-14. Block-I X-band TWM pump package.

In 1980 D. L. Trowbridge reported, "Four X-band traveling-wave maser (TWM) systems with effective input noise temperatures of 3.5 K and bandwidths varying from 65 to 108 MHz have been supplied to the Deep Space Network [23]. These TWMs are used on the 64-m antennas at Deep Space Stations 14, 43, and 63 at 8420 MHz to meet the requirements of the Voyager Saturn encounter. The TWMs use shortened and cooled signal input waveguide to reduce noise and are equipped with superconducting magnets and solid state pump sources to provide the required stability performance."

These Block II X-band TWMs were the first in a series of DSN TWMs to use ½-wavelength comb elements attached to a ruby slab in the style of a printed circuit. A goal of this design was to simplify the machining process needed to fabricate slow-wave structures with the needed precision. A new box-type of superconducting magnet was developed to accommodate a TWM volume sufficient to produce a greater gain-bandwidth product. A new staggered-height resonant isolator was developed to provide the needed bandwidth with reduced forward loss [24]. The "insides" of the TWM are shown in Fig. 3-15. The Block II TWM assembly, mounted in the superconducting magnet on the 4.5-K refrigerator is shown in Fig. 3-16. The low-noise input waveguide assembly and vacuum window are shown connected to the TWM and superconducting magnet assembly in Fig. 3-17.

Later modifications achieved a 100-MHz bandwidth for all Block II-A TWMs through the addition of a post amplifier and a gain-control attenuator.

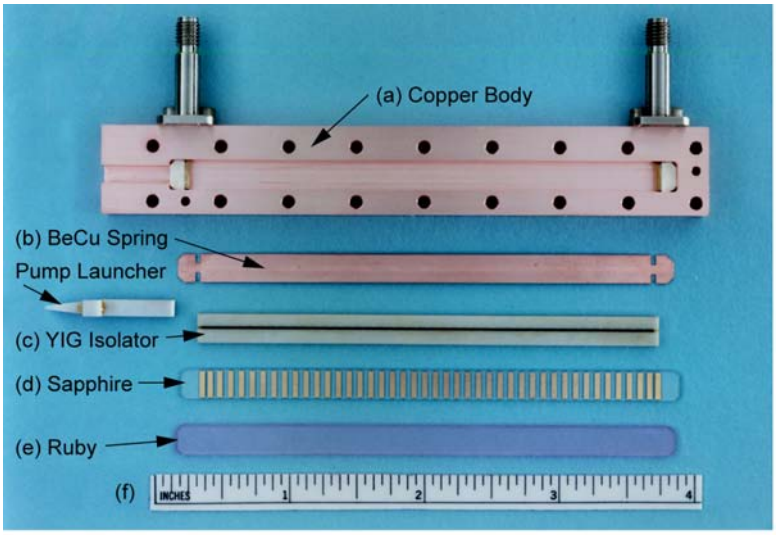


Fig. 3-15. Block III X-band TWM half-wave stripline comb structure components: (a) copper body; (b) copper-beryllium spring; (c) YIG isolator; (d) sapphire with slow-wave structure (SWS); (e) ruby; and (f) ruler for scale (in inches).

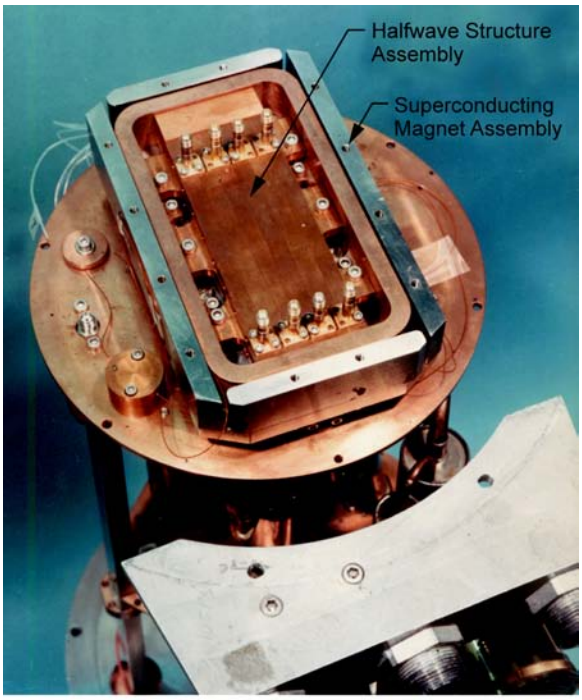


Fig. 3-16. Block II X-band TWM and superconducting magnet on 4.5-K refrigerator.

Further modifications to a Block II-A X-band TWM allowed the measurement of a total system temperature of 6.6 K with a cooled feedhorn pointed at zenith [25]. A 1.3-K effective input noise temperature was measured at the ambient interface to the TWM package. The modified Block II-A TWM was cooled to 1.6 K in superfluid helium. The X-band feedhorn was cooled, with a temperature profile that varied from about 1.6 K in the helium bath to about 300 K at the ambient interface. The ultra-low-noise X-band TWM was named the X-band ULNA. Further modifications to reduce the helium consumption rate removed the larger parts of the feedhorn from the dewar, increasing the effective input noise temperature to about 1.9 K. The X-band ULNA is used in DSS-13's Beam-Waveguide antenna to support planetary radar. The total system operating noise temperature at zenith is about 14 K in clear dry weather.

3.10 Reflected-Wave Masers

The investigation and optimization of push-push pumping techniques at 19 GHz and 24 GHz for future use with wide-bandwidth X-band TWMs was done in a well-matched ruby-filled waveguide in 1974. The well-matched ruby-filled waveguide was also tested as a reflection type amplifier in the 18- to 26.5-GHz range at temperatures between 1.9 K and 4.4 K [26]. The 0.178 cm

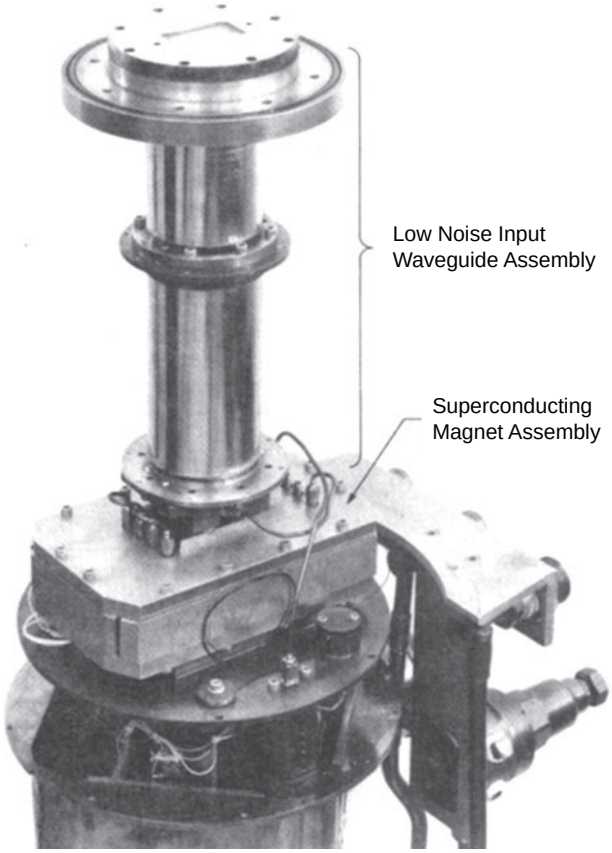


Fig. 3-17. Block-II X-band TWM with input waveguide assembly.

high by 0.356 cm by 5.08 cm (2 in.) long ruby filled section was matched to a standard WR42 (18-GHz to 26.5-GHz) waveguide with a 7.62-cm long tapered matching section. Magnetic shims were added to broaden the ruby linewidth and reduce the gain of the amplifying section to avoid regeneration. Gain values between 7 dB and 7.7 dB were observed across a 285-MHz bandwidth near 24.4 GHz at 4.4 K. Gain values between 15 dB and 17 dB were observed across a 250-MHz bandwidth near 24 GHz at 1.9 K. The ruby-filled waveguide, combined with a circulator and a pump source, formed a simple amplifier stage that could be tuned electronically across the entire 18-GHz to 26.5-GHz range. Several such stages used in series could achieve high gain with wide bandwidth.

A reflected-wave maser (RWM) was developed for use between 9 and 10 GHz with the support of grants from the California Institute of Technology President's Fund and the National Science Foundation. The maser group at JPL, and Professor Sheldon Schultz and Larry Flesner of the Physics

Department at the University of California, San Diego (UCSD), in La Jolla, California worked together on this task [27]. This maser improved the sensitivity of a transmission electron spin resonance spectrometer used in the UCSD Physics Laboratory.

A technology sharing contract between JPL and the National Radio Astronomy Observatory of the Associated Universities Inc. (NRAO, AUI) resulted in the development of "A Reflected-Wave Ruby Maser with K-Band Tuning Range and Large Instantaneous Bandwidth" [21]. This first K-band RWM was used on the NRAO 43-m (140-foot) diameter antenna at Greenbank, West Virginia. The effective input noise temperature of the maser-receiver was 13 K ±2 K, and the instantaneous bandwidth was 250 MHz near the center of the tuning range.

Craig Moore of NRAO, who had worked at JPL with the maser group to develop this RWM, later improved the performance of the K-band RWM, increasing the instantaneous bandwidth to more than 500 MHz. This bandwidth improvement was accomplished with a superconducting magnet that provided the needed magnetic field shape across the entire 15.24-cm (6-in.) length of the RWM structure [28]. Craig Moore shared his new magnet design with JPL, and subsequent K-band RWMs were built for, and in cooperation with:

1) Caltech's Owens Valley Radio Observatory,
2) Professor David Wilkinson and Dale Fixsen of Princeton University and Sam Gulkis of JPL for a balloon-borne maser measurement of the large-scale anisotropy of the cosmic background radiation, and
3) The Max Planck Institute for Radio Astronomy.

K-band RWMs were also used on the DSN's 70-m antennas to support radio astronomy applications. Additional maser-technology sharing contracts enabled work on masers for radio and radar astronomy for the Arecibo radio telescope in Puerto Rico and for the Haystack Observatory in Massachusetts.

An RWM was developed for the 32-GHz applications [29]. The Ka-band RWM used eight channels of ruby filled waveguide and a box-type 1.18-T (11,800-G) superconducting magnet, similar to the geometry of the magnet developed for the Block II X-band masers. The Ka-band RWM was cooled by a closed-cycle helium refrigerator like those used for other masers in the DSN. The maximum bandwidth achieved was 400 MHz with 21 dB of net gain. Four impact avalanche and transit time (IMPATT) oscillators, each providing about 100 mW near 66.4 GHz, were used to pump the maser. A pump power of 400 mW was not adequate to pump and saturate the relatively large volume of ruby across a 400-MHz bandwidth. Reducing the pump frequency modulation bandwidth increased the net gain to 40 dB with a 150-MHz bandwidth. At 40 dB gain, the Ka-band RWM demonstrated an effective input noise temperature of about 13 K at the ambient waveguide interface.

3.11 Ka-Band and the Return to Cavity Masers

Establishment of Ka-band allocations for deep-space-to-Earth (downlink) telecommunications from 31.8 GHz to 32. 3 GHz and Earth-to-deep-space (uplink) telecommunications from 34.2 GHz to 34. 7 GHz created 500-MHz wide bands for future Solar System exploration missions. Telecommunications and navigation needs would be served well by DSN antenna systems having high efficiency and low-noise performance at Ka-band. Low-noise amplifier development at Ka-band began in the 1980s.

Attempts to develop ¼ wavelength and ½ wavelength comb-type traveling-wave masers at 32 GHz were not successful. The earlier comb-type slow-wave structures developed for S-band, X-band, and Ku-band showed an increasing demand for precise dimensions as the frequency increased. It was not a surprise that comb structure tolerances would be very difficult and possibly beyond our ability to meet.

The experience with comb-type slow-wave structures and the RWM at 32 GHz convinced us of the need to find a better approach for building Ka-band masers. The Mars Observer spacecraft would use the X-band, near 8420 MHz, for downlink telemetry from Mars. A simple, low-power Ka-band transmitter was planned for a Ka-Band Link Experiment (KaBLE) between Mars Observer and the new 34-m beam-waveguide DSN research antenna at DSS-13. The KaBLE downlink was created by multiplying the X-band downlink frequency by a factor of 4, producing signals near 33.68 GHz. The Mars Observer downlink frequency would shift across a 5-MHz range, from below 8420 MHz to above 8420 MHz, depending on the spacecraft mode. The spacecraft mode would be dependent upon the use of an uplink for the two-way mode, or it would automatically switch to the one-way mode when the spacecraft was not locked to an uplink. The Ka-band maser would need to cover a bandwidth somewhat greater than 40 MHz, centered at 33.68 GHz.

A two-cavity maser design was selected and developed to support the Mars Observer KaBLE at frequencies near 33.68 GHz [5]. Figure 3-18 shows a schematic diagram of the dual-cavity Ka-band maser. Each cavity was composed of a piece of ruby with an electrical length of 1.5 wavelengths in a 2.54 mm × 1.27 mm copper waveguide. The physical length of the ruby is about 3.8 mm. Each ruby extends about 0.3 mm into a reduced height WR30 signal waveguide that connects to a circulator. The circulators are commercially available WR28 waveguide units. Impedance transformers using 1/4 wavelength steps are used to match the WR28 to the reduced height WR30 waveguides. The external and loaded *Q* of each ruby-filled cavity is determined by the length of the ruby extension into the reduced-height WR30 signal waveguide. 69.5-GHz pump energy is supplied through the 2.54 mm × 1.27 mm waveguide connected to the ruby-filled section at the end that is opposite the signal waveguide connection.

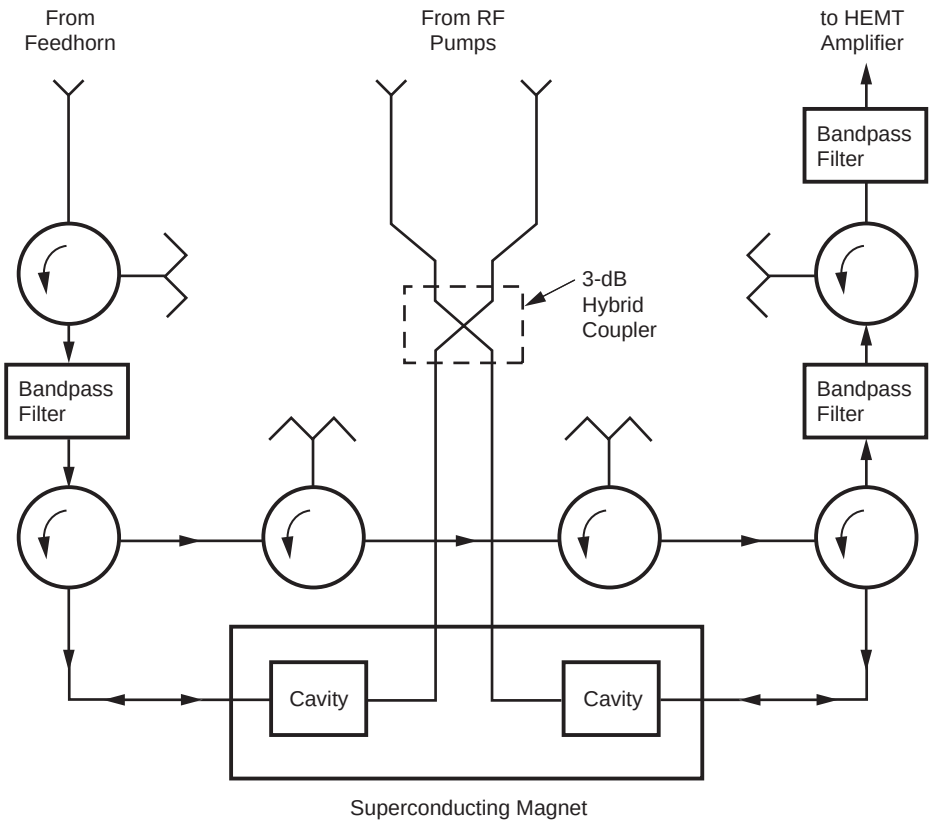


Fig. 3-18. Dual-cavity Ka-band maser schematic diagram.

The ruby-filled cavities were mounted inside a 1.24-T (12, 400-G) circular magnet, somewhat different in shape than the box-type magnet used for the 32-GHz RWM. The roughly square shape of the dual-cavity housing could have used a square superconducting coil, but a slightly larger circular coil was chosen to simplify the winding process. The niobium-titanium superconducting wire, surrounded by an iron alloy with high permeability (Hyperco-27), was similar to previously used box magnets in the respect that these geometries do not depend upon the use of additional superconducting shielding (usually niobium-tin) to prevent straying magnetic fields that destroy the magnetic-field uniformity. The feedhorn, connecting waveguide assemblies, and cryogenically cooled HEMT follow-up amplifiers were all located in the helium dewar with the maser.

The concentric liquid-helium double dewar was purchased from International Cryogenics, Inc. (Indianapolis, Indiana). Reduction of the vapor pressure above the liquid helium in the 22-liter inner dewar reduced the temperature of the super-fluid helium to 1.5 K. The vapors were used to cool

the feedhorn. A capillary tube assembly from the outer 80-liter 4.2-K dewar was designed to allow for continuous replenishment of the 1.5-K super-fluid helium. This configuration allowed for refilling helium without interrupting maser operation.

Two Hughes 47134H-1110 IMPATT Oscillators each provided 100-mW pump power near 69.5 GHz. These were frequency modulated across a 170-MHz bandwidth at a rate of 20 kHz. The electronic gain of the maser is 28 dB, the net gain is 25 dB, the inversion ratio is 2.0, and the –3 dB bandwidth is 85 MHz. The gain (with pumps on) and ruby absorption (with pumps off) is shown as a function of frequency in Fig. 3-19. The gain can be adjusted for higher or lower values with trim coils that change the magnetic field gradient, thereby artificially increasing or decreasing the ruby line-width and the maser's bandwidth.

A 5 K ±1 K effective input noise temperature was measured at the maser package feedhorn aperture. The uncertainty is due to the tolerance on the sky brightness temperature during the measurement. Surface temperature and relative humidity measurements were used to model the atmosphere's contribution. The maser's measured effective input noise temperature agrees well with computations of 3 K for the maser in the 1.5 K bath at the input to the first circulator, 0.3 K for cooled waveguide components, and 1.7 K for the feedhorn and Kapton window loss. The effective noise temperature of each cavity is calculated to be 2.2 K at an electronic gain of 14 dB and a net gain of 13.5 dB.

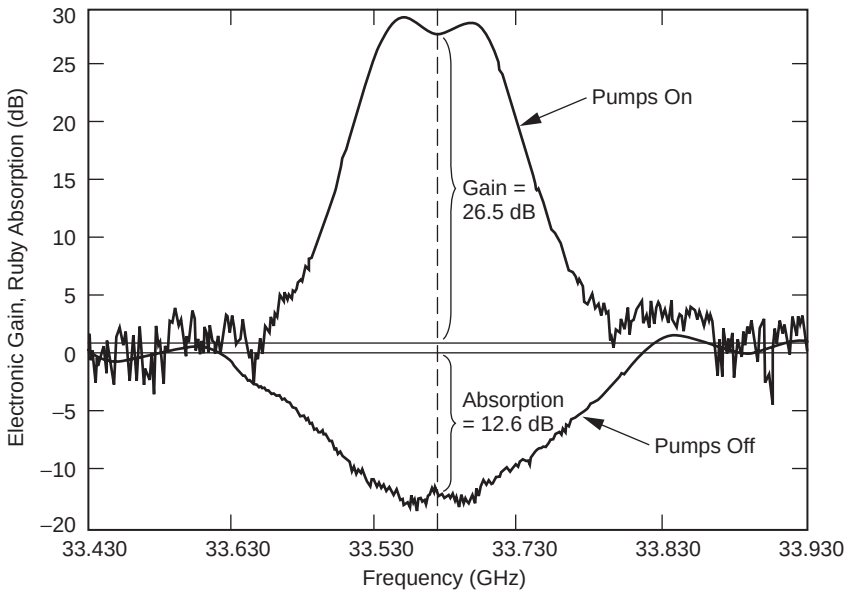


Fig. 3-19. Electronic gain with RF pumps on, and ruby absorption with RF pumps off, for the dual-cavity 33.68-GHz (Ka-band) maser.

3.12 Analysis of Maser Designs

As discussed above, construction of Ka-band traveling-wave masers (TWM) using comb type slow-wave structures (SWS) is not practical. A maser employing several stages of reflection type coupled waveguide cavities or a traveling-wave design using c oupled cavities is c onsidered to be the best approach to a chieve instantaneous bandwidths of 500 MHz. Computer programs capable of modeling the passive characteristics of such structures have existed for some time [30]. However, the authors are not a ware of any program capable of modeling the ruby gain. Development of such a program was undertaken at JPL.

The approach was based on the well-known mode-matching technique. A concise description of the waveguide mode-matching formulation can be found in reference [31]. The results of the program are the microwave scattering parameters. By using a generalized scattering matrix formulation, the interaction between waveguide cavities includes higher order modes if present. Antenna engineers at JPL developed a mode-matching program for the analysis of corrugated feedhorns [32]. This program was the starting point for the maser design program. The effect of the resonant spin system in the ruby was added by appropriately modifying the propagation constants of the waveguide modes used in the mode-matching formulation.

Each waveguide mode propagation constant is given by

$$\gamma^2 = k_c{}^2 - k^2 = k_c{}^2 - \omega^2 \mu_{\text{eff}} \varepsilon_{\text{eff}} \tag{3.12-1}$$

where $k_c{}^2$ is the square of the cut-off wavenumber, determined by the cross sectional dimensions of the waveguide and k^2 is the square of the wave vector magnitude. μ_{eff} and ε_{eff} are the effective scalar permeability and permittivity of the material filling the waveguide, and ω is the angular frequency. The impedance of each H-mode and E-mode used in the mode-matching analysis is also modified according to

$$Z_H = \frac{j\omega\mu_{\text{eff}}}{\gamma} \quad \text{and} \quad Z_E = \frac{\gamma}{j\omega\varepsilon_{\text{eff}}} \tag{3.12-2}$$

In his a nalysis of the traveling wave maser, Siegman [6] calculates the change in the propagation constant due to changes in the material in which the slow wave circuit is embedded. Siegman's result for the change in propagation constant, due to a complex susceptibility in an isotropic magnetic material, is given by

$$\Delta\gamma = \beta - \beta_o - j\alpha = \frac{\omega}{2v_g}(\chi' - j\chi'') \tag{3.12-3}$$

In this equation, β, β_o, α are the phase constant in the presence of magnetic resonance, the phase constant without magnetic resonance, and the attenuation constant with magnetic resonance. The attenuation constant in the absence of magnetic resonance is assumed to be zero, i.e., $\alpha_o = 0$. Siegman's analysis shows that the effect of the ruby spin system on the wave vector of the medium is described by the magnetic susceptibility. Following Siegman, we assume the analytic expression will remain valid within some limits as $\mu, \varepsilon, \gamma, k$ become complex quantities.

Since the magnetic susceptibility of many maser materials (ruby for example) is a tensor quantity, a technique is required to appropriately reduce this tensor quantity to an equivalent complex scalar. (In our program, the permittivity tensor is replaced by an equivalent frequency independent complex scalar. This is simpler than the permeability since no gain or resonant behavior is involved with the permittivity).

Lax and Button investigated the change in the propagation constant of an infinite waveguide of arbitrary cross section loaded with a medium which fills it wholly or partially in the transverse plane, but is uniform in the axial direction [33]. They found,

$$\gamma + \gamma_o^{\,*} = \frac{j\omega\int(\mathbf{H_0}^{*}\cdot\Delta\bar{\bar{\mu}}\cdot\mathbf{H_1}) + (\mathbf{E_0}^{*}\cdot\Delta\bar{\bar{\varepsilon}}\cdot\mathbf{E_1})dS}{\int(\mathbf{E_0}^{*}\times\mathbf{H_1} + \mathbf{E_1}\times\mathbf{H_0}^{*})dS} \tag{3.12-4}$$

where $\Delta\bar{\bar{\mu}}$ and $\Delta\bar{\bar{\varepsilon}}$ is the change in the permeability and permittivity tensors, respectively due to the perturbation. The propagation constants γ and γ_o are with and without the dielectric material. $\mathbf{H_0}$, $\mathbf{H_1}$ and $\mathbf{E_0}$, $\mathbf{E_1}$ are the unperturbed and perturbed microwave magnetic and electric fields, respectively. The integrations are over the waveguide cross-section. Because the waveguide cross section was arbitrary, we are effectively determining the change in the propagation constant due to the change in the wave vector of the medium, albeit surrounded by a metallic enclosure.

In our case, we choose the "unperturbed waveguide" to be that of the ruby filled waveguide without a resonant spin system. Since there is no change in the permittivity, we can set $\Delta\bar{\bar{\varepsilon}} = 0$. Since the ruby spin system is dilute, we can approximate the perturbed fields by the unperturbed fields. Thus, Eq. (3.12-4) simplifies to,

$$\gamma + \gamma_o^* = \frac{j\omega \iint \mathbf{H_0}^* \cdot \overline{\overline{\Delta\mu}} \cdot \mathbf{H_0} ds}{\int (\mathbf{E_0}^* \times \mathbf{H_0} + \mathbf{E_0} \times \mathbf{H_0}^*) ds} \tag{3.12-5}$$

The denominator is four times the power flow down the guide, which is equal to the energy stored times the group velocity. In the numerator $\overline{\overline{\Delta\mu}}$ can be replaced by $\mu_o \overline{\overline{\chi}}$. Thus,

$$\gamma + \gamma_o^* = \frac{j\omega\mu_o \iint \mathbf{H_0}^* \cdot (\overline{\overline{\chi}}) \cdot \mathbf{H_0} ds}{4 v_g \iint \frac{\mu_o \mathbf{H_0}^* \cdot \mathbf{H_0}}{2} ds} \tag{3.12-6}$$

where we have expressed the energy storage in terms of the magnetic field.

For an anisotropic material, the susceptibility tensor can be broken into real and imaginary parts.

$$\overline{\overline{\chi}} = \overline{\overline{\chi'}} - j\overline{\overline{\chi''}} \tag{3.12-7}$$

They are given in terms of the full susceptibility tensor by

$$\overline{\overline{\chi'}} = \frac{1}{2}(\overline{\overline{\chi}} + \overline{\overline{\chi_t}}^*) \quad \text{and} \quad -j\overline{\overline{\chi''}} = \frac{1}{2}(\overline{\overline{\chi}} - \overline{\overline{\chi_t}}^*) \tag{3.12-8}$$

where $\overline{\overline{\chi_t}}^*$ is the hermitian conjugate of $\overline{\overline{\chi}}$.
Therefore, we can write Eq. (3.12-6) as

$$\gamma + \gamma_o^* = \frac{j\omega}{2 v_g} \frac{\iint \mathbf{H_0}^* \cdot (\overline{\overline{\chi'}} - j\overline{\overline{\chi''}}) \cdot \mathbf{H_0} ds}{\iint \mathbf{H_0}^* \cdot \mathbf{H_0} ds} \tag{3.12-9}$$

Equating the imaginary parts of the propagation constants yields

$$(\beta - \beta_o) = \frac{\omega}{2 v_g} \frac{\iint \mathbf{H_0}^* \cdot (\overline{\overline{\chi'}}) \cdot \mathbf{H_0} ds}{\iint \mathbf{H_0}^* \cdot \mathbf{H_0} ds} \tag{3.12-10}$$

Equating the real parts (assuming $\alpha_o = 0$) yields

$$\alpha = \frac{\omega}{2v_g} \frac{\iint \mathbf{H_0}^* \cdot (\overline{\overline{\chi''}}) \cdot \mathbf{H_0} ds}{\iint \mathbf{H_0}^* \cdot \mathbf{H_0} ds} \tag{3.12-11}$$

Comparing this to Eq. (3.12-3), we find that the effective scalar susceptibility corresponding to the tensor susceptibility is given by

$$\chi'_{\text{eff}} - j\chi''_{\text{eff}} = \frac{\iint \mathbf{H_0}^* \cdot \left(\overline{\overline{\chi'}} - j\overline{\overline{\chi''}}\right) \cdot \mathbf{H_0} ds}{\iint \mathbf{H_0}^* \cdot \mathbf{H_0} ds} \tag{3.12-12}$$

For the case of ruby, the full susceptibility tensor is given by

$$\overline{\overline{\chi}} = \frac{\pi\mu_o}{\text{h}} (\text{g}\beta)^2 \left(n_i - n_j\right) \left\{ h\left(f - f_o\right) - jg\left(f - f_o\right) \right\} \overline{\overline{S}} \tag{3.12-13}$$

where

h = Planck's constant, 6.626×10^{-34} joule-second (J-s)

μ_0 = permeability of free space, 1.26×10^{-6} newtons/(amperes)^2 (N/A^2)

g = g-factor of the free electron 2.0023, (unitless value)

β = Bohr magnetron, 9.27×10^{-24} joules/tesla (J/T)

Also, $(n_i - n_j)$ is the difference in the number of spins per unit volume in states i and j. The line-shape functions describe the variation of the susceptibility away from the resonant frequency f_o. Two line-shape functions are used and are denoted by $g(f - f_o)$ and $h(f - f_o)$, respectively. They are given by

$$h(f - f_o) = \frac{4\left(f_o - f\right)}{\pi\left(\Delta f_L\right)^2 + 4\pi\left(f - f_o\right)^2} \tag{3.12-14}$$

$$g(f - f_o) = \frac{2\Delta f_L}{\pi\left(\Delta f_L\right)^2 + 4\pi\left(f - f_o\right)^2} \tag{3.12-15}$$

The quantity Δf_L is the linewidth of the ruby. The quantity $\overline{\overline{S}}$ is called the spin tensor, and it must be evaluated using the methods of quantum mechanics. The anisotropic nature of $\overline{\overline{\chi}}$ is contained within it. The spin tensor is given by,

$$\overline{\overline{S}} = \vec{S}^{ij}\left(\vec{S}^{ij}\right)^{*}, \tag{3.12-16}$$

where $\vec{S}^{ij}\left(\vec{S}^{ij}\right)^{*}$ is a dyadic product formed using the quantum mechanical spin vector operator. The indices *i*, *j* refer to the quantum states between which a given transition takes place. They can refer to a s ignal, pump or to an idle transition. Using the Dirac bra-ket notation, the states are represented by $< i \mid$ and $\mid j >$. The spin vector operator is given by

$$\begin{aligned} \vec{S}^{ij} &= \langle i|\vec{S}|j\rangle \\ &= \langle i|S_x\hat{x} + S_y\hat{y} + S_z\hat{z}|j\rangle \\ &= \langle i|S_x|j\rangle\hat{x} + \langle i|S_y|j\rangle\hat{y} + \langle i|S_z|j\rangle\hat{z} \end{aligned} \tag{3.12-17}$$

or

$$\vec{S}^{ij} = S_x{}^{ij}\hat{x} + S_y{}^{ij}\hat{y} + S_z{}^{ij}\hat{z} \tag{3.12-18}$$

Dropping the superscript, *ij*, we can write

$$\overline{\overline{S}} = \vec{S}\left(\vec{S}\right)^{*} = \begin{pmatrix} S_x \\ S_y \\ S_z \end{pmatrix}\begin{pmatrix} S_x{}^{*} & S_y{}^{*} & S_z{}^{*} \end{pmatrix} = \begin{pmatrix} S_xS_x{}^{*} & S_xS_y{}^{*} & S_xS_z{}^{*} \\ S_yS_x{}^{*} & S_yS_y{}^{*} & S_yS_z{}^{*} \\ S_zS_x{}^{*} & S_zS_y{}^{*} & S_zS_z{}^{*} \end{pmatrix} \tag{3.12-19}$$

In this form, $\overline{\overline{S}}$ is hermitian. Using Eq. (3.12-8), it can be shown that

$$\overline{\overline{\chi'}} = \frac{\pi\mu_o}{\mathrm{h}}(\mathrm{g}\beta)^2\left(n_i - n_j\right)h\left(f - f_o\right)\overline{\overline{S}} \tag{3.12-20}$$

and

$$\overline{\overline{\chi''}} = \frac{\pi\mu_o}{\mathrm{h}}(\mathrm{g}\beta)^2\left(n_i - n_j\right)g\left(f - f_o\right)\overline{\overline{S}}. \tag{3.12-21}$$

A concise description of the low-lying quantum states, $< i \mid$, $\mid j >$ is made possible through the concept of an "effective spin Hamiltonian". This approach includes such effects as the Zeeman splitting of the quantum states due to applied magnetic fields. It also describes the splitting of energy levels due to the electrostatic field of surrounding atoms. An excellent discussion of this concept can be found in [34]. The eigenvalues of the spin Hamiltonian matrix

are the energies of the discrete quantum states available to the spins. The difference in energies divided by Planck's constant determines the resonant transition frequencies. The eigenvector associated with an eigenvalue is a representation of the quantum state having that energy.

The presence of the crystalline electric fields makes the form of the spin Hamiltonian dependent on the orientation of the coordinate system. Taking ruby as an example, if the crystal c-axis is chosen along the z-direction, then the spin Hamiltonian, H_s, is given by

$$H_s = \mathrm{g}_1 \beta H_z S_z + \mathrm{g}_2 \beta \left(H_x S_x + H_y S_y \right) + \mathrm{D}\left[S_z{}^2 - \frac{1}{3} S(S+1) \right] \qquad (3.12\text{-}22)$$

Here g_1 and g_2 are spectroscopic splitting factors, β is the Bohr magneton, and $\overrightarrow{H_{dc}} = (H_x, H_y H_z)$ is the applied static magnetic field.

Personnel at Bell Telephone Laboratories derived a Hamiltonian for which the z-axis is along the applied static magnetic field [35]. The ruby crystal c-axis is specified by the polar angle, θ, with respect to the dc magnetic field and an azimuthal angle, φ, with respect to the x-axis. Their result is

$$\begin{aligned} H_s = & \left(\mathrm{g}_1 \cos^2\theta + \mathrm{g}_2 \sin^2\theta \right) \beta H_z S_z \\ & + \mathrm{D}\left(\cos^2\theta - 1/2 \sin^2\theta \right) \left[S_z{}^2 - 1/3 S(S+1) \right] \\ & + \mathrm{D}(1/2)(\cos\theta \sin\theta) \left[e^{-i\varphi} \left(S_z S_+ + S_+ S_z \right) + e^{i\varphi} \left(S_z S_- + S_- S_z \right) \right] \\ & + \mathrm{D}(1/4) \sin^2\theta \left(e^{-2i\varphi} S_+{}^2 + e^{2i\varphi} S_-{}^2 \right) \end{aligned} \qquad (3.12\text{-}23)$$

Here $S_+ = S_x + iS_y$, $S_- = S_x - iS_y$, and $i = \sqrt{-1}$. We use the values for the spectroscopic splitting factors $\mathrm{g}_1 = 1.9817$ and $\mathrm{g}_2 = 1.9819$, and the zero-field splitting $\mathrm{D} = -3.8076 \cdot 10^{-17}$ ergs, published by the National Bureau of Standards [36].

The coordinate system appropriate to this Hamiltonian is shown in Fig. 3-20. From the point of view of maser design analysis, it makes more sense to fix the direction of the applied field and let the direction of the ruby c-axis be unrestricted. The result is a more complex expression for the spin Hamiltonian. However, since a digital computer performs the calculation, the additional complexity is not a concern. Eq. (3.12-22) can be shown to be almost exactly equal to Eq. (3.12-23). We have neglected terms involving the difference between g_1 and g_2 because they are nearly equal.

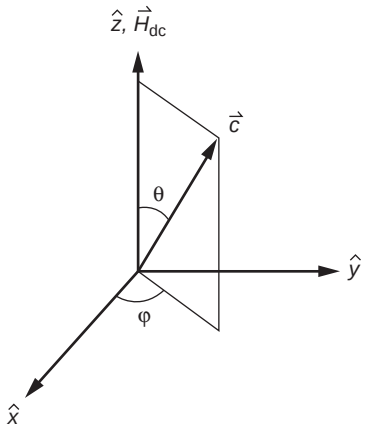


Fig. 3-20. Coordinate system for calculation of ruby spin Hamiltonian.

The components of the spin vector operator, S_x, S_y, S_z are matrices given below. The quantity D represents one half of the "zero-field splitting." This is the initial splitting of the four quantum states of a spin 3/2 chromium ion into two doublets in the absence of a magnetic field. The quantity $S(S+1)$ is the eigenvalue of the operator $S^2 = S_x{}^2 + S_y{}^2 + S_z{}^2$.

In addition to choosing a coordinate system for the spin Hamiltonian, we must choose a representation for the spin operators. This means choosing a set of base states in terms of which the spin quantum states can be expressed. The usual choice for a spin system is the set of states that are simultaneous eigenstates of the total angular momentum squared and the projection of the angular momentum along some axis, usually the z-axis. In this representation, the matrices representing S^2 and S_z are diagonal. We adopt this convention. For a spin $S = 3/2$ system, such as the Cr^{+3} ion in ruby, they are given as 4-by-4 matrices. In particular,

$$S^2 = 15/4 \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad S_z = 1/2 \cdot \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \tag{3.12-24}$$

In this representation, the matrices representing the spin operators S_x and S_y are given by

$$S_x = 1/2 \cdot \begin{bmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{bmatrix} \text{ and}$$

$$S_y = 1/2 \cdot \begin{bmatrix} 0 & -\sqrt{3}i & 0 & 0 \\ \sqrt{3}i & 0 & -2i & 0 \\ 0 & 2i & 0 & -\sqrt{3}i \\ 0 & 0 & \sqrt{3}i & 0 \end{bmatrix} \qquad (3.12\text{-}25)$$

In summary, knowing the spin Hamiltonian given in Eq. (3.12-23), the eigenvectors can be calculated. With the eigenvectors and the spin operators given in Eqs. (3.12-24) and (3.12-25), we can calculate the spin vector given in Eq. (3.12-17). With the spin vector, the spin tensor is calculated using Eq. (3.12-19). With the spin tensor, the real and imaginary parts of the susceptibility tensor are calculated using Eqs. (3.12-20) and (3.12-21). Finally, knowing the RF magnetic field in the regions containing ruby allows the effective scalar susceptibility to be evaluated using Eq. (3.12-12). (The RF magnetic field is calculated from the scattering parameters using the same mode-matching program). This allows the effective magnetic permeability to be calculated using

$$\mu_{\text{eff}} = \mu'_{\text{eff}} - j\mu''_{\text{eff}} = \mu_o \left(1 + \chi'_{\text{eff}} - j\chi''_{\text{eff}}\right) \qquad (3.12\text{-}26)$$

Finally, the effective permeability is used in Eq. (3.12-1) to determine the propagation constant.

The program has given good agreement with experiments. The K-band and Ka-band RWMs were appropriate for analysis by this program, and the analysis results were consistent with the previously measured performance. Designs using waveguide cavities filled with ruby, having the same loaded-Q as the S-band coaxial cavities used in the early JPL cavity masers were analyzed. Again, the results were consistent with the performance measured 40 years before this analysis tool was developed.

A Ka-band waveguide coupled-cavity maser design was optimized by analyzing various combinations of impedance transformers, irises, empty cavities, and a ruby filled cavity [37]. The program analyses these designs but does not generate the design. The design must be produced based on the designer's knowledge and use of other microwave design tools. The process might be described as automated cut-and-try, with no actual fabrication taking

place. When the analysis of any given design shows the desired result, it is then time for fabrication.

Fabrication of the coupled-cavity design was completed, and subsequent RF measurements showed performance that was not as predicted by the analysis. The analysis program was then used to make changes to obtain the same results as produced by the RF measurements. One change was needed in the impedance matching transformer. The height of one step was changed by 0.076 mm, about 2 percent of the 3.556-mm empty waveguide height. Inspection of the structure showed that there had been a machining error of the size and location, in the impedance transformer, just as indicated by the analysis program.

Subsequent Ka-band maser designs were produced with the goal of simplifying the machining process. A cascaded-cavity TWM design was developed [38]. A drawing of the design is shown in Fig. 3-21. A unique

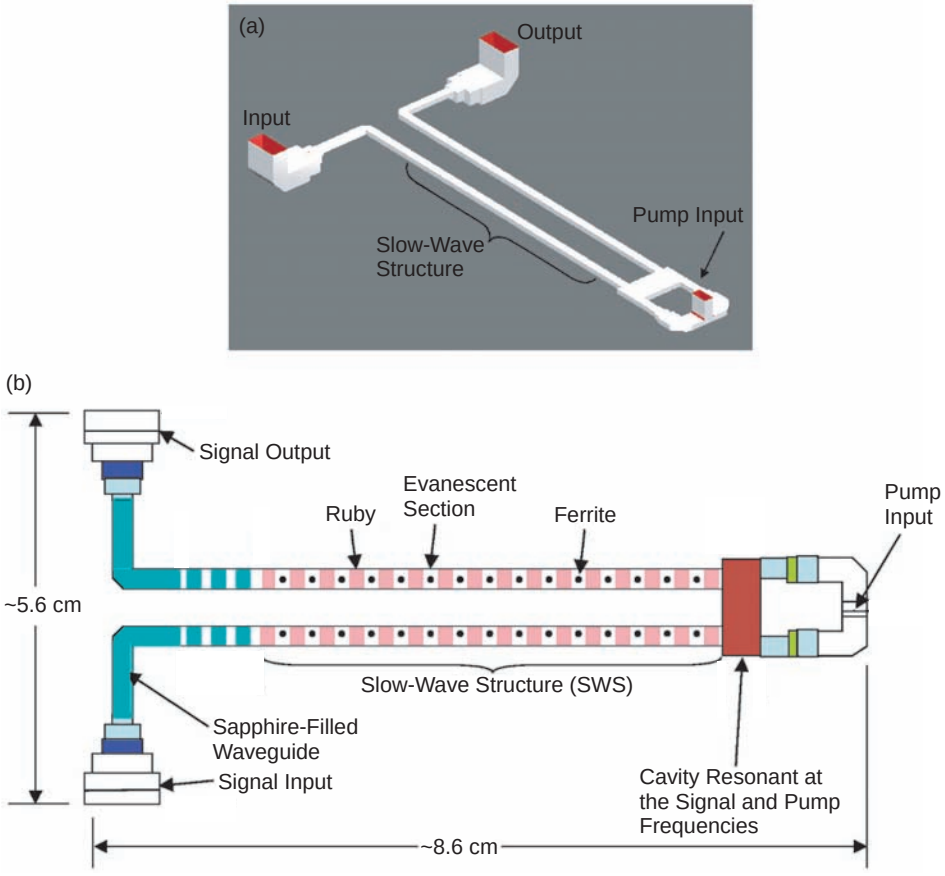


Fig. 3-21. 32 GHz cascaded-cavity TWM with (a) perspective view and (b) schematic view (not to scale).

feature of this design is the smooth waveguide structure that contains the ruby resonators, without irises. Sections of waveguide containing resonance isolators and a low dielectric constant material are beyond cut-off at the signal frequency and the pump frequency. The ruby resonator dimensions are such that the resonators form a slow-wave structure with pass-bands that exist at both the signal frequency and the pump frequency. This is done to maximize the pumping efficiency, as well as the signal frequency gain-bandwidth product. A "U-turn" cavity is used at one end of the structure enabling the introduction of pump power at that end. The signal input and output connection of the two side-by-side structures are at the end opposite the U-turn cavity.

The approach of using a smooth waveguide structure without irises was also used to design a coupled-cavity reflection type amplifier. A sketch of this design is shown in Fig. 3-22. The cavity closest to the circulator is made of sapphire and serves as a coupling cavity for the signal. Its dimensions are such

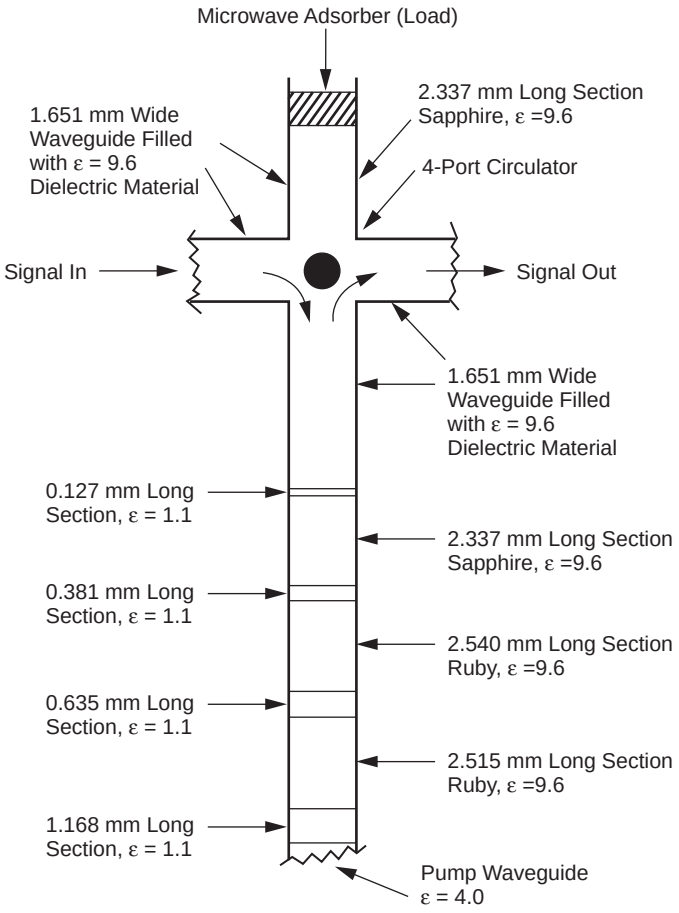


Fig. 3-22. 32-GHz coupled-cavity reflection-type amplifier.

that it reflects pump power, maximizing the pumping efficiency. The following two cavities are ruby and are resonant at both the signal and pump frequencies. Analysis of the coupled-cavity stage shows about 10 dB gain with a 500-MHz bandwidth from 31.8 GHz to 32.3 GHz when cooled to 1.5 K. Four stages can be used in series to produce 40 dB gain with a 500-MHz bandwidth. The analyzed frequency response for the four-stage smooth-wall coupled-cavity maser is shown in Fig. 3-23.

The simple coupled-cavity maser design is also well suited for laboratory measurements requiring the ultimate sensitivity. Calculations of the coupled-cavity maser noise temperature show effective input noise temperatures of 1.983 K when cooled to 1.0 K, 2.343 K when cooled to 1.5 K, 2.573 K when cooled to 1.8 K, and 2.754 K when cooled to 2.0 K. Connected to a termination in the super-fluid helium bath, the total operating system noise temperatures would be 2.404 K, 3.204 K, 3.713 K, and 4.083 K at bath temperatures of 1.0 K, 1.5 K, 1.8 K, and 2.0 K, respectively. The h*f*/k quantum noise at 32 GHz is 1.536 K, which represents 63.9 percent of the total system temperature in the 1.0-K bath and 37.6 percent of the total system temperature in the 2.0-K bath. Laboratory measurements verifying the thermal and quantum noise temperature quantities are suggested.

These ruby maser designs are available for future applications when stringent deep space telecommunications link needs arise at 32 GHz, as they have historically at S-band and X-band.

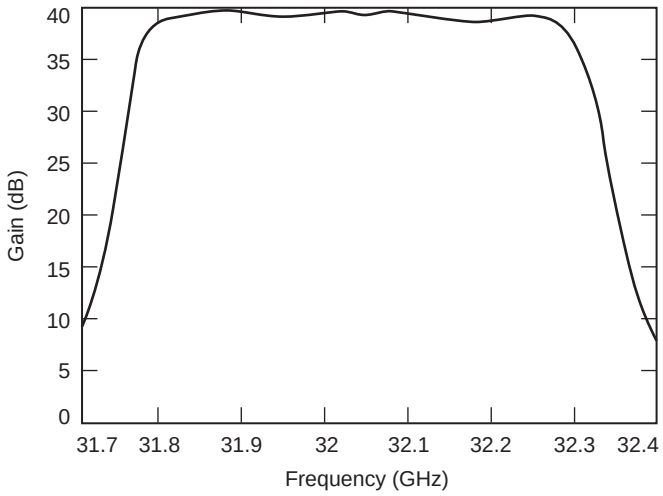


Fig. 3-23. Analyzed frequency response of four-stage 32-GHz coupled cavity reflection-type amplifier.

References

[1] W. R. Corliss, *A History of the Deep Space Network*, NASA CR-151915, National Aeronautics and Space Administration, Washington, District of Columbia, May 1, 1976.

[2] T. Sato and C. T. Stelzried, "An Operational 960 Mc Maser System for Deep Space Tracking Missions," *IRE Transactions on Space Electronics and Telemetry*, vol. SET-8, pp. 164–170, June 1962.

[3] R. C. Clauss, "A 2388Mc Two-Cavity Maser for Planetary Radar," *Microwave Journal*, vol. 8, pp. 74–77, May 1965.

[4] M. S. Reid, R. C. Clauss, D. A. Bathker, and C. T. Stelzried, "Low-Noise Microwave Receiving Systems In A W orldwide Network of Large Antennas," *Proceedings of the IEEE*, vol. 61, pp. 1330–1335, 1973.

[5] J. S. Shell, R. C. Clauss, S. M. Petty, G. W. Glass, M. S. Fiore, J. J. Kovatch, J. R. Loreman, D. E. Neff, R. B. Quinn, and D. L. Trowbridge, "Ruby Masers for Maximum G/T_{op}," *Proceedings of the IEEE*, vol. 82, no. 5, pp. 796–810, May 1994.

[6] A. E. Siegman, *Microwave Solid State Masers*, McGraw-Hill Book Company, New York, New York, 1964.

[7] W. S. Chang and A. E. Siegman, *Characteristics of Ruby for Maser Applications*, Tech. Rept. 156-2 Stanford Electronics Laboratories, Stanford University, California, September 30, 1958.

[8] A. R. Kerr, "Suggestions for Revised Definitions of Noise Quantities, Including Quantum Effects," *IEEE Transactions on Microwave Theory and Techniques*, vol. 47, no. 3, pp. 325–329, March 1999.

[9] C. T. Stelzried, R. C. Clauss, and S. M. Petty, "DSN Receiving Systems' Operating Noise Temperature Measurements," *The Interplanetary Network Progress Report 42-154, April–June 2003*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–7, August 15, 2003. http://ipnpr.jpl.nasa.gov/progress_report/

[10] R. Clauss and E. Wiebe, "Low-Noise Receivers: Microwave Maser Development," *The Deep Space Network Progress Report for November and December 1973,* Technical Report 32-1526, Vol. XIX, Jet Propulsion Laboratory, Pasadena, California, pp. 93–98, February 15, 1974. http://ipnpr.jpl.nasa.gov/progress_report/

[11] R.W. DeGrasse, E. O.Schulz-DuBois, H. E. D. Scovil, "The Three-Level Solid State Traveling Wave Maser," *The Bell System Technical Journal*, vol. 38, no. 2, pp. 305–334, March 1959.

[12] S. Okwit, F. R. Arams, and J. G. Smith, "Electronically Tunable Traveling-Wave Masers at L and S-Bands," *Proceedings of the IRE*, vol. 48, p. 2025, May 1960.

[13] D. L Trowbridge, "Block III Maser Implementation Program," *The Deep Space Network Progress Report for September and October 1973*, Technical Report 32-1526 Vol. XVIII, Jet Propulsion Laboratory, Pasadena, California, pp.130–135, December 15, 1973. http://ipnpr.jpl.nasa.gov/progress_report/

[14] J. W. Layland, "ICE Telemetry Performance," *The Telecommunications and Data Acquisition Progress Report 42-84, October–December 1985*, Jet Propulsion Laboratory, Pasadena, California, pp. 203–213, February 15, 1986. http://ipnpr.jpl.nasa.gov/progress_report/

[15] J. W. Layland and L. L. Rauch, "The Evolution of Technology in the Deep Space network: A History of the Advanced Systems Program," *The Telecommunications and Data Acquisition Progress Report 42-130, April–June 1997,* Jet Propulsion Laboratory, Pasadena, California, pp. 1–44, August 15, 1997. http://ipnpr.jpl.nasa.gov/progress_report/

[16] C. T. Stelzried, "Operational Performance of an H-Band Coupled Cavity Transmission Maser," *Microwave Journal*, vol. 10, pp. 103–106, March 1967.

[17] S. M. Petty and R. C. Clauss, "X-Band Traveling Wave Maser," *IEEE Transactions on Microwave Theory and Technology*, vol. MTT-16, no. 1, pp. 47–48, January 1968.

[18] R. Berwin, E. Wiebe, and P. Dachel, "Superconducting Magnet for a Ku-Band Maser," *The Deep Space Network Progress Report for July and August 1971*, Technical Report 32-1526, Vol. V, Jet Propulsion Laboratory, Pasadena, California, pp. 109–114, October 15, 1971. http://ipnpr.jpl.nasa.gov/progress_report/

[19] R. Clauss and R. Quinn, "Low Noise Receivers, Microwave Maser Development," *The Deep Space Network Progress Report for July and August 1971*, Technical Report 32-1526, Vol. V, Jet Propulsion Laboratory, Pasadena, California, pp. 102–108, October 15, 1971. http://ipnpr.jpl.nasa.gov/progress_report/

[20] R. W. Berwin, *Paramagnetic Energy Levels of the Ground State of Cr^{+3} in Al_2O_3 (Ruby)*, JPL Technical memorandum 33-440, Jet Propulsion Laboratory, Pasadena, California, January 15, 1970.

[21] C. R. Moore and R. C. Clauss, "A Reflected-Wave Ruby Maser with K-band Tuning Range and Large Instantaneous Bandwidth," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-27, no. 3, pp. 249–256, March 1979.

[22] D. L. Trowbridge, "X-Band Traveling Wave Maser Amplifier," *The Deep Space Network Progress Report for July and August 1973*, Technical Report 32-1526 Vol. XVII, Jet Propulsion Laboratory, Pasadena, California, pp. 123–130, October 15, 1973. http://ipnpr.jpl.nasa.gov/progress_report/

[23] D. L. Trowbridge, "X-band, Low-Noise, Traveling-Wave Maser," *The Telecommunications and Data Acquisition Progress Report 42-60, September and October 1980*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–38, pp. 126–132, December 15, 1980. http://ipnpr.jpl.nasa.gov/progress_report/

[24] R. C. Clauss and R. B. Quinn, *Resonant Isolator for Maser Amplifier*, New Technology Transmittal NASA-CASE-NPO-15201-1 (NASA-15201), National Aeronautics and Space Administration, Pasadena Office, Pasadena, California, 1983.

[25] D. L. Johnson, S. M. Petty, J. J. Kovatch, and G. W. Glass, "Ultralow Noise Performance of an 8.4-GHz Maser–Feedhorn System," *The Telecommunications and Data Acquisition Progress Report 42-100, October–December 1990*, Jet Propulsion Laboratory, Pasadena, California, pp. 100–110, February 15, 1990. http://ipnpr.jpl.nasa.gov/progress_report/

[26] R. C. Clauss, "Low Noise Receivers: Microwave Maser Development," *The Deep Space Network Progress Report 42-25, November and December 1974*, Jet Propulsion Laboratory, Pasadena, California, pp. 65–68, February 15, 1975. http://ipnpr.jpl.nasa.gov/progress_report/

[27] L. D. Flesner and Sheldon Schultz, "Simple Waveguide Reflection Maser with Broad- Tunability," *Review of Scientific Instruments*, vol. 48, no. 8, pp. 1104–1105, August 1977.

[28] C. R. Moore, "A K-Band Ruby Maser with 500-MHz Bandwidth," *IEEE Transactions on Microwave Theory Techniques*, vol. MTT-28, no. 2, pp. 149–151, February 1980.

[29] J. Shell and D. Neff, "A 32 GHz Reflected Wave Maser Amplifier with Wide Instantaneous Bandwidth," *1988 MTTS International Microwave Symposium Digest* (1988 Vol. II [MYSYM]), pp. 789–792, 1988.

[30] *WASP-NET, Waveguide Synthesis Program for Waveguide Networks*, Microwave Innovation Group, Bremen, Germany.

[31] G. L. James, "Analysis and Design of TE_{11}-to-HE_{11} Corrugated Cylindrical Waveguide Mode Converters," *IEEE Transactions on Microwave Theory and Techniques*, MTT-29, no. 10, pp. 1059–1066, October 1981.

[32] D. J. Hoppe, "Modal Analysis Applied to C ircular, Rectangular, and Coaxial Waveguides," *The Telecommunications and Data Acquisition Progress Report 42-95, July–September 1988*, Jet Propulsion Laboratory, Pasadena, California, pp. 89–96, November 15, 1988. http://ipnpr.jpl.nasa.gov/progress_report/

[33] B. Lax and K. J. Button, *Microwave Ferrites and Ferrimagnetics*, McGraw-Hill Book Co., New York, New York, 1962

[34] A. Abragam and B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions*, Dover Publications, Inc., New York, New York, 1986.

[35] E. O. Schulz-DuBois, "Paramagnetic Spectra of Substituted Sapphires-Part I: Ruby," *Bell System Technical Journal*, vol. 38, pp. 271–290, January 1959.

[36] T. Chang, D. Foster, and A. H. Kahn, "An Intensity Standard for Electron Paramagnetic Resonance Using Chromium-Doped Corundum (Al_2O_3:Cr^{3+})," *Journal of Research of the National Bureau of Standards*, vol. 83, no. 2, pp. 133–164, March–April 1978.

[37] J. S. Shell and R. C. Clauss, "A 32 GHz Coupled Cavity Maser Design," *The Telecommunications and Mission Operations Progress Report 42-142, April–June 2000*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–22, August 15, 2000. http://ipnpr.jpl.nasa.gov/progress_report/

[38] J. Shell and R. Clauss, "A 32-Gigahertz Traveling-Wave Maser Design," *The Interplanetary Network Progress Report 42-159*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–22, November 15, 2004. http://ipnpr.jpl.nasa.gov/progress_report/

[39] J. C. Fletcher and E. R. Wiebe, *Multistation Refrigeration System*, United States Patent 4,077,231, March 7, 1978.

[40] J. C. Fletcher and E. R. Wiebe, *Helium Refrigerator*, United States Patent 3,914,950, October 28, 1975.

[41] J. C. Fletcher, E. R. Wiebe, and R. C. Clauss, *Refrigerated Coaxial Coupling*, United States Patent 3,902,143, August 26, 1975.

[42] G. M. Low and E. R. Wiebe, *Helium Refrigerator and Method for Decontaminating the Refrigerator*, United States Patent 3,656,313, April 18, 1972.

[43] E. R. Wiebe, *Closed-Cycle Refrigerator for Masers*, NPO-13839 (NASA Technical Reports Server Document ID 19770000056), Jet Propulsion Laboratory, Pasadena, California, 1977.

[44] E. R. Wiebe, *Measuring Cryogenic-Refrigerator Cooling Capacity*, NPO-13435 (NASA Technical Reports Server Document ID 19770000411), Jet Propulsion Laboratory, Pasadena, California, 1978.

[45] E. R. Wiebe, *Automatic Thermal Switch Accelerates Cooling-Down of Cryogenic System* (NASA Technical Reports Server Document ID 195650000067), Jet Propulsion Laboratory, Pasadena, California, 1965.

[46] W. Higa and E. R. Wiebe, *Spiraled Channels Improve Heat Transfer Between Fluids* (NASA Technical Reports Server Document ID 119650000289, Jet Propulsion Laboratory, Pasadena, California, 1965.

[47] R. C. Clauss, *Transmission Line Thermal Short*, United States Patent 3,474,357, October 21, 1969.

[48] T. O. Paine and R. C. Clauss, *High-Gain Broadband Traveling Wave Maser*, United States Patent 3,486,123, December 23, 1969.

[49] G. M. Low, R. B. Quinn, and R. C. Clauss, *Maser for Frequencies in the 7-20 GHz Range*, United States Patent 3,676,787, July 11, 1972.

[50] J. C. Fletcher and R. C. Clauss, *Reflected-Wave Maser*, United States Patent 3,978,417, August 31, 1976.

[51] J. C. Fletcher, E. R. Wiebe, and R. C. Clauss, *Refrigerated Coaxial Coupling*, United States Patent 3,902,143, August 26, 1975.

[52] R. A. Frosch, R. C. Clauss, and R. B. Quinn, *Dielectric-Loaded Waveguide Circulator for Cryogenically Cooled and Cascaded Maser Waveguide Structures*, United States Patent 4,187,470, February 5, 1980.

[53] R. C. Clauss and R. B. Quinn, *Resonant Isolator for Maser Amplifier*, United States Patent 4,399,415, August 16, 1983.

Chapter 4
Cryogenic Refrigeration Systems

Robert C. Clauss

4.1 Introduction

Use of cryogenic cooling by the Deep Space Network (DSN) includes both open-cycle refrigeration (OCR) and closed-cycle refrigeration (CCR) systems. The temperatures achieved by these systems range from 1.5 kelvins (K) to about 80 K, depending upon the type of system used. These cryogenic systems are used to cool low-noise preamplifiers and some of the antenna feed system components for the DSN's receivers. Liquid nitrogen (LN_2) was used to cool reference loads (resistive terminations) used for noise temperature measurements, and liquid helium (LHe) was used to cool reference loads and antenna-mounted ruby masers. Russell B. Scott in *Cryogenic Engineering* [1] explains many aspects of cryogenic technology in terms that are easily understood. Progress since 1959, has given us many types of CCR systems that can be used for cooling low-noise microwave amplifiers.

Cryogenic refrigeration is a term that may be applied to the process of cooling equipment and components to temperatures below 150 K. The net capacity of a cryogenic refrigeration system at a particular temperature is the amount of heat that can be applied to a "cold station" in the system without warming the station above that particular temperature. The cold station may be a bath of cryogenic fluid, or the cold station may be a conductive surface cooled to the bath temperature to which equipment may be fastened.

Cryogenic refrigeration systems are different from the refrigeration equipment we encounter in our everyday environment. The refrigerants used in cryogenic systems are often helium (He), hydrogen (H_2), or nitrogen (N_2). Insulation techniques used to minimize heat leaks into the cooled parts of the

systems usually depend on the use of high-vacuum technology, radiation shields, and structural materials with low thermal conductivity. Systems that use stored cryogens such as liquid helium, liquid hydrogen, or liquid nitrogen in a container called a "dewar" are usually refilled on a periodic basis. Solidified gases (such as hydrogen or methane) can also be used for cooling purposes, much as solid carbon dioxide (dry ice) is used to refrigerate perishable foods during shipment, but this has not been done in the DSN.

Development of cryogenic refrigeration equipment and systems for laboratory, military, and commercial purposes began many years before development of the ruby masers and other low-noise equipment used in the DSN. This was fortunate, but the personnel developing the DSN's low-noise amplifiers did not have all of the knowledge and expertise needed for developing or purchasing cryogenic equipment and systems. There was much to be learned and many pitfalls to be avoided.

Techniques and materials needed for the efficient transfer of electrical power and microwave signals from a room-temperature environment into a cryogenic environment are often not compatible with the techniques needed to provide adequate thermal isolation. The development of very low-loss microwave input transmission lines and waveguides with high thermal isolation was challenging. Vacuum seals or windows in the transmission lines or waveguides could be degraded by condensation collecting on surfaces that were cooled by conduction or radiation. Many problems like these were waiting to be solved during the early years of cryogenically-cooled low-noise amplifier (LNA) development for the DSN.

Transferring liquid helium into an open-cycle dewar required the use of a vacuum-jacketed transfer line. Inadequate insulation in the transfer line would cause an expensive failure, wasting the precious liquid helium. Our leader and teacher, Dr. Walter H. Higa, noted, in 1960, that the cost of a liter (L) of liquid helium was about the same as the cost of a liter of good Scotch whisky. About 6 L of liquid helium a day were used for each antenna-mounted maser. Construction or procurement and maintenance techniques for liquid-helium transfer lines were learned the hard way.

Difficulties encountered during the development and field use of the eventually successful antenna-mounted open-cycle liquid-helium-dewar systems provided incentives for an alternative approach. Dr. Higa wrote, in a 1962 memorandum, "The inconvenience of having to refill a dewar is quite obvious, and much effort is being expended to perfect a closed-cycle refrigerator (CCR) for maser applications." The memo included the photograph of a liquid-helium transfer at the apex of a 26-meter (m) antenna shown in Chapter 3, Fig. 3-1.

The early closed-cycle helium refrigerator systems were not without problems. Early model 210 Cryodynes® purchased from the Arthur D. Little Corporation (ADL) in Cambridge Massachusetts, experienced frequent gear

failures in the drive units of the refrigerators. The compressors that supplied high-pressure helium gas to the antenna-mounted Cryodynes® often contaminated the helium with the lubricant used in the compressor. The expression "oil carry-over" was used often when reporting Cryodyne® warm-ups (failures). The Joule-Thomson (JT) counter-flow heat exchangers in the Cryodyne® had been plugged with lubricant that had solidified.

Our early learning period of using, maintaining, servicing, repairing, and modifying CCRs to cool masers on antennas in the field lasted from 1961 to 1966. Many problems were solved. Then, the development and production of the Model 340 A DL Cryodyne® in 1965 provided the world with a reliable two-stage Gifford-McMahon (GM) CCR. This 15-K CCR had adequate capacity for use with a JT counter-flow heat exchanger system developed by Dr. Higa and Ervin R. Wiebe [2]. This new CCR provided reliable 4.5-K refrigeration for DSN masers beginning in 1966.

The costs and difficulties experienced developing and using the early cryogenically-cooled masers seemed high, but the value of the low-noise masers proved to be more than worth the cost and effort. The expense to provide and operate larger antennas, or more antennas, that could be used to equal the existing antenna and maser-receiving system's figure of merit (G/T_{op}) was more than the cost of building and operating the masers by factors of tens to hundreds. Advances in cryogenic refrigeration technology continued, enabling improvements that helped to maximize the performance of deep space missions.

4.2 Advantages of Using Cryogenic Cooling

Cooling microwave components and LNAs to cryogenic temperatures enables significant reductions in the operating noise temperature (T_{op}) of receiving systems. The sensitivity of a receiving system is directly proportional to A/T_{op}, where A is the receiving antenna's effective collecting area. For example, when $T_{\mathrm{op}} = 80\ \mathrm{K}$, an array of four identical antennas and receivers is needed to equal the sensitivity of one such antenna and receiver with a T_{op} of 20 K. In 1965, when 26-m-diameter antennas were in operation and 64-m-diameter antennas were being built for the DSN, ruby traveling wave masers (TWM) cooled by 4.5-K Cryodynes® were the logical economical choice of LNAs for DSN receivers.

Ruby cavity masers were used on antennas in the Deep Space Instrumentation Facility (DSIF) and the DSN beginning in 1960. The performance of a maser at DSN frequencies (below 40 gigaherz [GHz]) improves as the bath temperature is reduced. The bath temperature term (T_b), is often used for the thermodynamic temperature of a maser whether it is

immersed in a liquid-helium bath, or cooled by conduction with a closed-cycle refrigerator. The maser noise temperature is proportional to the bath temperature, and the maser electronic gain in decibels (dB) varies inversely with the bath temperature. The temperature dependence of masers is explained mathematically in Chapter 3. All DSN ruby masers are cooled to temperatures below 5 K. Transistor LNAs used in the DSN are less temperature dependent than masers and are often cooled to temperatures between 5 and 20 K.

Figure 4-1 is a photo of two X-band feedhorns mounted on an X-band maser during noise temperature measurements. T_{op} was measured by switching from the ambient load (located between the two horns) and either horn. The measurement was used to determine the noise temperature difference of the corrugated feedhorn and a smooth feedhorn. The corrugated feedhorn resulted in a T_{op} measurement that was 0.1 K lower than that measured with the smooth horn.

A high vacuum inside the CCR reduces heat transfer from the ambient vacuum housing to the cryogenically cooled assembly within the housing. A vacuum window in the X-band signal waveguide entering the CCR seals the system from the atmosphere while passing the microwave signals to the LNA system within the vacuum housing. The CCR vacuum window withstands the atmospheric pressure and attenuates the incoming signal by about 0.01 dB. This loss at 300 K adds 0.7 K to the X-band maser's effective input noise temperature as measured at the ambient interface of the maser package.

DSN X-band systems use waveguide components, including a diplexer to accommodate a transmit capability, a filter for out-of-band radio-frequency interference (RFI) and transmit signal rejection, and a polarizer preceding the receiver's LNA. These components cause loss and noise, thereby degrading the receiver's sensitivity. The microwave loss of components made of good quality electrolytic copper (oxygen-free high conductivity (OFHC) or electrolytic half-hard) drops by a factor of two when cooled from 290 K to 80 K, and by a factor of three when cooled from 290 K to 20 K. The microwave loss does not change below 20 K. These factors (ratios) are independent of frequency in the 2-GHz to 34-GHz range measured. The direct current (dc) resistance of electrolytic half-hard copper drops by a factor of about 9 when cooled from 290 K to 80 K and by a factor of about 100 when cooled from 290 K to 20 K. The microwave loss does not vary as the square root of the dc resistance at cryogenic temperatures. The process used to determine the noise contribution of components preceding LNAs was described in an earlier chapter and is not repeated here. Table 4-1 describes the advantages of cooling feed-system components preceding an X-band maser.

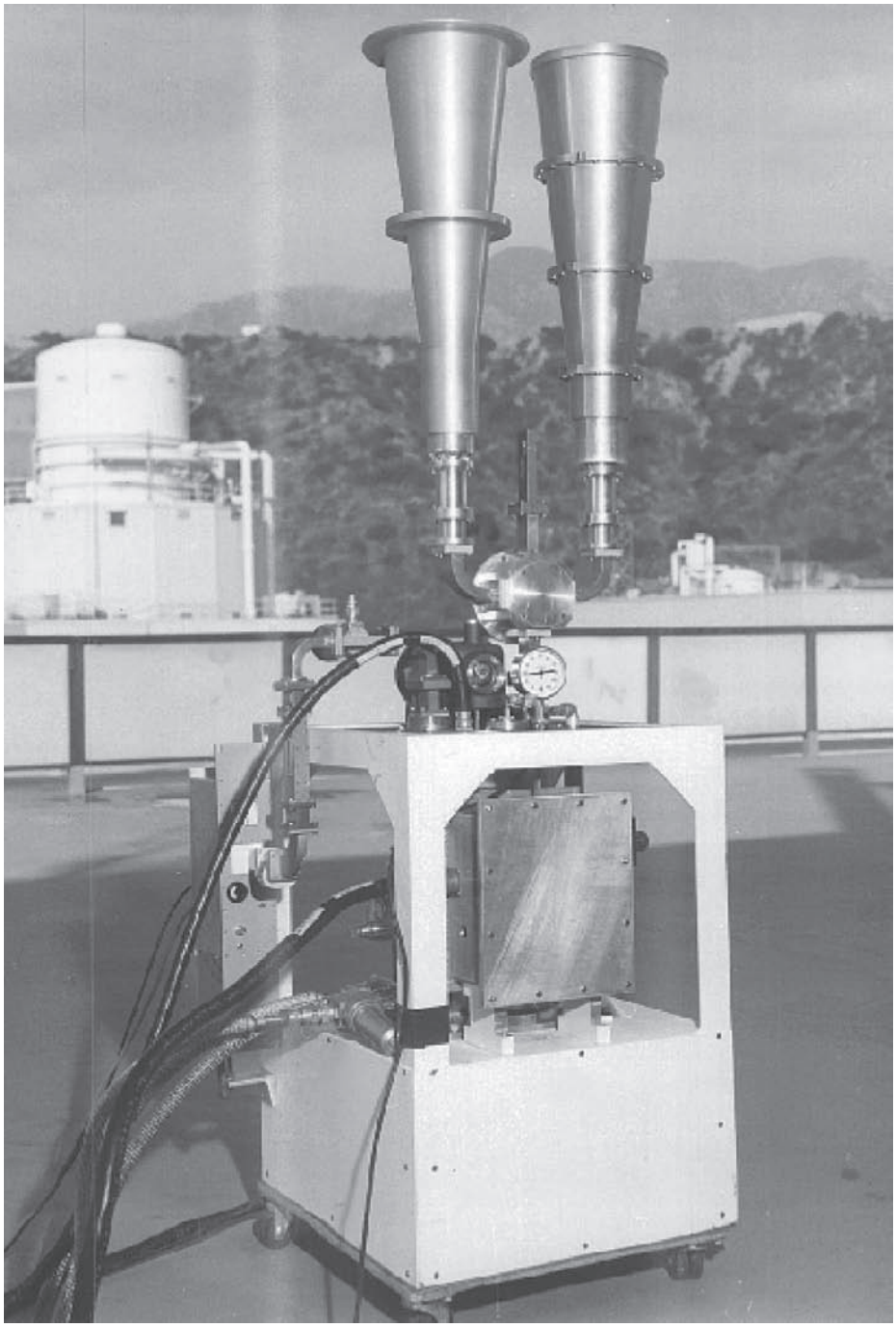


Fig. 4-1. X-band TWM with X-band feed-horns.

The system noise temperature (T_{op}) values shown in Table 4-1 are reasonable for 70-m DSN antennas currently operating at 8420 m egahertz (MHz) (X-band) at an elevation angle of 30 degrees (deg) in clear dry weather. The noise from the antenna at the feedhorn input is 12 K. Component thermodynamic temperatures of 300 K, 80 K, 40 K, 20 K, and 4 K are used. Table 4-1 T_{op} values are calculated at the feedhorn input, in the space just above the feedhorn.

Figure 4-2 shows three of the feed-system components (a) outside of the CCR vacuum housing at room temperature and (b) the same components inside the CCR vacuum housing at various cryogenic temperatures. The maser input noise temperature at the vacuum window ambient interface shown in Fig. 4-2 (a) is 3.0 K. The examples used in Table 4-1 gives the combined insertion loss of the three components at the various temperatures. T_{op}, the sensitivity changes, and the maximum data-rate percentages are calculated values, with 100 percent being the reference rate when the feed-system components are at ambient temperature. Ambient-temperature feed-system components at S-band and X-band were used in the DSN in most operational systems prior to the year 2002.

4.3 Open-Cycle Refrigeration

Liquid helium and liquid nitrogen in open-cycle dewars have been used in the DSN to c ool resistive terminations (loads) for calibrations, LNAs, and microwave feed-system components. The normal boiling point of liquid nitrogen is 77.395 K at 760 millimeters of mercury (mm Hg) (1.000 atmosphere (atm), 101.32 kPa (kilopascals)), and the temperature varies with vapor pressure from 64 K at 109.4 mm Hg (0.1439 atm, 14.6 kPa) to 84 K at 1539 mm Hg (2.025 atm, 205 kPa), and up to a critical temperature of 126.1 K

Table 4-1. System noise temperature, receiving system sensitivity change, and maximum data rate as a function of component physical temperature and loss.

Feed-System Components Temperature (K)	Feed-System Components Loss (dB)	System Noise Temperature (K)	Sensitivity Change (dB)	Maximum Data Rate (%)
300	0.24	32.729	0.000	100 (Reference rate)
80	0.12	17.800	2.645	183.9
40	0.10	16.475	2.981	198.7
20	0.08	15.902	3.135	205.8
4	0.08	15.603	3.217	209.8

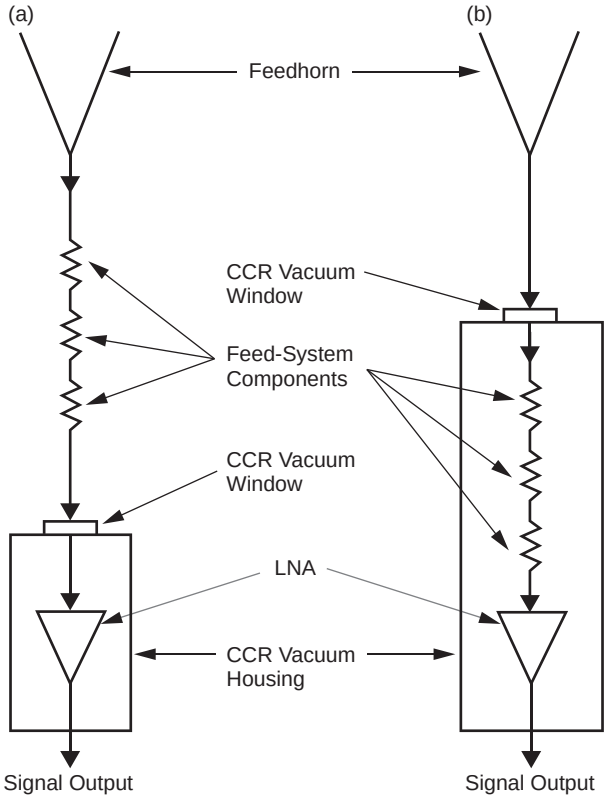


Fig. 4-2. Feed-system components (a) outside of the CCR vacuum at room temperature and (b) inside the CCR vacuum housing and cooled.

at the critical pressure of 25,454.2 mm Hg (33.49 atm, 3393 kPa, or 492.23 pounds per square inch absolute (psia)) [1]. Liquid nitrogen is available at a relatively low cost that ranges from about 20 cents to 2 dollars per liter in United States currency, depending on the circumstances. Liquid nitrogen has sufficient cooling capacity for many applications.

Measurement techniques used to determine the effective input noise temperature of a receiver (when the receiver is connected to an ambient load and then to a liquid-nitrogen-cooled load) were described in Chapter 2. The accuracy of the measurement is dependent upon the accuracy of the knowledge of the noise temperature of the loads at the input to the amplifier system. The temperature of the liquid nitrogen is dependent upon the pressure of the vapor above the liquid. When this vapor pressure is determined by atmospheric pressure that supports a 760-mm column of mercury (101.3 kPa, 14.7 psia) at sea level, the temperature of the liquid nitrogen varies with altitude as shown in the graph below (Fig. 4-3).

The graph in Fig. 4-3 is based on the vapor pressure data found in reference [1], a portion of which is summarized in Table 4-2 for liquid nitrogen and liquid helium.

A reduction in pressure of 12.68 mm Hg (1.69 kPa) from 760 mm Hg to 747.32 Hg (101.32 kPa to 99.63 kPa) causes a reduction in the temperature of

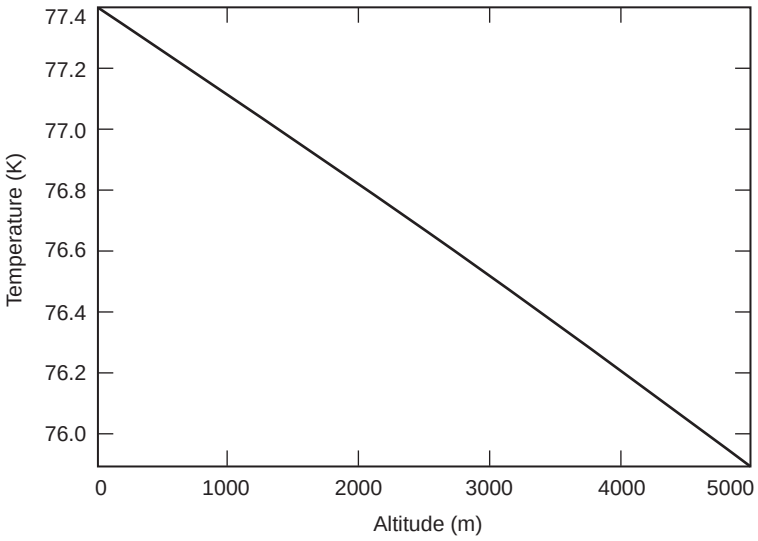


Fig. 4-3. Liquid nitrogen temperature versus altitude.

Table 4-2. Altitude, Boiling Point of Liquid Helium and Liquid Nitrogen, and Pressure

Altitude, feet (meters)	Temperature (K)		Pressure			
	Liquid Helium	Liquid Nitrogen	(mm) Hg	(psia)	(inches Hg)	kPa
0 (0)	4.216	77.395	760	14.70	29.92	101.3
1000 (304.8)	4.181	77.119	735	14.22	28.94	98.0
2000 (609.6)	4.145	76.832	710	13.73	27.95	94.7
3000 (914.4)	4.108	76.534	685	13.25	26.97	91.3
4000 (1219.2)	4.007	76.224	660	12.77	25.98	88.0
5000 (1524)	4.031	75.902	635	12.28	25.00	84.7

liquid nitrogen by 0.1 K. A reduction in the pressure of the same amount causes a 0.0127-K reduction in the temperature of liquid helium.

Liquid helium is used to cool components to temperatures below 4.2 K by using a vacuum pump to reduce the vapor pressure above the liquid helium in the container. As with liquid nitrogen, liquid helium is readily available on a commercial basis at a higher cost than liquid nitrogen. Today's cost of liquid helium ranges upward from $3 per liter in United States currency, depending on the circumstances.

The cooling capability of liquid helium, in terms of watts per liter (W/L), is lower than the cooling capability of liquid nitrogen by a factor of about 63. It takes 0.716 watt (W) for a period of one hour to convert 1 L of liquid helium at 4.2 K to gas. It takes about 45 W for a period of one hour to convert 1 L of liquid nitrogen to gas. For comparison, it takes about 628 W for a period of one hour to convert 1 L of water at 100 deg C to gas (steam). Dewars used to store liquid helium and liquid nitrogen are shown in Fig. 4-4.

Scott, in reference [1], page 215 writes, "The invention of the vacuum insulated vessel for liquefied gases by James Dewar in 1892 was a break-through in the field of thermal insulation that has not yet been matched by further developments. All of the advances since Dewar's time have been improvements on Dewar's original concept, usually by means of reducing radiant heat transfer by attaining surfaces of higher reflectivity or by interposing shields which reflect or intercept radiant energy." These words are still true today.

The dewars shown in Fig. 4-4 (at Goldstone) were commercially available for use with liquid nitrogen and liquid helium long before their first use at the DSS-11 DSIF Goldstone Tracking Station in 1960. The forklift planned for the transportation of dewars did not materialize, but the "HAPPY-TIME *RACER*" (little red wagon) shown in Fig. 4-4, requested by Walter H. Higa and provided by Charles T. Stelzried, was adequate for the task. These dewars can be used to store liquid helium and liquid nitrogen for many weeks with minimal loss. The liquid-helium dewar contains a liquid-nitrogen tank that surrounds the helium tank to intercept radiation from the ambient environment. Styrofoam containers can be used to hold liquid nitrogen for short periods. A Styrofoam coffee cup will hold liquid nitrogen for many minutes, and a Styrofoam container (such as the one shown in Fig. 4-5) will hold liquid nitrogen for several hours.

Figure 4-5 shows manual switching between an ambient load and a liquid-nitrogen temperature load. Dr. Walter H. Higa is holding a Styrofoam bucket containing liquid nitrogen and a calibration load. Robert S. Latham is facing the camera, and the author has his back to the camera. The walls of the Styrofoam bucket were about 2.5 centimeters (cm) (1 in.) thick, providing adequate insulation for the liquid-nitrogen-cooled calibration load. The noise temperature of a cavity maser was being measured.

Fig. 4-4. Liquid helium and liquid nitrogen dewars on "HAPPYTIME *RACER*" utility vehicle with the cherry picker (High Ranger) and cherry-picker fiberglass basket in the background.

The cavity maser and dewar system shown in Fig. 3-3 was typical of the dewars used for antenna-mounted operation of masers during the 1960-through-1965 time period. The dewar operated during antenna motion as the antenna was moved to track planets or spacecraft from about 10 deg above the eastern horizon to about 10 deg above the western horizon. The hold time for the 3-L liquid-helium tank was about 36 hours during the antenna-mounted tracking operations. Refills were accomplished on a daily basis using the High Ranger (cherry picker) partially shown in Fig. 4-4 (and completely shown in Fig. 3-1).

Fig. 4-5. Noise temperature measurement (manually switching between ambient and liquid nitrogen temperature loads) showing use of styrofoam container for storing liquid nitrogen.

This first dewar used to cool a 960-MHz cavity maser to 4.2 K on a DSIF antenna in 1960 was designed by Dr. Walter H. Higa. The dewar and maser were built and serviced in the field by members of his group. The dewar had a liquid-helium capacity of 3 L and a liquid-nitrogen capacity of 7 L. The liquid-helium tank was surrounded by the liquid-nitrogen tank, and aluminum radiation shields were attached to the top and bottom of the liquid-nitrogen tank. This geometry accommodated the application of the magnetic field needed for maser operation by a large permanent magnet located at ambient temperature, outside of the dewar.

Neoprene O-ring seals were used in several locations, between the vacuum housing and the top and bottom plates, and at the flanges supporting the cryogen fill-tubes and the coaxial transmission line. Neoprene is not impervious to helium gas. Initial evacuation of the dewar was followed by filling the dewar with liquid helium. The 4.2-K temperature caused residual gases in the vacuum jacket of the dewar to be trapped (solidified) on the surface of the liquid-helium tank. The residual gas pressure in the vacuum jacket dropped to less than

1×10^{-6} torr (1.333×10^{-4} Pa) for many hours. Unfortunately, small amounts of helium gas diffused through the neoprene O-rings and could not be trapped. Within a few days, the accumulation of helium gas in the vacuum jacket of the dewar was sufficient to destroy the insulating properties of the vacuum jacket. A Vac-ion® pump that could pump air at the speed of 5 liters-per-second (Lps) and helium at a speed of nearly 1 Lps was installed on the dewar, and this pump solved the problem.

The 36-hour liquid-helium lifetime of the 3-L tank in the dewar described above indicates a total heat leak and maser pump energy load into the helium bath of about 0.06 W. The liquid-nitrogen lifetime was about the same as the liquid-helium lifetime, indicating a heat load of about 9 W. The liquid-helium and liquid-nitrogen tanks were refilled on a daily basis. The exiting cold helium gas could have been used to reduce the amount of heat entering the dewar conductively along the coaxial signal transmission line, but this particular design did not take advantage of helium-vapor cooling.

The dewar design described above was used for the 2388-MHz single-cavity maser in the 1961 Venus Radar experiment and for several 960-MHz masers supporting spacecraft missions to Venus and the Moon. A slightly larger dewar with a similar design was used for the two-cavity 2388-MHz maser for the reception of radar signals from Venus and Mars in 1962 and 1963.

Commercially available liquid-helium dewars that were suitable for antenna mounted maser operation became available about the time that DSN switched to cryodyne CCRs for cooling masers. One commercially available liquid-helium dewar was used on a 9-m-diameter antenna at the DSN's research station (DSS-13) to cool a multiple-cavity X-band maser purchased from the Hughes Aircraft Corporation in El Segundo, California.

4.4 Heat Transfer

Unwanted heat transfer by radiation and conduction into cryogenically cooled equipment should be minimized. The 9-W heat load observed on the liquid-nitrogen tank in the dewar designed by Dr. Higa is typical of cryogenic systems where a cold surface near liquid-nitrogen temperature is about 1/3 of a square meter (m^2). The insides of the ambient-temperature aluminum vacuum housing were polished mechanically, and the surface emissivity was probably less than 0.05.

Emissivity of the stainless-steel liquid-nitrogen tank was probably about 0.1. The emissivity of a cryogenically cooled surface is often degraded with time. Gases subliming on cold surfaces are capable of raising the emissivity to values close to 1. The emissivity of other room temperature surfaces (flanges, extensions, welds, and gaps) might have ranged from 0.1 to 1.

Scott wrote [1], "The rate at which a surface emits radiation is given by the Stefan-Boltzmann equation,

$$W = \sigma e A T^4 \tag{4.4-1}$$

where

e is the total emissivity at temperature T, K

A is the area,

σ is a constant having the value of 5.67×10^{-12} W/cm^2, K^{-4}

The net exchange of radiant energy between two surfaces is given by the expression

$$W = \sigma E A\left(T_2^4 - T_1^4\right) \tag{4.4-2}$$

where subscripts 1 and 2 refer to the cold and warm surfaces, respectively, and A is an area factor. In the case of cylinders or spheres, it will be taken as the area of the enclosed (inner) surface; in the case of parallel plates it is obviously the area of either surface. E is a factor involving the two emissivities."

Scott's further explanation [1] of the value of E is not repeated here. The complications of specular reflection and diffuse reflection, gray surfaces, and surface irregularities can produce errors that seem difficult, if not impossible, to define. Scott lists the "Emissivity (Total Normal)" of various materials in a later chapter [1], Chapter X, pp. 347 and 348, and explains, "It will be noted that the emissivity (or adsorptivity) of metals decreases with decreasing temperature. At a fixed temperature the emissivity must equal the adsorptivity. (If these differed, there could be a net transfer of heat between two surfaces at the same temperature—a violation of the second law of thermodynamics). The rate of radiant heat transfer between two surfaces at different temperatures, where the geometry permits multiple reflections, depends in a complicated way upon the emissivities of both surfaces and their absorptivities of the radiant energy emitted by the opposite surface as well as upon the geometric configuration. Fortunately, in the region of the electromagnetic spectrum of importance in cryogenic insulation, metals are approximately "gray bodies"; that is, their emissivities (and absorptivities) are almost independent of the wavelength of the radiation, so their emissivities for one distribution of wave lengths are nearly equal to their absorptivities for another distribution. Thus, the formula for radiant heat transfer given in Chapter 6 (p. 148 [1]), will give acceptable results."

One of the formulae on page 148 [1] is for diffuse reflection when long coaxial cylinders or concentric spheres are used to determine the value of E.

$$E = \frac{e_1 e_2}{e_2 + \dfrac{A_1}{A_2}\left(1 - e_2\right) e_1} \tag{4.4-3}$$

where A_1 is the area of the inner surface, A_2 is the area of the outer surface, e_1 is the emissivity of the colder inner surface, and e_2 is the emissivity of the warmer outer surface. This equation is suggested for use with the OCRs and CCRs used to cool preamplifiers in the DSN. The intent here is to provide a rough estimate of the radiation heat transfer that must have been part of the 9-W total heat load on the liquid nitrogen tank in Higa's dewar. Heat transfer by radiation is an important design consideration for CCRs as well as dewars. Using $A_1 / A_2 = 0.5$, $e_1 = 0.1$, and $e_2 = 0.05$ gives a value of about 0.0513 for E. Use of this value of E in Eq. (4.4-2) indicates heat transfer rates into the liquid nitrogen tank at 77.4 K of

(1) 10.13 W at an ambient temperature of 320 K (116.3 degrees Fahrenheit (deg F)),

(2) 7.82 W at an ambient temperature of 300 K (80.3 deg F), and

(3) 5.92 W at an ambient temperature of 280 K (44.3 deg F).

It is important to consider the environmental extremes to be encountered when designing such a system. The ambient environment at DSN tracking stations can vary between 266 K (19 deg F) and 322 K (120 deg F).

A change in the emissivity of the liquid-nitrogen tank from 0.1 to 0.2 would cause the radiation heat load from a 320-K surface to increase from to 10.13 to 13.62 W. The emissivity of a surface may vary as a function of time and conditions during use. The out-gassing of materials in a vacuum chamber and gases entering through leaks will change the emissivity of the surfaces on which these gases are solidified. The emissivity of these surfaces can increase to values far higher than those of the original clean configuration. Scott [1] lists the emissivities of smooth ice (H_2O) and glass at 0.94, paper at 0.92, white lacquer at 0.925, and candle soot at 0.952. Appearances can be deceiving. A white, or light-colored surface, is often not a good reflector of radiation at the wavelengths emitted by room-temperature surfaces.

Heat transfer from the ambient vacuum housing to the liquid-nitrogen and liquid-helium tanks by conduction through support structures, microwave transmission lines, waveguides, and wires used for electric circuits must be considered. Materials such as nylon, Perspex® (polymethylmethacrylate), Teflon® (polytetrafluoroethlyne), and fused quartz are very good insulators, especially at low temperatures. The thermal conductivities of materials at various temperatures are found in handbooks [3,4] and in graphs by Scott [1]. These thermal conductivity values for materials of interest for cryogenic applications are shown in Table 4-3.

Table 4-3. Thermal conductivity of various materials, milliwatts/cm K.

Material	Temperature (K)				
	4.2	20	76	194	273
Paper-fiberglass-foil layers	-	-	0.001	0.001	0.001
Polystyrene foam	-	-	0.33	0.33	0.33
Teflon	0.45	1.3	2.3	-	-
Fused quartz	1.3	1.6	4.8	9.5	14
Alumina	5	230	1500	480	-
Sapphire (36 deg to c-axis)	1100	35000	11000	-	-
Stainless steel (321)	3	20	80	130	140
Constantan	9	86	170	190	220
Beryllium copper	20	100	350	650	800
Steel, SAE 1020	130	200	580	650	650
50–50 lead–tin solder	150	550	505	510	-
Brass	23	120	390	700	1200
Aluminum 2024-T4	32	170	560	950	1300
6063-T5	330	1600	2300	2000	2000
1100	500	2400	2700	2200	2200
Phosphorus deoxidized copper	75	420	1200	1900	2200
Electrolytic tough pitch copper	3300	13000	5500	4000	3900
Copper, high purity	120000	105000	6600	4100	4000
Silver, 99.999% pure	144000	51000	5200	4200	4180
Silver solder	-	120	340	580	-
Helium gas	-	0.21	0.62	1.15	1.51
Nitrogen gas	-	-	-	0.18	0.26

The thermal conductivity value for paper-fiberglass-foil layers appears to be for a composite material that is sometimes called super insulation, but is normally referred to as multi-layer-insulation. The value shown may apply in a clean, hard-vacuum environment, but not when residual gas, condensed or frozen, forms a path for heat transfer through the layers.

Heat transferred by conduction through a material depends upon the thermal conductivity of the material (k), the length (L), and the cross-sectional area (A) of the material and the temperature difference (ΔT) across the length of the material. The rate of heat transfer is:

$$W = k\left(\frac{A}{L}\right)\Delta T \tag{4.4-4}$$

The thermal conductivity of materials often varies as a function of temperature, as shown in Table 4-3. The expression then becomes:

$$W = \frac{A}{L}\int_{T_c}^{T_h} k, dT \tag{4.4-5}$$

The integral has been solved for many common cryogenic materials and temperatures [5]. Heat transferred from the higher temperature (T_h) to the lower temperature (T_c) may be calculated in steps by using numerical integration with a sufficient number of steps. An example of the temperature profile and heat transfer through a 321 alloy stainless-steel rod having a length of 30 cm and a cross-sectional area of 1 cm^2 is shown in Fig. 4-6.

One watt is transferred from 300 K to 4 K, yet the temperature near the midpoint is 200 K, and the temperature is 100 K at a point close to the 4 K end. The same geometry of stainless steel transfers 0.088 W from 70 K to 4 K, and the 50-K point is near the midpoint. The center conductor of a coaxial transmission line made of thin-wall stainless-steel tubing stays at a higher temperature for a greater length into a cryogenic system than it would if the thermal conductivity of the stainless steel was constant as a function of temperature. The unfavorable temperature profile added to the loss and noise contributed by input transmission lines in early S-band masers.

4.5 Antenna-Mounted Operation

All OCRs and CCRs developed for use in the DSN between 1960 and 1990 were capable of operation at the primary or secondary (Cassegrainian) focal points of large fully steerable antennas. The DSN's construction of 34-m beam-waveguide (BWG) antennas after 1990 provided an ideal environment for cryogenically cooled equipment, for large transmitters that require liquid cooling, and for other electronics that benefit from operation in a stable, non-tilting environment. The BWG research antenna at DSS-13 was built for the purpose of developing Ka-band technology for future missions. The BWG antenna concept had been used successfully in many Earth terminals for commercial satellite applications. The 64-m Usuda (Japan) BWG antenna built

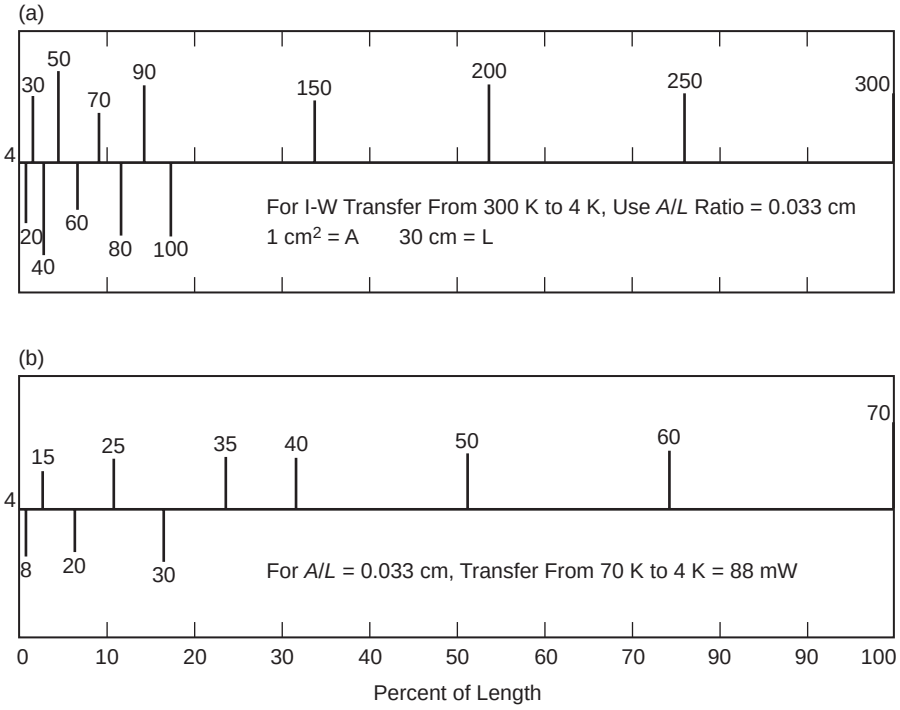


Fig. 4-6. Temperature profile of stainless steel for two conditions: (a) warm end 300 K and cold end 4 K; (b) warm end 70 K and cold end 4 K.

by Mitsubishi had demonstrated the low-noise characteristics needed for deep-space applications.

BWG antennas use a system of microwave reflectors (mirrors) to transfer the focal point of a large parabolic reflector antenna to a convenient location, such as a stationary room on the ground. The BWG concept is much like the optical coude telescope developed in 1888. The stationary environment of the DSS-13's front-end-area room is ideal for the use of liquid-helium dewars of the type shown in Fig. 4-7. This commercially available helium dewar was used in 1992 to cool a 33.7-GHz maser to 1.5 K for a Ka-band 'Link' experiment (KaBLE) with a spacecraft on its way to Mars. The 33.7-GHz maser was also used for dual-polarization radio astronomy observations. An X-band traveling-wave maser (TWM) operating at 1.8 K in a dewar similar to the one shown in Fig. 4-7 is used for planetary radar support at DSS-13.

The environment provided by super-fluid helium at temperatures below 2.19 K (liquid helium II) is ideal for masers and for calibration loads. The heat transport through liquid helium II is such that it is not possible to measure a temperature gradient anywhere in the bath. The temperature of any component

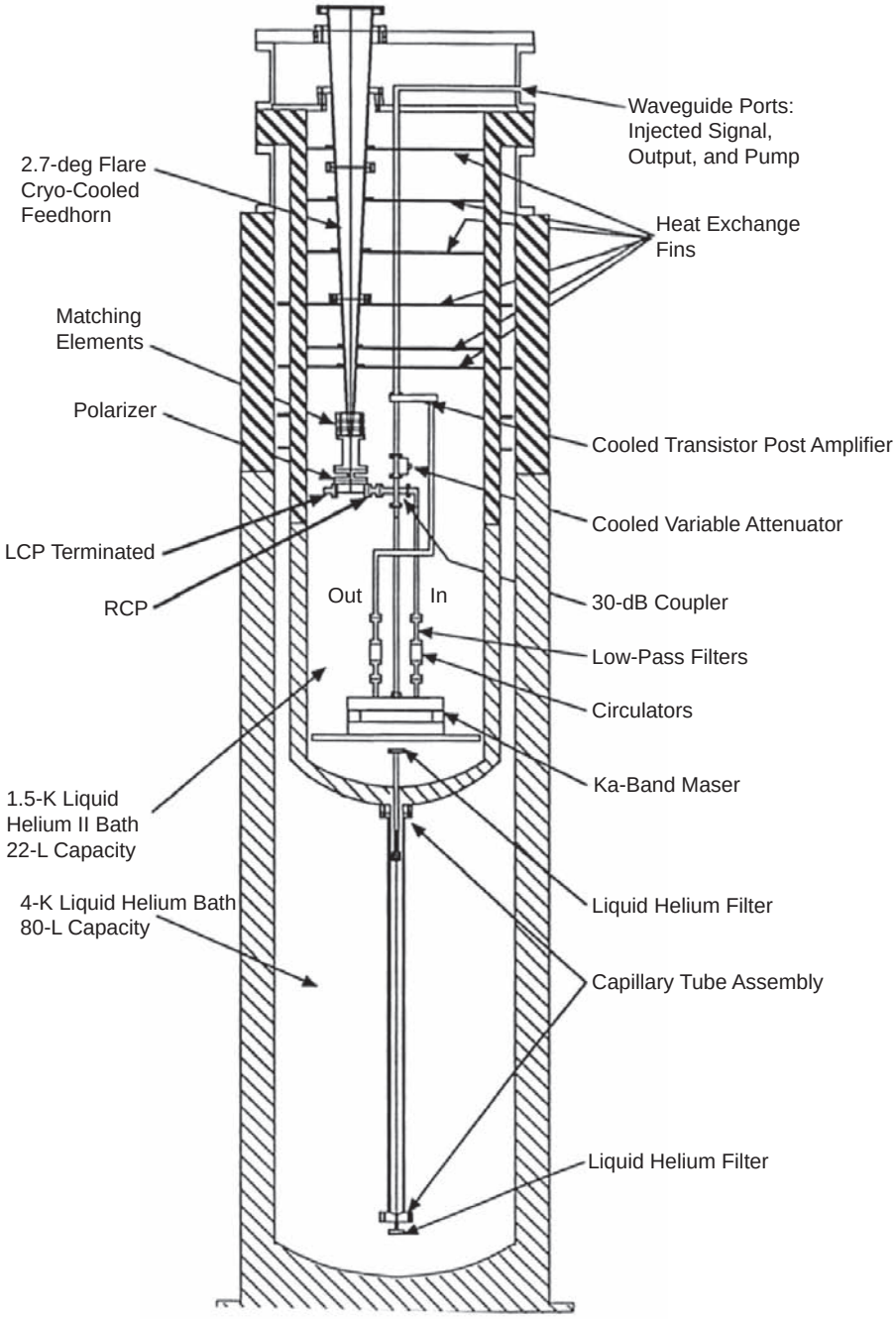


Fig. 4-7. Continuously cooled liquid-helium dewar with cooled feedhorn and Ka-band maser.

in the liquid-helium II bath is known precisely, based on the vapor pressure above the liquid (3.63354 mm at 1.5 K and 12.56124 mm at 1.8 K for example) (0.484 and 1.675 kPa, respectively).

4.6 Closed-Cycle Helium Refrigerators

The DSN's first CCR was a model 200 Cryodyne®. All of the CCRs used to cool operational DSN ruby masers between 1961 and the present time used the combination of a JT liquefaction process with a GM cycle CCR.

The JT process cools a gas passing through an expansion valve only when the gas is below the inversion temperature. For helium gas, the inversion temperature is near 40 K. Helium heats when passing through an expansion valve (called the JT valve) at temperatures above 40 K. Typical pressure on the high side of the JT valve in a DSN CCR is at 20 atmospheres (2.03 MPa). The low-pressure side of the JT valve is at 1.2 atmospheres (0.122 MPa).

Helium entering a JT valve at 300 K and 20 atmospheres (2.03 MPa) exits the JT valve at a pressure of 1.2 atmospheres (0.122 MPa) and at a temperature between 301 K and 302 K. Helium entering a JT valve at 40 K (close to the inversion temperature) and at a pressure of twenty atmospheres (2.03 MPa) exits the JT valve at a pressure of 1.2 atmosphere and at a temperature of 40 K.

Helium at 20 atmospheres (2.03 MPa) and 300 K enters the CCR and must be cooled to about 15 K to enable a JT cooling effect of about 2.5 K. The enthalpy of helium at 20 atmospheres (2.03 MPa) is about 1570 joules per gram (J/g) at 300 K and about 80 J/g at 15 K. The JT flow of DSN 4.5 K CCRs is about 0.12 grams per second (g/s) of helium. Removal of about 179 joules per second (J/s) at 15 K would be required to cool the helium to 15 K. This is equivalent to cooling a 179-W heater, leading to the use of counter-flow heat exchangers.

A gas stream flowing from 300 K to 15 K, and then returning from 15 K to 300 K, can use exhausting cold gas to cool the incoming gas stream. This can be accomplished by connecting the two separate sets of tubing used to contain the two streams in a manner that enables heat transfer from the incoming to the outgoing gas. The simplest form of a counter-flow heat exchanger connects two long tubes with solder so the thermal path between the two tubes is very short. The tubes are long, because conductive transfer from 300 K to 15 K must be minimized.

The short heat-transfer path through metal between the two passages of a counter-flow heat exchanger is easy enough to accomplish. The main problem is transferring heat from the gas to the metal walls of the passage. The thermal conductivity of helium is very low in comparison to metals, as shown in Table 4-3. The geometry of an efficient counter-flow heat exchanger becomes complex. The complexity is increased by the requirement for minimizing the pressure drop, especially in the low-pressure side of the heat exchanger.

A heat exchanger with 100-percent efficiency would result in a heat load of about 2.2 W on the 15-K station of the GM stage used to cool the JT flow in a 4.5-K CCR. This heat load is caused by the enthalpy difference between 300 K and 15 K of the incoming helium at a pressure of 20 atmospheres (2.03 MPa) and the exiting helium at a pressure of 1.2 atmospheres (0.122 MPa). Counter-flow heat exchangers with 96-percent efficiency were considered to be excellent for use with CCRs having a 1-W cooling capability at 4.5 K. The added heat load to the 15-K station due to heat exchanger inefficiency between 300 K and 15 K would be about 7 W. The use of one or two intermediate stages of GM cooling between 300 K and 15 K eliminates the excessive heat load on the 15-K stage of the GM cooler. An intermediate stage at 70 K in the JPL CCR reduces the heat load on the 15-K stage to about 3 W.

Helium entering a JT valve at 20 K and a pressure of 20 atmospheres (2.03 MPa) exits the JT valve at a pressure of 1.2 atmospheres (0.122 MPa) at a temperature of about 18.5 K. Helium entering a JT valve at 15 K and at a pressure of 20 atmospheres (2.03 MPa) exits the JT valve at a pressure of 1.2 atmospheres (0.122 MPa) at a temperature of about 12.5 K. Helium entering a JT valve at 10 K and at a pressure of 20 atmospheres (2.03 MPa) exits the JT valve at a pressure of 1.2 atmospheres (0.122 MPa) at a temperature of about 6.5 K.

Increasing the JT high-side pressure from 20 to 30 atmospheres (2.03 to 3.04 MPa) gives an improvement of about 0.5 K to the cooling effect when the temperature is 15 K. A reduction of the low-side pressure from 1.2 atmospheres to 0.5 atmospheres (0.122 to 0.051 MPa) improves the cooling effect by about 1 K when the high-side temperature is 15 K. The choice of 1.2 atmospheres (0.122 MPa) for the low-side JT pressure was made to maintain a higher-than-atmospheric pressure in the JT return line. This choice is based on the type and length of the gas lines going from the helium compressor located near the base of the antenna to the CCR mounted near the secondary focal point of a large cassegrainian antenna. The gas lines are typically about 100 m in length and use flexible sections to traverse the antenna axes.

Loss of helium pressure in the storage tank of a CCR compressor indicates the presence of a leak. Positive pressure in all helium lines prevents massive air contamination when a leak is large. The 1.2-atmosphere (0.122-MPa) helium pressure in the JT return line does not eliminate the danger of air entering the system in cases of small leaks or when using certain types of seals. Small leaks and O-ring seals made of rubber allow air to diffuse into the helium-filled line. The partial pressure of helium is less than that of air, so the advantage of having the JT return line above atmospheric helium pressure to avoid contamination is compromised. The partial pressure of a gas is the measure of thermodynamic activity of the molecules. Gases flow from a region of higher partial pressure to one of lower pressure; the larger this difference, the faster the flow.

Suitable liquefaction in the 4.5-K station following the JT valve is obtained by using a counter-flow heat exchanger with high efficiency between the CCR stage at 15 K and the JT valve. A high-efficiency counter-flow heat exchanger minimizes heat transfer between the two heat stations connected to ends of the heat exchanger. Removing heat from the 4.5-K station of a DSN CCR during the initial cool-down process requires a way of transferring heat around (bypassing) the counter-flow heat exchangers in the JT loop.

An attempt to explain the cooling and liquefaction that occurs as result of the JT effect begins here with the 15-K GM station and the 4.5-K JT station, both at 15 K during the CCR cool-down process. Helium at a pressure 20 atmospheres (2.03 MPa) and cooled to 15 K enters a high-efficiency counter-flow heat exchanger and travels to the JT valve. When reaching the JT valve, the helium expands through the valve to a decreased pressure of 1.2 atmospheres (0.122 MPa). The helium cools from 15 K to 12.5 K during the expansion. The 12.5-K gas removes heat from the final stage and travels through the low-pressure side of the counter-flow heat exchanger, removing heat from the incoming stream of helium. The heat exchanger cools the incoming stream from 15 K to a lower temperature at the high-pressure entrance to the JT valve, depending on the amount of heat being removed from the 4.5-K station. The helium reaching the JT valve, now at a temperature below 15 K, expands through the valve and drops to a temperature lower than 12.5 K. This progressive process continues and lowers the helium temperature until the expansion through the JT valve produces some liquid helium at 4.5 K.

Parasitic heat leaks and counter-flow heat exchanger inefficiency affect the percentage of liquid helium produced. The ideal case, with no external heat leak and 100-percent heat exchanger efficiency, would not result in 100-percent liquefaction. The 15-K temperature at which the helium enters the final JT loop counter-flow heat exchanger limits the liquefaction percentage. The liquid-vapor helium mixture flows through a series of perforated copper disks imbedded in the copper 4.5-K heat station and then into the low-pressure side of the counter-flow heat exchanger. Heat transferred to the final stage through supports, radiation, heat exchanger inefficiency, microwave transmission lines, and wiring is about 0.4 W. The net refrigeration capacity (reserve capacity) without the application of additional heat is about 1 W.

The 4.5-K JT stage produces liquid helium at a rate that is sufficient to fill a large part of the low-pressure side of the counter-flow heat exchanger with the 4.5-K liquid-vapor mixture. Without the application of additional heat to the 4.5-K stage, the boiling liquid and vapor mixture of helium fills more than 3/4 the length of the heat exchanger. The efficiency of the counter-flow heat exchanger between the 15-K stage and the 4.5-K stage drops as the active length of the heat exchanger is shortened. The active part of the counter-flow heat exchanger is that section nearest the 15-K station that contains only helium gas (no liquid) in the low-pressure side. The heat transferred from the 15-K

station through the shortened active section of heat exchanger to 4.5-K point increases until the final stage can no longer produce additional liquid. The temperature of the helium at 20 atmospheres (2.03 MPa) on the high-pressure side of the final heat exchanger is at 4.5 K for about 3/4th the length of the heat exchanger when no additional heat is applied to the 4.5-K station.

Application of sufficient heat to the 4.5-K JT stage empties the liquid–vapor mixture from the low-pressure side of the final counter-flow heat exchanger. The heat load is then adjusted to a level that converts all of the liquid helium produced to vapor (about 1 W). This amount of heat represents the reserve capacity of the JT stage. A resistive heating element is mounted on each stage of the refrigerator for the purpose of measuring reserve capacity. With the cold end of the heat exchanger at 4.5 K, and with 96-percent heat exchanger efficiency, the helium gas leaving the heat exchanger at the 15-K station, at 1.2 atmospheres (0.122 MPa) would be 14.6 K, about 0.4 K less than the temperature of the incoming gas at 20 atmospheres (2.03 MPa). The enthalpy of helium gas at 15 K and 20 atmospheres (2.03 MPa) is about 80 J/g. The enthalpy of helium gas at 14.6 K and 1.2 atmospheres (0.122 MPa) is about 92 J/g. The gas leaving the heat exchanger has an enthalpy that is about 12 J/g higher than the gas entering, thereby removing heat from the final stage. The helium flow through the JT loop of a DSN 4.5-K 1-W CCR is about 1.5 standard cubic feet per minute (scfm) or 0.12 grams/second (g/s). This shows a total heat removal of about 1.44 W at the 15-K end of the final counter-flow heat exchanger. Steady-state operation of the JT loop producing liquid helium at 4.5 K does not involve a temperature change at the JT valve. Heat removed from the 4.5-K station converts the liquid to vapor.

GM-cycle Cryodynes® were used to cool the JT loop to temperatures of 100 K, 35 K, and 15 K in the case of the Model 200 and 210 Cryodynes®, and to 70 K and 15 K in the case of the JPL-developed 4.5-K CCR [2]. The GM cycle was considered to be the most efficient and reliable combination available with the appropriate capacity at the time the ADL Model 200 and 210 Cryodynes® and JPL CCRs were developed. The combination of efficiency, reliability, and appropriate capacity is emphasized here. There may have been more efficient cryogenic refrigeration cycles that were not sufficiently reliable and there may have been more reliable cycles that were not sufficiently efficient. Large cryogenic refrigerators with hundreds of watts of cooling power at 4.2 K can be very efficient and reliable. These large systems do not scale down effectively to the size needed for the DSN's LNA applications.

The GM-cycle Cryodynes® and JT loops can use a helium compressor that is located a long distance from the cryogenic refrigerator. This is convenient for the DSN implementation on large antennas, as shown in Fig. 3-6. The compressor provides about 1.6 g/s of helium at a pressure of about 21 atmospheres (2.13 MPa). Most of this gas stream is used by the GM

Cryodynes®, returning to the compressor at pressures between 5 and 8 atmospheres (0.507 and 0.811 MPa), depending on the type of system used.

Figure 4-8 shows a schematic diagram of the JPL CCR with a two-stage GM refrigerator and a JT loop. The "supply" gas stream from the compressor is split at the refrigerator, with about 1.48 g/s at a pressure of 21 atmospheres (2.13 MPa) going to the CCR drive unit. About 0.12 g/s goes to a pressure regulator and then to the JT loop inlet at a pressure of about 20 atmospheres (2.03 MPa).

The GM cycle uses isentropic expansions of helium in combination with displacers and regenerators to achieve low temperatures. Helium at high pressure (about 21 atmospheres, 2.13 MPa) enters through the inlet valve when

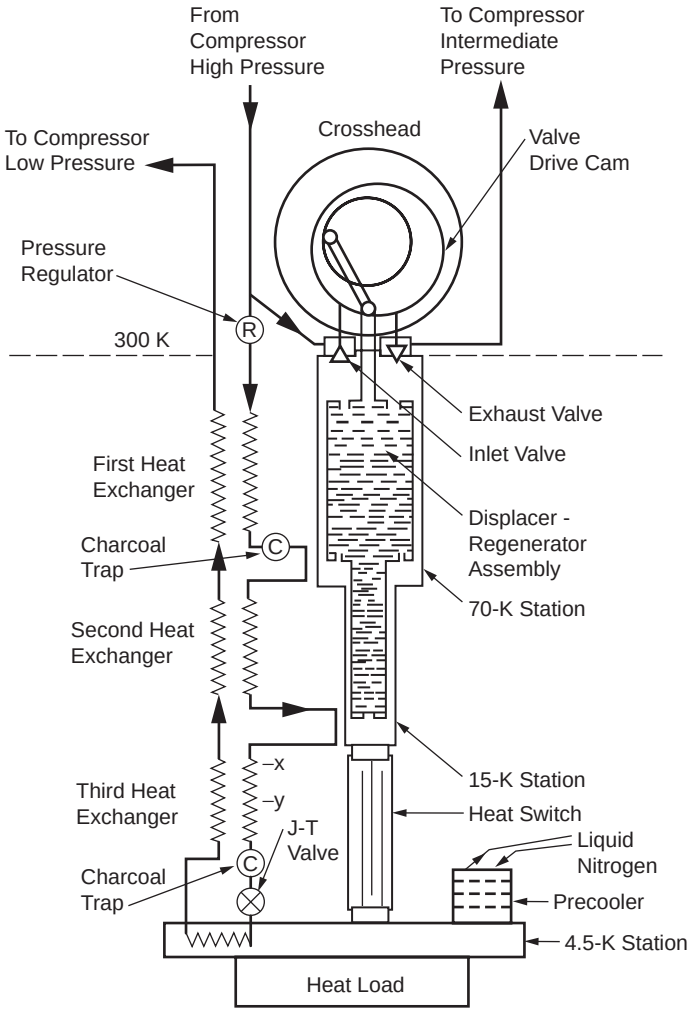


Fig. 4-8. JPL CCR schematic diagram.

the displacers are furthest away from the crosshead. The volume at the cold ends (shown as the bottom ends in Fig. 4-8) of the cylinder containing the displacer and regenerator assemblies are near minimum at this point in time. The volume above the regenerators is near maximum at this point in time. Pressurization of this volume causes heating, and the heat is later stored at the upper or warm ends of the regenerators.

The regenerators are made of materials having high specific heat in the temperature ranges appropriate to each stage. The goals of the regenerator designs are to (1) store heat, (2) transfer heat quickly between the helium and the regenerator material, (3) minimize the heat transfer from one end to the other end of the regenerator, (4) minimize the dead volume (void space) thus maximizing the filling factor of the regenerator material, and (5) minimize the restriction through the regenerator material.

After pressurizing the volume above the regenerators, including the entire volume of the displacers and regenerators, the displacers are moved towards the crosshead, causing the high-pressure helium gas to flow through the regenerator matrices into the increasing volume at the cold end of the regenerators. After the volume at the cold end of the regenerators is maximized and filled with high-pressure helium, the inlet valve is closed, and the exhaust valve is opened. The high-pressure gas in the now large volume at the cold end of the regenerators experiences an isentropic expansion. The cooled gas passes back through the regenerator matrices, removing the stored heat in the process. The displacer is then moved away from the crosshead, again minimizing the cold-end volume and causing the last of the cold gas, now at the intermediate pressure, to pass through the regenerator and further cool the cold end of the regenerators. Each repetitive cycle causes a significant drop in the temperature at the cold end of the regenerators until the ultimate low temperatures are reached. In the two-stage GM system of the JPL CCR, these temperatures are at about 70 K (1st stage) and 15 K (2nd stage) with the typical heat load of the JT loop, radiation shield, and conductive heat transfer through supports and microwave transmission lines. A radiation shield made of copper is attached to the 70-K station, and it surrounds the components at temperatures below 70 K.

The Model 200 and 210 Cryodynes® used a three-stage GM system with the JT loop. The temperatures of the three stages were at about 100 K, 35 K, and 15 K. The regenerator material in the 100-K stage, and in the 70-K stage of the JPL CCR, was brass screen. The regenerator material in the 35-K stage and 15-K stage of the GM systems was lead, in the form of small spheres. Three Micarta displacers of the Model 200 and 210 Cryodynes® were contained in independent stainless-steel cylinders, with regenerators in separate stainless-steel cylinders external to the displacers, connected with copper or stainless-steel tubing. The concentric assembly of the model 340/350 Cryodyne® used in the JPL CCR is shown in Fig. 4-8 consisted of Micarta displacers that contained the regenerator material. Seals, preventing helium gas flow around the

displacers, were made of Teflon with 25-percent fiberglass. These seals, not shown in Fig. 4-8, were located at the warm ends of each displacer. Schematic diagrams of the Model 200 and 210 Cryodynes® are not shown, but significant differences between these early DSN cryogenic refrigerators and the much more successful JPL CCR seem worthy of mention. Reliability problems and failures with the early systems resulted in the development of repairs, modifications, and upgrades that helped provide knowledge and insight, and leading to the development of the 4.5-K JPL-CCR that has been used in the DSN since 1966, more than forty years.

A lesson about early success with new equipment was learned. Trouble-free performance with the Model 200 Cryodyne® was demonstrated for 5000 hours on the 26-m research antenna at DSS-13. Subsequent experience with model 210 Cryodynes® on operational DSN antennas was less successful. Helium compressors with inadequate helium purification allowed migration of compressor lubricant into the CCRs. LB 400 Union Carbide compressor lubricant traveled from the compressors to the refrigerators in helium high-pressure supply lines without wetting the lines. Liquid appeared in the refrigerator return lines and in the intermediate pressure areas of the drive units. The regenerators in the GM CCRs were acting as agglomererators, clustering tiny oil droplets, which traveled like vapor with the helium gas through the supply lines. This condition indicated improper assembly of the agglomerators in the compressors used for the early-model 210 Cryodynes®.

The helium compressors used modified "3-HP" (2.24-kW) Copeland Freon® compressors. These modified Copeland Freon® compressors each draw about 7 kW electrical input power in this application. Coolant is needed to reduce the temperatures caused by compressing helium from a pressure of 1.1 atmospheres (0.11 MPa) to the intermediate pressure and then to the 21-atmosphere (2.13-MPa) supply pressure. A large quantity of lubricant is circulated through the compressor to keep the helium temperature below 80 C. The high-pressure helium gas and lubricant mixture is cooled to ambient temperature by the compressor's air-cooled heat exchangers. A properly assembled series of agglomerators causes the lubricant droplets to collect and flow to the bottom of the agglomerators. The lubricant is then returned to the compressor through a series of filters or screens and orifices, and the lubricant-free helium is sent on to a charcoal and molecular-sieve filter for further purification.

Unfortunately, the pipe thread connectors used to attach the sintered metal agglomerators of the early model 210 Cryodynes® compressors were not tightened sufficiently during assembly. The first-stage agglomerators fell apart during the early months of use. Second-stage agglomorators, made of fine glass wool (an artificial substitute for fine lamb's wool), were not packed with sufficient density. "Oil-carry-over" problems plagued these systems during the early days.

Other problems complicated the situation. Cryodyne® drive unit gear failures occurred within a few hundred hours of operation. This failure mode was not seen in model 200 Cryodyne® during the first 5000 hours of operation. The drive mechanism for the displacers used a very small gear at the end of the motor shaft to achieve a high reduction ratio. The small gear drove a large gear and a complex gear arrangement to approximate a square-wave displacer motion at 105 revolutions per minute (rpm). The load on the small gear caused rapid wear and early failures. The load on the gear train was aggravated by accumulations of lead and Micarta® dust mixed with compressor lubricant, forming a sticky mixture on the displacers. A "fix" to eliminate the source of lead dust was to plate the lead spheres with a shiny hard-surface metal. This "fix" added another contaminant; the plating flaked off of the lead spheres during the thermal cycling and vibration of operation.

The tight implementation schedule needed to support the first Mariner-Mars mission launched in November 1964 forced repairs, modifications, and upgrades to be done at JPL and at the DSN Stations, rather than sending the Cryodynes® back to ADL. With the support of ADL personnel, Walt Higa's group completed the work enabling the use of these systems for the Mariner-Mars mission. Much was learned in the process. Ervin (Erv) R. Wiebe at JPL was key to finding solutions to the problems. For example, Erv modified a drive unit to achieve a sinusoidal displacer motion and eliminated the need for the complex "square-wave" gear drive mechanism. The speed was slowed to simulate performance that would result from the use of a 72-rpm direct-drive motor. Capacity measurements showed performance equivalent to the square-wave drive that was previously thought to be ideal for the displacer motion. Several compressor and drive-unit fixes, modifications, and upgrades (including the installation of 72-rpm direct-drive motors) were implemented in a short time period at JPL and at the DSN tracking stations. The six TWM CCR systems, implemented in time to support the Mariner-Mars encounter in July 1965, performed without failure during the Mars picture-playback sequence in July and August 1965.

Walt Higa and Erv Wiebe were the leaders in the development of the JPL CCR that was first used in the DSN on the 64-m antenna at DSS-14 in 1966. Professor William Gifford consulted for our group during that development period. The team that made this development successful included Rex B. Quinn, a precision machinist, microwave technician and engineer, whose contributions to maser development, ultra-low-noise systems, and cryogenic refrigerators continued for decades. Robert S. (Bob) Latham is another of the team members that deserves mention; Bob's guidance in dealing with operational situations was most valuable.

The ADL Model 340 Cryodyne® was chosen to pre-cool a JT loop developed by Higa and Wiebe [6]. The efficiency of counter-flow heat exchangers developed by Higa and Wiebe resulted in a 1-W capacity at 4.5 K at

a JT flow of about 0.12 g/s. The first and second stages of the GM Cryodyne® operated at 70 K and 15 K, respectively, in combination with the JT loop. The use of a fixed JT valve simplified the operation and eliminated a failure mechanism encountered with the adjustable JT valve used in Model 200 and 210 Cryodynes®. The cool-down time for the JPL CCR was shortened and simplified through the use of a hydrogen-filled thermal switch invented by Erv Wiebe. The novel switch eliminated the need for a heat-exchanger bypass valve used during the cool-down process in model 210 Cryodynes®.

Figure 4-9 shows an S-band maser package as used on the 64-m antenna in 1966, with inserts showing key features. The upper right insert (a) shows a cut-away section of the inlet and outlet connections in a counter-flow heat exchanger. Details of the inner construction are shown in Fig. 4-10. Flexible phosphor-bronze "hose" with helical convolutions is used between snug-fitting inner and outer stainless-steel tubes. The region between the inner stainless-steel tube and the corrugated phosphor-bronze hose provides one passage through the counter-flow heat exchanger. The region between phosphor-bronze hose and the outer stainless-steel tube forms the other passage. The two gas streams, flowing in opposite directions, transfer heat from one to another. The thin wall between the two streams enables efficient heat transfer between the two streams, but the long length of the heat exchanger assembly minimizes heat transfer from end-to-end. The helical convolutions cause turbulence, helping to transfer heat through the helium and maximizing the efficiency of the counter-flow heat exchanger.

The restriction through the passage at low pressure affects the pressure of the JT loop at the liquid stage of the refrigerator. For this reason, Higa suggested another design using phosphor-bronze hose with helical convolutions. The geometry of this design is shown in Fig. 4-11. The low restriction in this design enabled increased JT flow, and this was later used to increase the capacity of DSN CCRs.

Insert (b) in Fig. 4-9 shows an exploded view of a thermal switch (three parts on the left side), a charcoal-filled filter (upper right), and the fixed JT valve. The small needle near the center of the insert is part of the fixed JT-valve. This valve in the JT loop, dropping the JT pressure from about 20 atmospheres to 1.2 atmospheres (2.03 to 0.122 MPa, respectively) consists of a hypodermic needle having an inner diameter of 0.007 inch (in.) (0.018 cm) partially filled with a stainless-steel wire having an outer diameter of about 0.006 in. (0.015 cm). The length of the tubing and wire is about 2 cm. The diameter of the wire can be reduced to decrease the restriction through the valve to provide JT flow of about 0.12 g/s during normal operation.

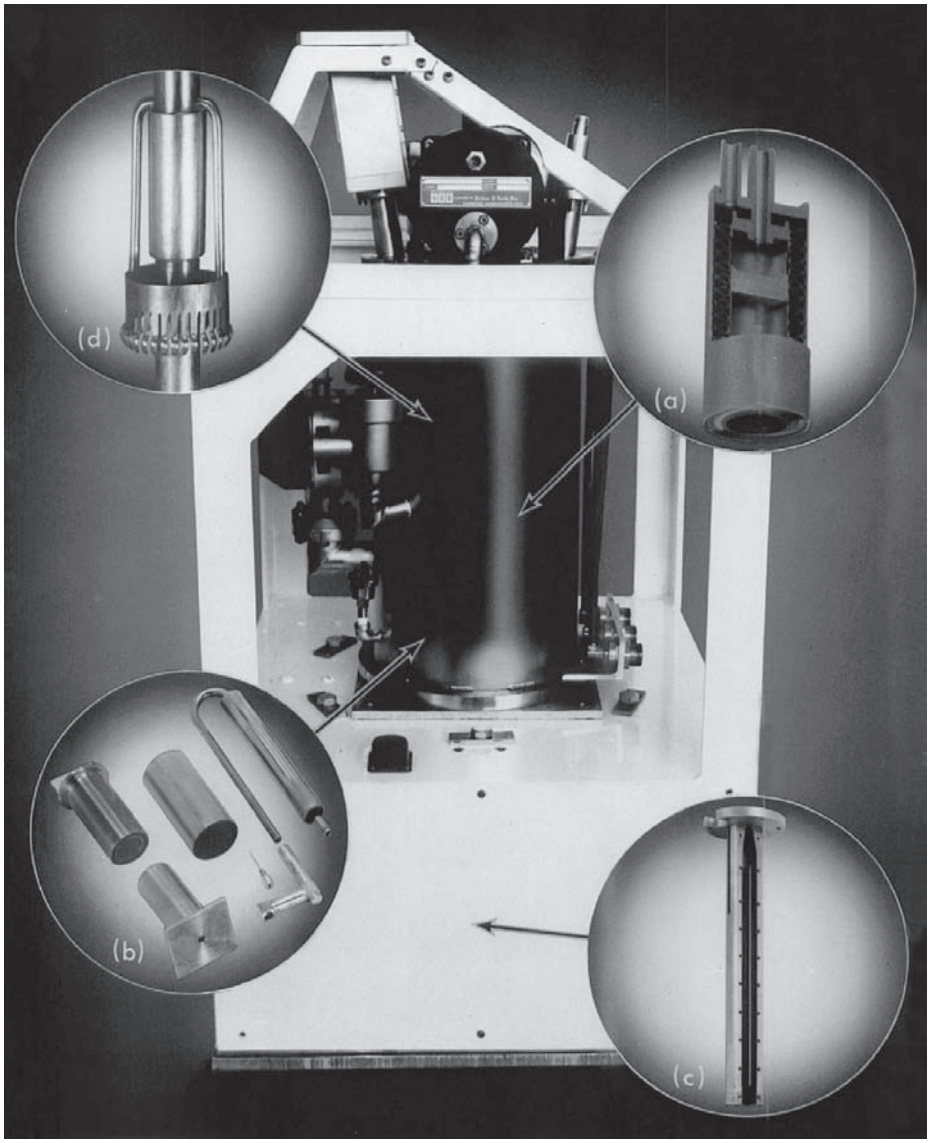
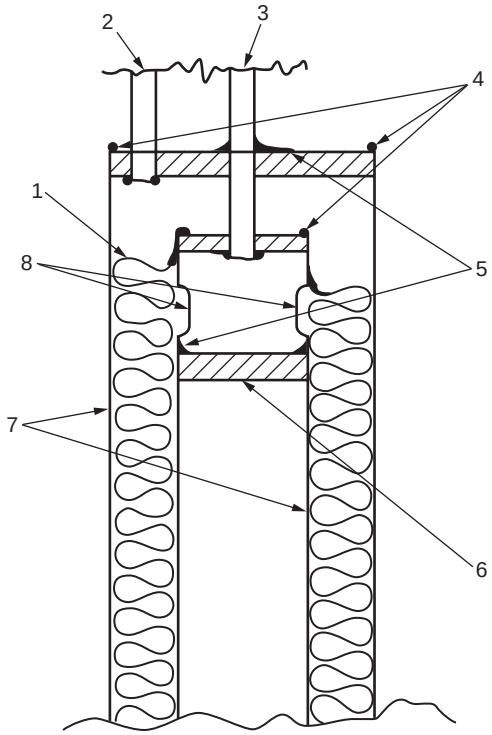


Fig. 4-9. S-band JPL TWM/CCR package with key features shown in expanded views (a) cut-away section of the inlet and outlet connections in the counter-flow heat exchanger; (b) charcoal-filled filter to the upper right of the insert, the JT valve housing and hypodermic-needle-part of the fixed JT-valve to the lower right, and the thermal switch parts (in the center, on the left side, and the bottom of the insert; (c) an S-band TWM comb structure (discussed in Chapter 3); and (d) a quarter-wavelength thermal conductor between the center conductor and outer conductor of an S-band coaxial input transmission line, transparent to the microwave signal.



Notes:

1. Helium gas or liquid flows in opposite directions through phosphor–bronze flexible hose with helical convolutions providing two long spiraling gas passages, one on either side of the convolutions. (Note: annular convolutions cannot be used for this application.)
2. Connection to passage between outer stainless steel tube and phosphor–bronze hose.
3. Connection to passage between inner stainless steel tube and phosphor–bronze hose.
4. Stainless steel welds shown with dot at joint.
5. Silver soldered (brazed) joints used to seal: plug (6) into ends of inner stainless steel tube, phosphor–bronze hose to inner stainless steel tube, and connection to inner passage (3) at point passing through the heat exchanger end caps.
6. Plug located at each end of inner stainless steel tube.
7. Thin-wall (typically 0.025 cm) stainless steel tubes positioned inside and outside of the phosphor–bronze hose with dimensions providing a snug fit on either side of the convolutions.
8. Short slots near ends of the inner stainless steel tube to allow gas to pass from the connection tube (3) to the inner spiraling passage between the inner stainless steel tube and the phosphor–bronze convolutions.

Fig 4-10. Construction details of counter-flow heat exchanger, which transfers heat from the incoming stream at high pressure to the exhausting stream at low pressure.

The fixed JT valve replaced an adjustable JT valve used in the model 200 and model 210 Cryodynes®. The adjustable JT valve of those systems used a long tapered stainless-steel needle in an orifice that was about 0.089 cm in diameter. The tapered needle was moved through the orifice by a long shaft attached to an external micrometer drive mechanism at an ambient surface of

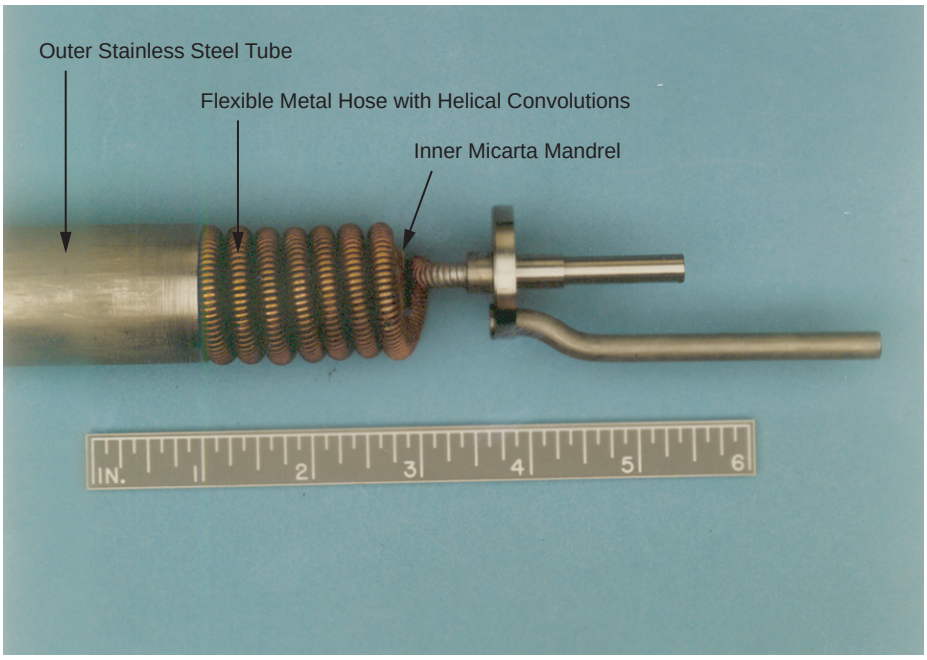


FIg. 4-11. Improved counter-flow heat exchanger (low restriction) (partially disassembled view).

the Cryodyne®. The micrometer adjustment for JT flow was manually controlled. High JT flow during CCR cool-downs was used to help to transfer heat from the final stage of the CCR to the upper stages. A manually controlled bypass valve sent the helium from the low-pressure side of the JT valve around the output pa ssages of the second-, third-, and final-stage counter-flow heat exchangers. This arrangement required the attendance of an operator at the antenna-mounted assembly during the CCR cool-down process. Improper adjustments of the valves sometimes caused the cool-down process to fail.

The CCR final-stage temperature following a successful cool-down was dependent upon the JT flow because the vapor pressure at the final stage was dependant on the restriction through the low-pressure side of the counter-flow heat exchangers. Closing the JT valve reduced the JT flow and the final-stage temperature. Reducing the JT flow also reduced the final-stage refrigerator capacity. The temptation of lowering the final-stage temperature often induced an operator to c lose the JT valve completely. The sudden CCR warm-up that resulted could be remedied by opening the JT valve. Unfortunately, the stainless-steel needle sometimes stuck in the stainless-steel orifice. The galling effect of forcing the two stainless-steel surfaces together was like a welding process. Major surgery was required to repair several CCRs.

The combination of the fixed JT valve and the hydrogen-filled thermal bypass switch (also shown in insert b) eliminated the need for an adjustable JT valve and a bypass valve. The use of a wire in a capillary tube as a JT valve was suggested by automotive air conditioning specialist Richard W. Clauss.

Higa and Wiebe summarized the first ten years of the use of the JPL CCRs in a presentation at the Cryocooler Applications Conference (sponsored by the National Bureau of Standards in Boulder, Colorado in October 1977) with a presentation entitled, "One Million Hours at 4.5 Kelvin" [6]. Higa's team was responsible for all aspects of the TWM/CCR system. Higa's plan was to avoid controversies between specialized groups by having one group that was responsible for all the necessary technologies, including cryogenics, microwaves, ferrite isolators, quantum physics, crystallography, high-vacuum systems, power supplies, monitor and control equipment, and noise temperature measurements. The entire team learned about the whole system, not just individual areas of expertise. Higa's plan and approach were successful.

The team changed with time. Cancer caused Erv Wiebe to slow down in 1976 and Frank McCrea took over the task of producing the 4.5-K CCRs for the new Block II X-band masers. These masers were needed to reduce the system noise temperature of the DSN 64-m antennas for the upcoming Voyager-Saturn encounters in 1980 and 1981. Frank McCrea devised and patented a choked X-band (WR125) waveguide input assembly for coupling signals into a cryogenically cooled LNA. This device (U. S. Patent 4,215,327) [7] enabled cooling the input waveguide to the first-stage temperature near 70 K within a small fraction of a millimeter from the ambient vacuum window. Coaxial stainless-steel support tubes surrounding the cold copper waveguide were connected in cascade to provide a folded low-conduction path from the room-temperature vacuum housing to the cold side of the choke joint in the waveguide. The new low-noise waveguide input assembly provided the major part of a 7-K noise temperature reduction, enabling increased data rates from Voyagers I and II at Saturn in 1980 and 1981.

Walt Higa retired and Sam Petty, an experienced engineer and group leader, became the group's supervisor in 1981. Sam continued leading the team for another 24 years. Block II-A masers systems and cooled high-electron-mobility transistor (HEMT) systems were implemented to support Voyager encounters at Uranus and Neptune. A cooperative effort with the National Radio Astronomy Observatory (NRAO) in the mid-1980s resulted in the production of low-noise X-band HEMTs for use at the NRAO Very Large Array (VLA) during the Voyager-Neptune encounter. A variety of additional low-noise amplifier systems, cooled by 4.5-K JPL CCRs, 15-K GM CCRs, and dewars operating at 1.5 K and 1.8 K were implemented in the DSN between 1981 and 2006.

Cooperative maser and CCR work with Craig Moore and Howard Brown of the NRAO in the mid-1970s contributed to the development of improved

masers and more powerful refrigerators. Mike Britcliffe and Ted Hanson continued this work, increasing the capacity of the 1-W 4.5-K JPL CCR to 2 W in the 1980s. The 1965 version of the CCR was modified by installing the improved heat exchanger design shown in Fig. 4-11 and increasing the JT flow to about 0.18 g/s. The details of the 2-W 4.5-K CCR and the compressor with greater capacity are reported in Telecommunication and Data Acquisition Progress Reports [8,9]. Britcliffe and Hanson also investigated and developed a prototype CCR system using two JT expansion stages, achieving 700-mW capacity at 2.5 K [10]. Figure 4-12 shows a simplified flow diagram of the "5-HP" (3.73-kW) compressor.

The concentric two-stage model 340 Cryodyne® developed by ADL and marketed in 1965 was originally referred to as a parametric amplifier (par-amp) cooler. Parametric amplifiers were used as LNAs during that time period and achieved fairly low-noise temperatures (near 50 K) when cooled to 15 K or 20 K. Use of the model 340, and later the model 350 Cryodynes®, in combination with a JT loop to provide cooling at 4.5 K was not anticipated by the staff at ADL who developed the concentric two-stage Cryodynes® (Fred F. Chellis, Walter H. Hogan, Robert W. Stuart, and others). Their unique and efficient design soon found far greater use in high-vacuum systems (cryo-pumps) than for cooling LNAs. Microwave field-effect transistors after 1975 and then high-electron-mobility-transistors (HEMTs) after 1983 showed significant performance improvements at cryogenic temperatures, and the two-stage Cryodynes® were used to cool transistor amplifiers in many radio-astronomy and space-communications facilities around the world. Implementation and operation of masers continued between 1975 and 1995 because of the low-noise temperature advantage during those years.

Emphasis on better reliability and lower costs after 1995 resulted in implementation of fewer maser systems and more HEMTs, taking advantage of commercially available two-stage GM CCRs. The simplicity of the two-stage concentric GM CCR became even more attractive after 1995 when superior regenerator materials enabled 4-K performance. The lead regenerator material had limited the useful minimum temperature of the second stage to about 15 K. James S. Shell of JPL purchased early models of the two-stage GM cycle 4-K CCRs built by Leybold Cryogenics North America (Hudson, New Hampshire) and by Sumitomo (Tokyo, Japan), beginning evaluations in February 1997. The Sumitomo two-stage 4-K CCRs were selected for use in the DSN. Operation of an X-band maser cooled by a commercially available two-stage Sumitomo SRDK-415 GM CCR was demonstrated [11]. A Sumitomo Heavy Industries two-stage GM CCR [12] was recently used at JPL to cool a dual-channel X-band HEMT to 6.5 K; the measured input noise temperature was 4.4 K at the ambient interface to the amplifier package [13].

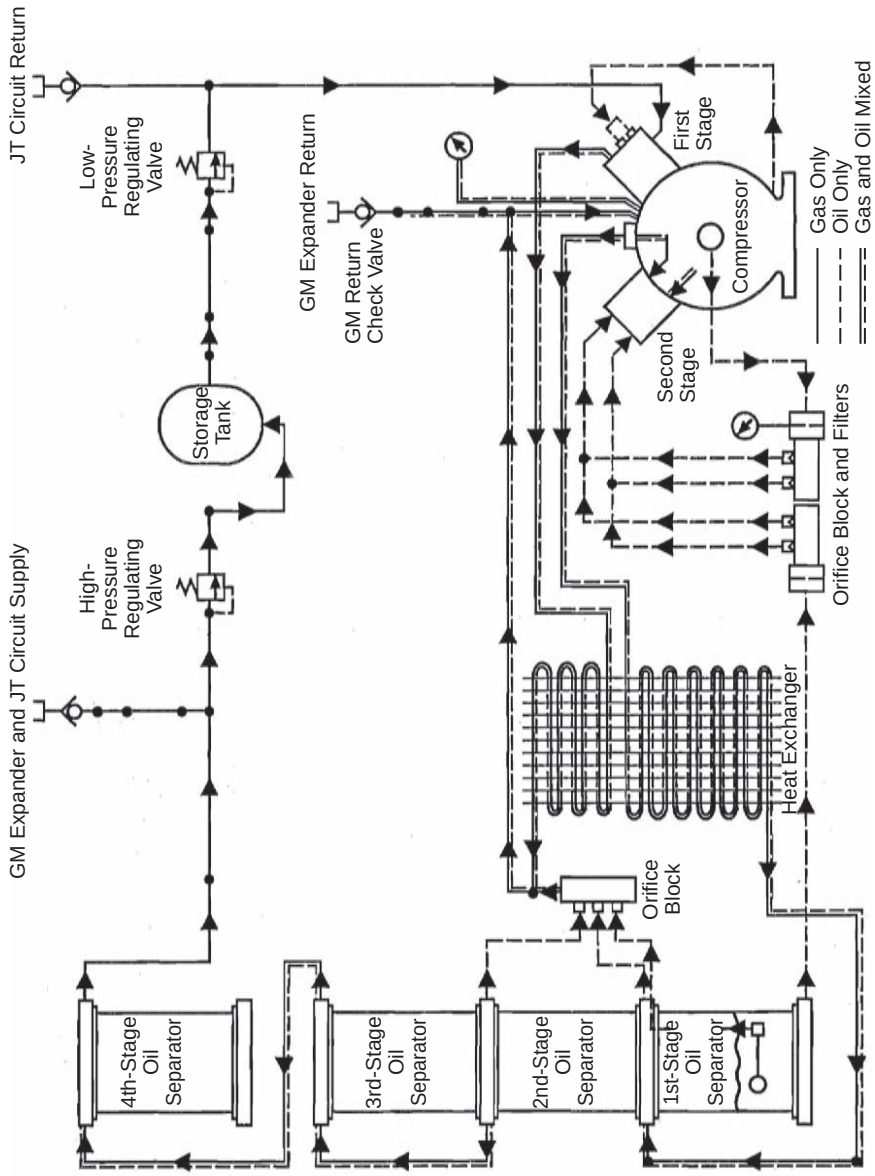


Fig. 4-12. Simplified flow diagram of the "5-HP" (3.73-W) compressor. (This modified Freon® compressor draws about 7 kW electrical input power in this application.)

4.7 Conclusion

Many at JPL contributed to the development of low-noise systems for the DSN. Robertson (Bob) Stevens was our section leader in 1959, then our division manager, and then the chief engineer for JPL's Telecommunications and Data Acquisition Office from 1979 to 1992. Bob would not let us give up on the goal of improving sensitivity through the reduction the DSN's system noise temperature, even if the reduction was only a fraction of a kelvin.

The development and use of cryogenic refrigeration systems for cooling LNAs and feed system components to improve the sensitivity of DSN receiving systems for deep space mission support was essential and cost effective. The development and implementation of even more advanced cryogenic systems is now practical and logical. Commercially available 4-K GM CCRs can be used to liquefy helium and to intercept heat leaks to a final stage at a temperature near 1.5 K. These advanced cryogenic refrigeration systems will enable the use of LNAs having noise temperatures within 1 K of the quantum limit at all DSN frequencies.

References

[1] R. B. Scott, *Cryogenic Engineering*, Princeton, New Jersey: D. Van Nostrand Company, Inc., 1959.

[2] W. H. Higa and, E. Wiebe, "A Simplified Approach to Heat Exchanger Construction for Cryogenic Refrigerators," *Cryogenic Technology*, vol. 3, pp. 47–51, March/April 1967.

[3] American Institute of Physics Handbook, Second Edition, McGraw-Hill Book Company Inc. New York, New York (Library of Congress Catalog Card Number 61-16965), 1963.

[4] *Handbook of Chemistry and Physics*, Forty-First Edition, Chemical Rubber Publishing Company, Cleveland, Ohio (Library of Congress Catalog Card No. 13-11056), Reprinted February 1960.

[5] R. F. Barron, *Cryogenic Systems*, 2nd ed., Oxford University Press, New York, New York, 1985.

[6] W. H. Higa and E. Wiebe, *One Million Hours at 4.5 Kelvins*, National Bureau of Standards Technical Publication No. 508, pp. 99–107, April 1978.

[7] R. A. Frosch and F. E. McCrea, *Support Assembly for Cryogenically Coolable Low-Noise Choked Waveguide*, United States Patent 4,215,327, July 29, 1980.

[8] M. Britcliffe, "Two-Watt, 4-Kelvin Closed Cycle Refrigerator Performance," *The Telecommunications and Data Acquisition Progress Report 42-91*, July–September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 312–317, November 15, 1987. http://ipnpr.jpl.nasa.gov/progress_report

[9] T. R. Hanson, "Helium Compressors for Closed-Cycle, 4.5-Kelvin Refrigerators," *The Telecommunications and Data Acquisition Progress Report 42-111*, July–September 1992, Jet Propulsion Laboratory, Pasadena, California, pp. 246–253, November 15, 1992. http://ipnpr.jpl.nasa.gov/progress_report

[10] M. Britcliffe, "A Closed-Cycle Refrigerator for Cooling Maser Amplifiers Below 4 K elvin," *The Telecommunications and Data Acquisition Progress Report 42-98*, April–June 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 141–149, August 15, 1989. http://ipnpr.jpl.nasa.gov/progress_report

[11] M. Britcliffe, T. Hanson, and J. Fernandez, "An Improved X-Band Maser System for Deep Space Network Applications," *The Telecommunications and Mission Operations Progress Report 42-141*, January–March 2000, Jet Propulsion Laboratory, Pasadena, California, pp. 1–10, May 15, 2000. http://ipnpr.jpl.nasa.gov/progress_report

[12] "SRDK Series Closed Cycle Refrigerator," Sumitomo Heavy Industries, LTD., Tokyo, Japan. http//shicryogenics.com/products/4kcryocoolers.jssx

[13] E. M. Long, "A Dual-Channel 8- to 9-Gigahertz High-Electron Mobility Transistor (HEMT) Low-Noise Amplifier (LNA) Package for the Goldstone Solar System Radar," The *Interplanetary Network Progress Report*, vol. 42-165, Jet Propulsion Laboratory, Pasadena, California, pp. 1–8, May 15, 2006. http://ipnpr.jpl.nasa.gov/progress_report

Chapter 5
HEMT Low-Noise Amplifiers

J. Javier Bautista

5.1 Introduction—Semiconductor Conductivity

A large variety of electronic devices (such as diodes, transistors, mixers, attenuators, and photo-detectors, to name a few) are based on the unique electrical conducting properties of semiconductors. Semiconductors are solid-state conductors whose electrical conductivity ranges approximately between 10^{2} and 10^{-9} mho per centimeter (mho/cm) (or its inverse, resistivity between 10^{-2} and 10^{9} ohm-cm) at room temperature [1]. These conductivities are higher than an insulator's but lower than a metal's. Although most pure single-crystal semiconductors are insulators at absolute zero with conductivities less than 10^{-14} mho/cm, it is impurities and imperfections that allow manipulation of semiconductor's device properties. In fact, it is the ability to precisely engineer and finely tune a semiconductor's conductivity at the atomic layer level that is essential to the production of state-of-the-art high electron mobility transistor (HEMT) devices [2].

Important semiconductors like silicon (Si), germanium (Ge), gallium arsenide (GaAs), indium phosphide (InP) and their ternary derivatives, like aluminum gallium arsenide (AlGaAs), indium gallium arsenide (InGaAs), and indium aluminum arsenide (InAlAs), form a crystalline diamond (or zincblende) lattice structure bonded by covalent forces. Each atom has four nearest neighbors lying at the corners of a tetrahedron, and it shares an electron with each atom. The covalent attractive potential between atoms is created by the shared electron pairs of opposite spin. The three-dimensional models for the diamond and zincblende structures are shown in Fig. 5-1.

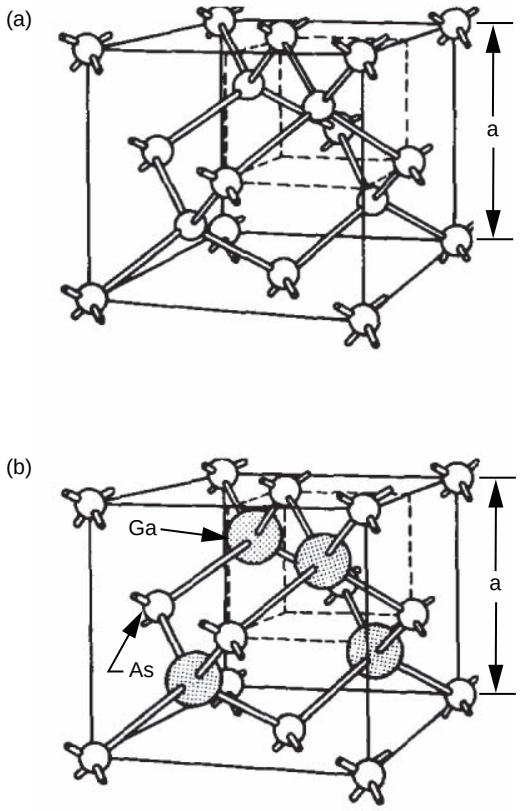


Fig. 5-1. Three-dimensional periodic distribution of atomic sites or direct lattice of important semiconductors and their representative elements or compounds for (a) diamond (carbon [C], germanium [Ge], silicon [Si], etc.) and for (b) zincblende (gallium arsenide [GaAs], gallium phosphide (GaP], etc.). Note: a is the lattice constant.

5.1.1 Charge Carrier and Energy Band Gap

Semiconductors can have two types of charge carriers that contribute to the conductivity, electrons and holes. In pure intrinsic semiconductors at room temperature, thermal energy can free an electron (negative charge) from its bonding or lattice site, leaving behind a vacant positively ionized atom. The vacancy is called a hole, and is equivalent to a charge carrier of positive charge moving in a direction opposite to the electron's motion.

Unlike electrons in free space, which have a continuous set of energy states, electrons in a solid have forbidden bands of energy. At low energies electrons

in a solid behave much like free, classical electrons. However, as the electron energy increases, they interact with and are scattered by the lattice, exhibiting their wave nature. This occurs at energies where the electron's de Broglie wavelength, $\lambda = h / p$ (Planck's constant/electron momentum), approaches the inter-atomic spacing, a, producing a band of forbidden energy states. A graphical representation for the electron energy as a function of wave vector ($k = 2\pi / \lambda$) for an ideal one-dimensional solid is shown in Fig. 5-2.

5.1.2 Charge Carrier Transport Properties

In general, the band structure or energy-momentum (E-k) relationship for a semiconductor is obtained by solving the Schrödinger equation of an approximate one-electron problem using a variety of numerical methods (e.g., the orthogonalized plane wave method, the pseudopotential method, and the **k•p** method) and verified by experimental measurements (such as optical absorption and cyclotron resonance techniques). The intrinsic transport properties that determine microwave- and millimeter-wave performance, such as electron mobility (μ_e), peak velocity (v_p) and effective mass (m^*), can thus be calculated and verified by measurements at room and cryogenic temperatures.

In general, near the bandgap edge the electron effective mass is inversely proportional to the second derivative of the energy with respect to wave-vector,

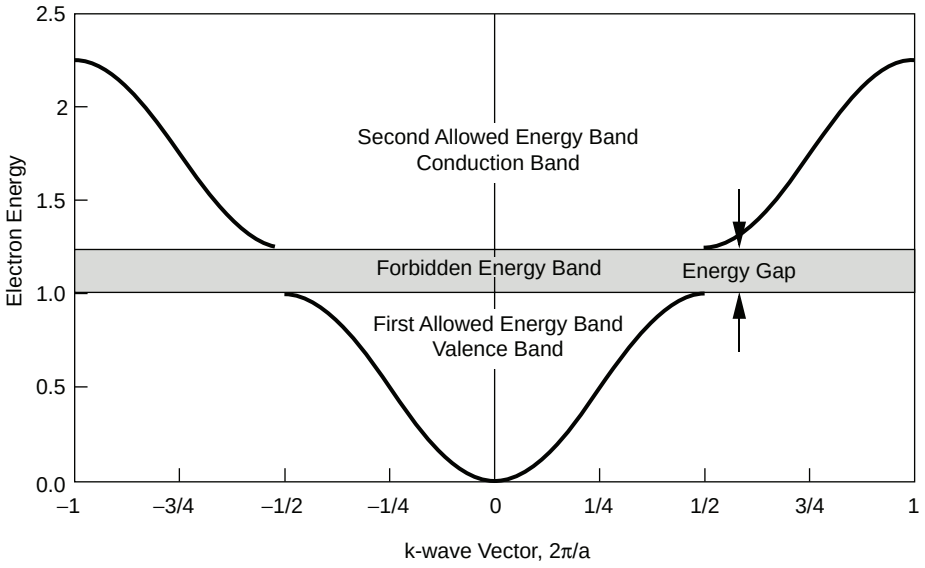


Fig. 5-2. Electron energy plotted as a function of wave-vector k (2π/a) for an electron in a one-dimentional solid of lattice constant a showing the energy gap.

$$m^* = h^2 \left(\frac{\partial^2 E}{\partial k^2} \right)^{-1} \tag{5.1-1}$$

Thus for large values of $\partial^2 E / \partial k^2$, m^* can be much less than the free electron mass. In addition, below room temperature the electron mobility is approximately inversely proportional to the product of the effective mass and physical temperature, that is $\mu_e \propto (m^* T)^{-1}$. While the electron velocity is in turn proportional to the product of the electron mobility and applied electric field,

$$\mathbf{v} = \mu_e \boldsymbol{\varepsilon} \tag{5.1-2}$$

Although silicon is the dominant material in use today for the manufacture of transistors and is the most mature technology, the III–V semiconductor compounds have far superior transport properties, especially at high frequencies and low temperatures. In fact, the m^* of GaAs and InP is less than an order of magnitude smaller than Si's, and their μ_e's are more than two orders of magnitude greater than Si's at liquid nitrogen temperatures (77 K). In fact, the maximum intrinsic frequency of operation for a HEMT device is dir ectly proportional to the electron velocity. Listed in Table 5-1 for comparison is the band gap energy (E_g), electron mobility (μ_e), peak velocity (v_p), and lattice constant (a) for a number of important semiconductors.

Table 5-1. Electronic properties of important semiconductors at room temperature [3].

Semiconductor	E_g (eV)	μ_e (cm^2/Vs)	v_p (10^7cm/s)	a (Å)	a (nm)
Ge	0.66	3900	0.6	5.65	0.565
Si	1.12	1500	1.0	5.43	0.543
GaAs	1.42	8500	1.8	5.65	0.565
InP	1.35	4600	2.4	5.87	0.587
$Ga_{0.15}$ $In_{0.85}As$	1.20	9500	2.9	5.85	0.585
$Ga_{0.47}$ $In_{0.53}As$	0.75	15,000	3.4	5.85	0.585
InAs	0.36	33,000	4.4	6.06	0.606
InSb	0.17	80,000	5.0	6.48	0.648

5.1.3 Donor and Acceptor Impurities

Electronic properties of semiconductors are engineered and optimized during material growth through the introduction of impurities, a process known as "doping." In Si and Ge phosphorous (P) or boron (B) are used as dopant impurities, while Si is used in GaAs and InP. Dopants with a valence of five (like P) replace the host atom Si or Ge adding an extra valence electron (donor state), while those with a valence of three (like B) reduce the number of valence electrons by one (acceptor state). However, in GaAs and InP the dopant Si can either introduce a donor state by replacing Ga (or In) or an acceptor state by replacing As (or P). The donor state is more energetically favorable, and therefore dominant.

Dopant impurities introduce discrete energy levels, donor and acceptor levels, within the forbidden energy band of the intrinsic semiconductor. These impurities states are similar to hydrogen atomic states with the exception that the orbital electrons or holes move through a dielectric medium. Recalling that the ionization energy for the hydrogen atom is

$$E_H = \frac{m_e e^4}{8\varepsilon_o^2 h^2} = 13.6 \text{ eV} \tag{5.1-3}$$

The ionization energy for the donor state E_d can be calculated by replacing the electron mass, m_e, with its effective mass, m^*, and the free space permitivity, ε_0, with the permitivity of the semiconductor, ε_s, yielding

$$E_d = \left(\frac{m^*}{m_e}\right)\left(\frac{\varepsilon_0}{\varepsilon_s}\right)^2 E_H \tag{5.1-4}$$

The ionization energy for donors calculated with this expression are 0.006, 0.025, and 0.007 eV for Ge, Si, and GaAs, respectively. The ionization energies can be calculated in a similar manner for acceptor states with comparable values (as measured from the valence-band edge): 0.015, 0.05, and 0.05 eV for Ge, Si, and GaAs, respectively. The donor energy states lie close to the conduction band edge, while acceptor states lie close to the valence band edge. In fact, the energy required to ionize these hydrogen-like atomic states can be estimated using this simple model. Since both are loosely bound states, both can be easily ionized by thermal energy resulting in excess electrons in the conduction band and excess holes in the valence band.

5.1.4 Heterojunction—HEMT versus MESFET

Until the invention of the HEMT, the most widely used III–V transistor for both microwave and high-speed digital applications was the GaAs metal semiconductor field effect transistor (MESFET). However, since electrons must transit through the doped channel in a MESFET, it does not take full advantage of the high mobilities in GaAs. The result is more than a 50-percent reduction in electron mobility, since ionized dopants scatter electrons. Hence, separation of the dopant channel from the electron transit channel is key to the superior noise, gain and frequency performance of the HEMT. For comparison, the cross sections of a GaAs HEMT and a MESFET are shown in Fig. 5-3 and a comparison of their material properties are shown in Table 5-2.

The model HEMT structure can be formed of two distinct semiconductor layers [4]. The bandgap difference results in the formation of conduction and valence band discontinuities at the layer interface or heterojunction creating a quantum well in the conduction band. The wider band gap semiconductor is

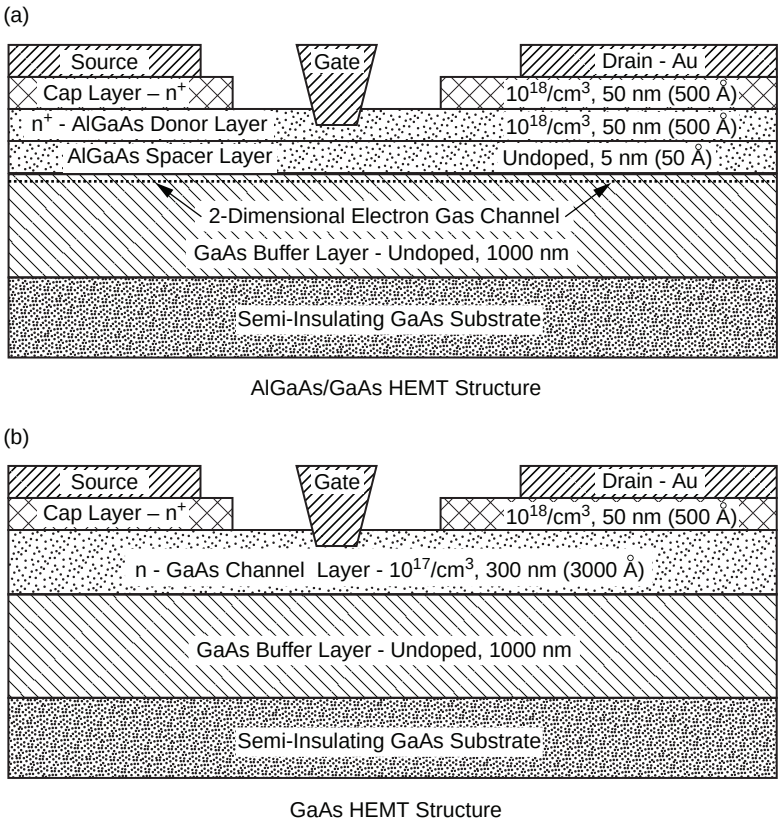


Fig. 5-3. Cross sectional diagrams comparing structures of an (a) AlGaAs/GaAs HEMT and a (b) GaAs MESFET [3].

Table 5-2. Material properties of conventional HEMT and GaAs MESFET structures.

Material Properties	HEMT	MESFET
Sheet charge density ($1/cm^3$)	10^{18}	10^{17}
300-K electron mobility ($cm^2/V\text{-}s$)	8,500	4,000
77-K electron mobility ($cm^2/V\text{-}s$)	80,000	6,000

* Assume a spacer thickness of 5 nm (50 Å)

doped with donors while the smaller band gap material is left undoped. The conduction band electrons move from the donor layer to the undoped layer forming a two-dimensional electron gas (2-DEG) along the heterojunction. The band gap discontinuities are energy barriers spatially confining the electrons.

5.2 The Many Acronym-ed Device (MAD)—A Brief HEMT History

The commercial HEMT evolved from the GaAs and AlGaAs superlattice (multiple heterorstructures) research conducted in the late 1960s by Leo Esaki and Ray Tsu at the IBM Thomas J. Watson Research Center in Yorktown Heights, New York [5]. It was not until 1978 when Raymond Dingle, Horst Stormer, and Arthur Gossard at Bell Laboratories, Murray Hill, New Jersey, first demonstrated the high mobilities possible in these superlattices that world-wide interest developed at a number of university, industrial, and government laboratories [6,7]. Among them were Cornell University, University of Illinois, University of Michigan, University of Tokyo, University of Duisberg, Germany, Rockwell International, General Electric Company, TRW Inc., Fujitsu Ltd., Japan, Thomson CSF, France, U.S. Naval Laboratory, and the U.S. Air Force Avionics Laboratory, to name but a few. The first cryogenic, microwave HEMTs were reported in 1983 by Thompson CSF, France [8], and Fujitsu, Ltd., Japan [9]. The HEMT is known by other names, such as modulation-doped field effect transistor (MODFET), two-dimensional electron gas field effect transistor (TEGFET), selectively-doped heterostructure transistor (SDHT), and heterojunction field effect transistor (HFET) reflects the number of laboratories involved world-wide in its development and the device property they chose to emphasize.

5.2.1 HEMTs in the Deep Space Network and Radio Astronomy—Voyager at Neptune

One of the great technical challenges faced by the Jet Propulsion Laboratory's (JPL's) Deep Space Network (DSN) is to receive signals from spacecraft that are up to billions of kilometers from Earth. Since the transmitted power is limited and fixed, the communication burden is placed on the ground-based antenna receive systems, which must detect an extremely weak signal in

the presence of a nearly overwhelming amount of noise. The Deep Space Network (DSN) approach is to simultaneously maximize the received signal collection area and receiver sensitivity. This key figure of merit for a telecommunications link is the ratio of antenna gain to operational noise temperature (G/T_{op}) of the system.

For the Voyager encounter with Neptune, the 27 antennas of the National Radio Astronomy Observatory's (NRAO) Very Large Array (VLA) in Socorro, New Mexico, were successfully arrayed with the antennas of the DSN's Goldstone complex in California at 8.4 GHz. The VLA was equipped with cryogenic, HEMT low-noise amplifiers (LNAs) while the DSN was operated with its extraordinarily sensitive cryogenic, ruby maser LNAs.

Since masers were too expensive, the VLA was to be originally equipped with cryogenic, GaAs field effect transistor (FET) LNAs. However, in 1983 during the definition phase of the VLA/Goldstone Telemetry Array project, the first cryogenic, HEMT was announced. The results, reported by Thompson CSF [8] and Fujitsu [9], suggested that the cryogenic HEMT was capable of significantly lower noise than FETs at cryogenic temperatures. Their results further suggested that much of the system noise performance lost with the use of FETs instead of masers could be regained if the new HEMT device could be developed in time for VLA implementation. In 1984, a cooperative program was initiated among NRAO, JPL, General Electric Co. (GE), and Cornell University to develop a HEMT device and HEMT amplifiers optimized for cryogenic use in the 1- to 10-GHz range. GE was contracted by JPL to design and fabricate the devices, while NRAO assumed responsibility for device evaluation and amplifier development at 8.4 GHz for the VLA. JPL was similarly responsible for device evaluation and amplifier development for DSN applications at 2.3 GHz. During the program, GE HEMT noise temperatures and device yields steadily improved. At the start of the program in 1985, the best device cryogenic noise temperature demonstrated at 8.4 GHz was 8.5 K with an associated gain of 12 dB. Near the end of the program in 1986, the device noise temperature had dropped to 5.5 K with more than 14-dB associated gain [10].

5.2.2 InP HEMT LNAs in the Deep Space Network

To date, the noise, gain, and maximum frequency of InP HEMTs at room temperature is steadily improving as the technology is being internationally developed and commercialized. Although device (commercial and research) noise temperatures continue to fall at ambient, there is no guarantee that an attendant improvement at cryogenic temperatures will be realized. To develop ultra-low-noise microwave amplifiers for cryogenic applications, one must have a reliable source of state-of-the-art cryogenic devices and the capacity to accurately characterize them at the device or wafer level at cryogenic

temperatures. The work on cryogenic, InP HEMTs was based on another partnership among TRW, Inc., the Georgia Institute of Technology (GIT), and the Jet Propulsion Laboratory (JPL). TRW was responsible for device fabrication and optimization, GIT cryogenic device noise parameter characterization, JPL cryogenic device scattering parameter characterization, LNA module development, and receiver implementation.

The development of cryogenic, InP HEMTs has enabled the demonstration of state-of-the-art LNA modules that are yielding noise temperatures less than ten times the quantum noise limit from 1 to 100 GHz (0.5 to 50 K). The state-of-the-art noise temperature of cryogenic, HEMT-based amplifiers has steadily improved since the invention of the HEMT. Notable examples at physical temperatures near 20 K are 5.5 K at 8.4 GHz in 1986 [10], 15 K at 43 GHz in 1993 [11], and 30 K at 102 GHz in 1999 [12].

The DSN is in the process of implementing this technology to meet its current and future (2010) navigation, telemetry, radar, and radio science needs at 8.4 and 32 GHz. To date, typical InP HEMT LNA modules developed for the DSN have demonstrated noise temperatures of 3.5 K at 8.4 GHz and 8.5 K at 32 GHz. Front-end receiver packages employing these modules have demonstrated operating system noise temperatures of 17 K at 8.4 GHz on a 70-m Cassegrain antenna and 39.4 K at 32 GHz on a 34-m beam wave-guide (BWG) antenna, both at zenith [13].

5.3 HEMT Growth Technology

The material optimization of HEMT structures, sometimes referred to as bandgap engineering, is complex and typically focused on increasing the electron's mobility, velocity, and density and on improving its confinement to the channel layer. The material quality is key to achieving optimal device performance and can be significantly degraded by impurities, defects, heterojunction roughness, and lattice mismatch strain. These criteria dictate relatively slow growth rates of 1 to 2 atomic layers per second (epitaxial growth) of lattice matched materials on heated semi-insulating GaAs or InP wafers. Molecular beam epitaxy (MBE) [14–16] and metal-organic chemical vapor deposition (MOCVD), sometimes referred to as metal-organic vapor phase epitaxy (MOVPE) [17,18], are the two primary techniques used for HEMT epitaxial growth. (For digital and power applications a high transconductance over a large gate bias range is desirable, while for microwave devices the peak transconductance is more important.)

5.3.1 Molecular Beam Epitaxy

MBE growth of HEMT structures such as AlGaAs/GaAs, AlGaAs/InGaAs/GaAs, and InAlAs/InGaAs/InP HEMTs is performed in a vacuum chamber held at ultra-low pressures (less than 10^{-8} torr

(1.33×10^{-6} Pa)). High purity elemental sources (e.g., Al, Ga, and In) are evaporated while As and antimony (Sb) are sublimated when heated to high temperatures in furnaces known as effusion or Knudsen cells by atomic and molecular beams, respectively. For example, the epitaxial growth of GaAs is performed by heating and evaporating atomic Ga from the melt, while solid arsenic is heated to sublimate As_4 which is "cracked" by further heating to produce molecular, As_2. The furnace temperatures are adjusted so that the flux ratio of gallium to arsenic is 2 to 1 . The atomic (Ga), and molecular (As_2), beams are aimed at a heated, substrate (GaAs or InP) and adsorbed on the heated substrate. They react to form GaAs; facilitating the growth of GaAs one atomic layer at a time. Typical growth rates range from 0.1 to 5 μm per hour. During growth n-type (Si) and p-type (beryllium, Be) dopants (heated in effusion cells) can be incorporated into the layers. The epilayer composition is controlled with shutters placed in front of the sources and by variation of the substrate temperature. In addition, epilayers must be closely lattice matched to avoid crystalline defects, such as dislocations. The lattice constant and the energy gap of some III–V semiconductor compounds are plotted in Fig. 5-4, and those compounds lattice matched to GaAs and InP are noted.

5.3.2 Metal-Organic Chemical Vapor Deposition (MOCVD)

Gaseous sources, such as trimethyl or triethyl organometallics, of the elemental compounds are used for the epitaxial layer growth of HEMT

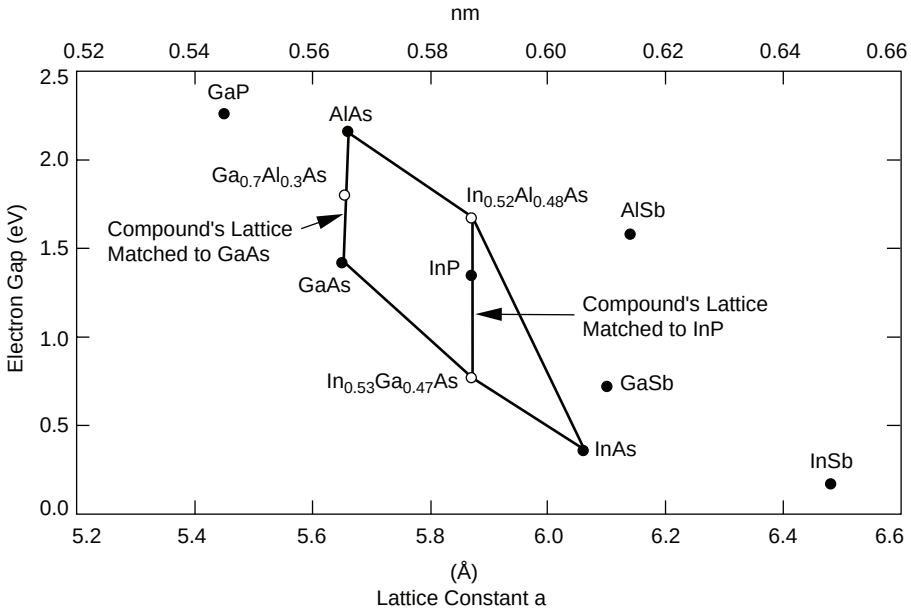


Fig. 5-4. The energy gap of III–V compounds and their ternary derivatives as a function of lattice constant.

structures in MOCVD. The gaseous compounds are "pyrolyzed" at the heated surface of a substrate like GaAs. For example, to grow GaAs trimethylgallium, $(CH_3)_3Ga$, (abbreviated TMGa) carried by hydrogen gas and gaseous arsine, AsH_3, are passed over the substrate surface. The adsorbed radicals react to produce GaAs and methane gas. Silane (SiH_4), diluted with hydrogen can be added to the mix to incorporate donor impurities. MOCVD growth rates can be as low as 1 μm/h to as high as 30 μm/h. The main advantage of MOCVD is high throughput and multi-wafer growth capability. Another important advantage is the ability to grow various indium- and phosphorus-containing compounds using phosphine (PH_3). However, in comparison to MBE, it is a greater challenge to achieve the same level of thickness and doping uniformity using MOCVD because of vapor-phase effects, such as source gas-phase depletion, turbulence, and convection.

Another advantage of MBE over MOCVD is the ability to use powerful in situ characterization techniques, such reflection high-energy electron diffraction (RHEED) for monolayer counting and compositional measurement of layers as well as Auger electron spectroscopy (AES), quadrupole mass spectrometry, and automatized ellipisometry [16]. In addition, MBE achieves higher compositional resolution, accuracy, and uniformity across the wafer compared to MOCVD. However, MBE is more complex and expensive, and has lower throughput, while MOCVD is relatively simple and less expensive, and is capable of a higher throughput capacity. Through improvements in the design of MBE systems, significant progress has been made in improving MBE surface morphology and throughput. Production MBE machines have been commercially available for more than a decade, currently costing on the order of one million dollars with a comparable amount required for operation and maintenance per year.

To summarize, MBE provides the ultimate control for growing heterostructures with precise layer composition, doping profile, and sharp layer transition. To date, the best device and circuit results have been demonstrated with MBE grown HEMTs.

5.4 HEMT Materials Evolution—From GaAs to InAs

As previously mentioned, the first successful HEMTs were based on the lattice-matched heterostructure AlGaAs/GaAs. A few years later, In was added to the carrier channel to improve device performance, that is, to increase electron mobility (μ_e), increase the frequency of operation and lower the noise. The InGaAs channel layer inserted between the AlGaAs and GaAs is not lattice matched to either compound, GaAs or AlGaAs, but compressed to match them at their interface; accounting for the device name, pseudomorphic HEMT (PHEMT). The thickness of the InGaAs layer (between 50 to 200 Å [5–20 nm], depending on In concentration) is chosen so that most of the compressional

strain is taken up by the InGaAs layer [19]. To further push device performance, the In concentration was increased from the range of 15–20 percent to 65 percent, the spacer material changed to InAlAs, and InP substrates used to accommodate the larger lattice mismatch [20]. Pure InAs as the carrier channel is currently being investigated as the next logical step to produce the ultimate in HEMT device performance. The advantage of this material is the high electron mobility (30,000 cm^2/Vs at 300 K) and velocity (4×10^7 cm/s) and a large conduction band offset between InAs and AlSb (1.35 eV) [21,22].

Although the layer number, composition, and thickness vary depending on the desired properties, all HEMT layer structures have the same essential feature, a vertical heterojunction. The heterojunction spatially separates charge carriers from donors and confines them to the channel layer where the electron momentum is quantized in the "vertical" direction but is continuous in the horizontal direction. In the following discussion, the design guidelines for an optimal low noise AlGaAs/GaAs HEMT device are given, then a summary and comparison is presented of some of the essential enhancing features of PHEMTs and lattice-matched InP HEMTs.

5.4.1 Optimized Low-Noise AlGaAs/GaAs HEMT Structure

As shown in Fig. 5-3a, the conventional AlGaAs/GaAs HEMT structure is grown on a GaAs semi-insulating substrate with the following epitaxial layers: an undoped buffer and GaAs channel layer, an undoped AlGaAs spacer layer, a heavily doped (n+) AlGaAs donor or gate-barrier layer, and an n+ GaAs capping or ohmic contact layer. These layers are essential for fabricating and understanding the operation of a HEMT device. Depending on the application (for example, low noise, power, or digital), modifications and refinements to the basic structure are necessary to obtain optimum device performance. Some of the modifications and refinements to fabricate low-noise, high-gain devices are presented in the following paragraphs.

Under normal bias conditions the drain-to-source electric field can inject electrons beyond the 2-DEG channel into the GaAs buffer layer, contributing excess drain current, resulting in gain reduction and degradation of the device noise performance. Introduction of a high band-gap AlGaAs buffer layer before the GaAs buffer suppresses the buffer layer drain-to-source leakage current by creating an energy barrier in the conduction band to reduce electron injection into the buffer, while reducing the velocity of injected electrons [23]. The use of an AlGaAs buffer, however, results in buffer-channel interface roughness that reduces the mobility in the device channel [24]. The interface roughness can be improved by incorporating a thin GaAs smoothing or AlGaAs/GaAs superlattice buffer [25] layer between the buffer and the channel. A superlattice buffer, thin alternating layers of differing materials sharing the same crystalline

lattice, is very effective at confining carriers to the 2-DEG channel without sacrificing the material quality.

The thin spacer layer separating electrons from their donors is to reduce the scattering of electrons by the positively charged donors. This is done by placing a thin spacer layer of undoped AlGaAs with a thickness ranging from 20 to 50 Å (2–5 nm) between the AlGaAs donor and the GaAs channel layer to separate the negatively charged 2-DEG from the ionized dopant atoms. At room temperature, a thin spacer layer of approximately 20 Å (2 nm) is preferred for low-noise and power devices due to the reduced parasitic source resistance and the increased transconductance and current density. A thicker spacer, conversely, provides a higher electron mobility with a smaller charge density in the channel. At cryogenic temperatures the noise performance of a HEMT is strongly dependent on the spacer thickness, and a thickness of 40 Å (4 nm) has been determined to be optimum due to the large increase in electron mobility and velocity [26].

In order to eliminate parallel conduction in the AlGaAs donor layer, this layer must be completely depleted by both the AlGaAs/GaAs heterojunction and the Schottky gate. The donor layer is typically uniformly doped with Si at a doping level of approximately 10^{18} atoms/cm^3. The high doping level makes possible the small spacing between the gate and the carrier channel. A higher doping level results in a higher sheet charge density in the channel, increasing transconductance (g_m), unity current gain frequency (f_T), and current density, at the expense of a lower breakdown voltage. Fortunately, high sheet charge density and breakdown voltage can be achieved with planar-doping, sometimes also referred to as δ-doping or pulse-doping [27,28]. The planar-doping layer is a monolayer of Si approximately 5 Å (0.5 nm) thick with a doping level of approximately 5×10^{12}/cm^2 located just above the spacer. The use of planar-doping also allows a lower doping level in the AlGaAs layer for the gate barrier, increasing the breakdown voltage without sacrificing the channel sheet charge density.

The AlAs mole fraction, x, in the $Al_xGa_{1-x}As$ donor layer is another important parameter controlling low-noise performance. At room temperature the conduction band discontinuity, ΔE_c, at the $Al_xGa_{1-x}As/GaAs$ heterojunction interface is linearly dependent on the AlAs mole fraction given by [29]

$$\Delta E_c = 0.8806x, \text{ for } x < 0.47 \tag{5.4-1}$$

This expression indicates that an increase in the AlAs mole fraction will result in an increase in the total carrier sheet charge density. In addition, a higher AlAs mole fraction reduces scattering of carriers from the carrier channel to the AlGaAs layer and minimizes the excess modulation of the

AlGaAs layer [30]. However, the high AlAs mole fraction reduces the doping efficiency in the channel and for $x > 0.25$ leads to the production of the DX centers [31]. DX centers are deep donor states related to the band structure of AlGaAs and are responsible for the persistent photoconductivity (PPC), threshold voltage shift, and drain I–V collapse at low temperatures [32,33]. Although, DX centers do not pose a problem at room temperature, they severely limit the performance of cryogenic low-noise microwave HEMTs.

To provide a good ohmic contact to the charge carrier channel, the GaAs capping layer is approximately 500 Å (50 nm) thick and typically heavily doped with Si at approximately 10^{18} atoms/cm^3. This reduces the device source resistance and protects the AlGaAs donor layer from surface oxidation and depletion. Higher doping levels and a thicker capping layer would simultaneously reduce the device source resistance and the effectively shorten the source-to-drain spacing, resulting in very high electron velocity, transconductance (g_m), and unity current gain frequency (f_T) in the device. However, this also significantly reduces the device breakdown voltage and also increases the device drain output conductance and drain-to-gate feedback capacitance. It is also more difficult to obtain uniform gate recess in a HEMT with a thicker capping layer. The capping layer has to be completely recessed under the gate to eliminate parallel conduction in the GaAs layer.

5.4.2 The GaAs Pseudomorphic HEMT—A1GaAs/InGaAs/GaAs PHEMT

In 1986 the GaAs pseudomorphic HEMT was introduced as a high performance alternative to the AlGaAs/GaAs HEMT. Substitution of InGaAs for GaAs as the two-dimensional electron gas channel improves transport properties due to the higher mobility of InGaAs and stronger electron confinement associated with the quantum well at the heterojunction. Thus, injection of electrons back into the AlGaAs from the InGaAs is significantly reduced; thereby improving the transport properties. The larger conduction-band discontinuity at the AlGaAs/InGaAs heterojunction allows a higher sheet charge density and hence a higher current density and transconductance. Additionally, the electron mobility and peak velocity can be further improved by increasing the indium concentration (Fig. 5-5). The room temperature mobility of this type of PHEMT structure is generally 5000 to 7000 cm^2/Vs with a 2-DEG concentration of $1–3 \times 10^{12}$ cm^{-2}.

Although InGaAs is not lattice matched to either the AlGaAs donor or the GaAs buffer layers (Fig. 5-4), the strain associated with the lattice mismatch can be elastically accommodated within the InGaAs layer. For example, for a PHEMT structure like $\mathrm{AlGaAs/In_xGa_{1-x}As/GaAs}$ where x is in the range of 0.15 to 0.20, the InGaAs layer must be smaller than the critical thickness ~150 Å (15 nm) . Above the critical thickness, lattice dislocations form, while

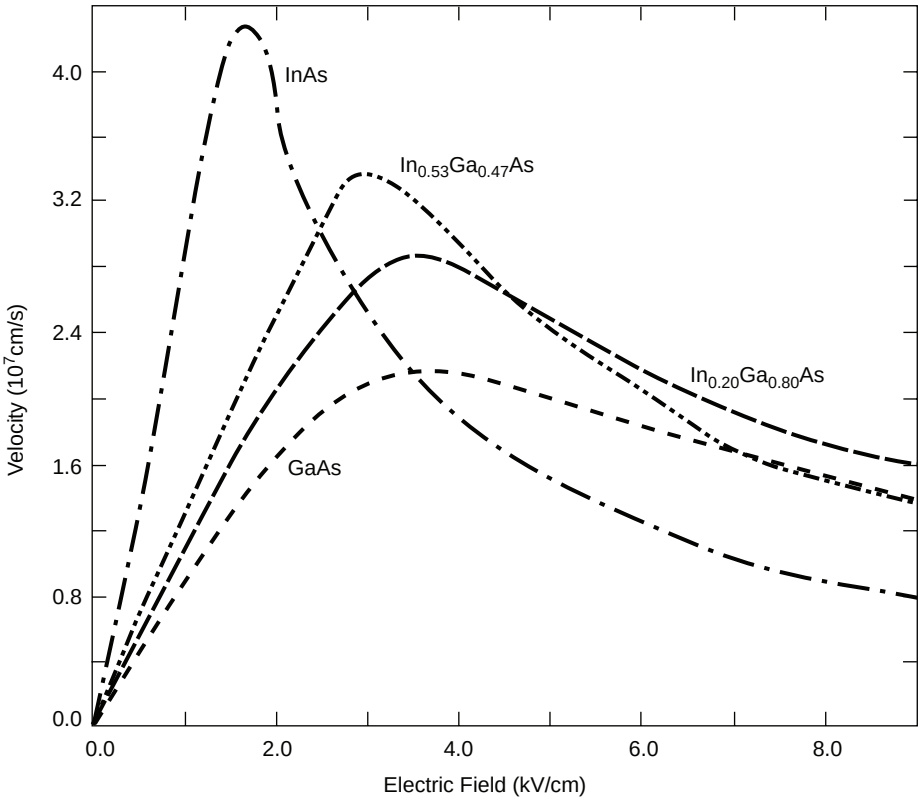


Fig. 5-5. Electron velocity as a function of electric field for variety of In concentrations of InGaAs.

for a thickness less than ~ 50 Å (5 nm), quantum size effects substantially reduce electron confinement and increase electron scattering [17,34].

In 1990 the state-of-the-art performance for a 0.1-μm gate length device PHEMT was an $f_{\max}$ of 290 GHz [35], an f_T of 130 GHz, and a minimum noise figure of 2.1 dB with an associated gain of 6.3 dB at 94 GHz. Pseudomorphic technology is quite mature, and microwave monolithic integrated circuits (MMICs) based on this type of material are common and routinely exceed this performance [36].

5.4.3 InAlAs/InGaAs on an InP HEMT

A quick check of Fig. 5-4 (higher conduction band discontinuity) and Fig. 5-5 (higher electron velocity) suggest that increasing the In concentration in the carrier channel of the PHEMT will result in further improvements in electron carrier confinement and transport properties. Unfortunately, increasing the indium concentration in $In_xGa_{1-x}As$ also increases the lattice constant.

Increasing the In concentration to the highest possible value is desirable; however, the higher lattice mismatch strain between $In_xGa_{1-x}As$ and GaAs cannot be elastically accommodated. Thus, GaAs is not suitable as substrate material for such high indium concentration HEMT structures.

Fortunately, the ternary compounds, $In_{0.52}Al_{0.48}As$ and $In_{0.53}Ga_{0.47}As$ are lattice matched to InP, and they can thus be used as high bandgap and low bandgap material, respectively, in the HEMT structure. Low breakdown voltages due to impact ionization in the InGaAs channel, limit this material to low-noise applications. HEMT devices using this material structure have very good high-gain and low-noise performance at high frequencies. Additionally, this structure also eliminates deep-level trapping sites (DX centers), providing a solution to the persistent photoconductivity effect at cryogenic temperatures. That is, unlike AlGaAs, InAlAs does not suffer from DX center effects and can be doped with Si up to 1×10^{19} cm^{-2} [37].

In fact, the mobility in lattice matched InP HEMT structures varies from 8,000 to 12,000 cm^2/Vs with a 2-DEG concentration of 2 to 4 $\times 10^{12}$ cm^{-2}. In 1990 state-of-the-art performance for InP-based HEMTs displayed a maximum operating frequency (f_{max}) of 455 GHz and a maximum available gain of 13.6 dB at 95 GHz for a 0.15-μm gate length HEMT [38]. Another InP-based HEMT had a minimum noise figure of 1.7 dB with an associated gain of 7.7 dB at 93 GHz [39].

Materials with still higher indium concentrations in the channel have superior electron transport characteristics. There are several approaches for the incorporation of more indium into the carrier channel: pseudomorphic InGaAs channel with an In concentration above 0.53 [40], insertion of a thin $In_xGa_{1-x}As$ $(x > 0.53)$ layer in the channel region [41–43], change of material system to InAs/AlGaSb [44], and a graded channel [45,46]. In most of these approaches the thickness of the high-In concentration InGaAs layer must be kept thin to prohibit formation of dislocations. An InP-based pseudomorphic HEMT had a transconductance of 1700 millisiemens/millimeter (mS/mm) with an f_T of 305 GHz for a 65-nm gate length device [47].

An approach that does not limit the channel thickness to the critical layer thickness is to lattice mismatch the substrate to the channel with a buffer that transforms the lattice constant of the substrate to that of a high indium concentration channel. Various methods to realize this buffer have been demonstrated such as graded InGaAlAs [48,49], graded $In_xGa_{1-x}As$ [50], and step-graded $In_yGa_{1-y}As$ [51] buffer layers. In [50], a 0.4-μm gate $In_{0.29}Al_{0.71}As/In_{0.3}Ga_{0.7}As$ HEMT on GaAs with an f_{max} and an f_T of 115 and 45 GHz, respectively, was presented. The transconductance and the current density were 700 mS/mm and 230 mA/mm. In [51], a 0.16-μm

$In_{0.5}Al_{0.5}As/In_{0.5}Ga_{0.5}As$ HEMT on GaAs substrate had an f_{max} and an f_T of 147 and 67 GHz, respectively. The transconductance and current density were 1060 mS/mm and 550 mA/mm.

To date, InP-based HEMTs have been shown to be the best performing three-terminal devices [52], with excellent performance in the microwave and millimeter-wave range. The combination of high gain and low noise has been demonstrated by many devices and circuits having operating frequencies as great as 100 GHz and higher [53–56].

5.4.4 InAlAs/InGaAs on GaAs HEMT—Metamorphic HEMT or MHEMT?

A drawback of the InP substrate is that it is a fairly young technology compared to GaAs. InP substrates are more expensive, more fragile, and more difficult to etch. Currently, only 2-inch (in.) and 3-in. (5- and 7.6-cm) high quality InP substrates are available at relatively high cost, whereas 6-in. (15.2-cm) GaAs substrates are readily available. Additionally, wafer thinning and backside processing technologies are more mature for GaAs. To combine the advantages of the GaAs substrate with the advantages of InP based HEMTs, metamorphic InGaAs/InAlAs quantum-well structures or metamorphic high electron mobility transistors (MHEMTs) were developed.

In MHEMTs the lattice constant of GaAs is transformed into the InP lattice constant with an appropriate buffer (typically 1 to 2 μm thick (1000 to 2000 nm)), on which the lattice-matched InP HEMT layer is grown. MHEMTs using quaternary buffers (such as AlGaAsSb) and a ternary buffer (such as InAlAs on GaAs) have show perfomance comparable to latticed InP HEMTs at room temperature. The quality of the final heterostructure, and thus the device performance, depends fully on the buffer type and quality [57]. Table 5-3 is a summary of the evolution of the first HEMT structure to current HEMT structure technology.

5.5 Device Fabrication

This section outlines the major processing steps in the fabrication of HEMT devices and microwave monolithic integrated circuits (MMICs). These processing steps are mostly based on, and hence quite similar, to GaAs MESFET processing. The major steps following growth of the heterostructure material are surface preparation and cleaning, front-side processing, and backside processing. Front-side processing steps include mesa or device isolation, ohmic contact formation, gate formation, metallization, and device passivation, while back-side processing includes substrate thinning via-hole formation and dicing. The processes are all performed in a clean room environment, varying in grade from Class 10,000 for the least critical steps of

Table 5-3. Summary of HEMT structures.

Layer	HEMT (x<0.25)	GaAs PHEMT	InP HEMT	InAs HEMT	MHEMT
Cap	Heavily doped GaAs	Heavily doped GaAs	Heavily doped InGaAs	InAs $In_{0.4}Al_{0.6}As$ AlSb	Heavily doped InGaAs
Donor Schottky	Heavily doped $Al_xGa_{1-x}As$ Si doping	Heavily doped $Al_xGa_{1-x}As$ Si doping	$In_yAl_{1-y}As$ Si δ doping	AlSb Te δ doping	$In_yAl_{1-y}As$ Si δ doping
Spacer	$Al_xGa_{1-x}As$	$Al_xGa_{1-x}As$	$In_yAl_{1-y}As$	AlSb	$In_yAl_{1-y}As$
Channel	GaAs	$In_xGa_{1-x}As$	$In_xGa_{1-x}As$	InAs	$In_xGa_{1-x}As$
Buffer-1		GaAs	InP	AlSb	$Al_{0.5}Ga_{0.5}As_{1-z}Sb_z$ or InAlAs
Buffer-2			InAlAs	$Al_{0.7}Ga_{0.3}As$	
Buffer-3				AlSb	
Buffer-4				GaAs	
Substrate	GaAs	GaAs	InP	GaAs	GaAs

wafer preparation to Class 100 for the most critical steps, such as ohmic contact and gate formation.

5.5.1 Wafer Preparation and Cleaning

To maintain optimum fabrication conditions and insure high device yields, cleaning operations are performed before all major steps during device processing. Relatively benign solvents, acids, bases, and rinses are used to remove contaminants such as organic materials, metals, and oxides.

For example, organic solvents effectively remove oil, grease, wax, photo resist, and electron-beam resist without affecting the HEMT device and circuit materials. The most common cleaning method is to im merse and agitate the wafer in the heated solvent. Solvents are then removed with alcohol that is, in turn, rinsed off with filtered, de-ionized water.

Acids are used for the wet etch removal of III–V semiconductor material as well as the removal of metal and oxide contaminants. The presence of thin interfacial dielectric layers, such as oxides, cause poor ohmic contact and Schottky barrier formation. These oxides can be dissolved with wet bases or removed with plasma etching techniques. Following cleaning and rinsing, the wafer is carefully dried to avoid leaving solvents or water stains.

5.5.2 "Hybrid" Lithography

To achieve high speed, low noise temperature, and high power-added efficiency at high frequencies, HEMTs require very short gate lengths. The intrinsic maximum frequency of operation, $f_{\max}$, is the figure of merit used to evaluate HEMT performance which is given by

$$f_{\max} = \frac{v_e}{2\pi L_g} \tag{5.5-1}$$

where v_e is the electron velocity and L_g is gate length. Fabricating very short gates is a challenge that requires well-developed lithography and pattern-transfer techniques.

Masks for the lithography of HEMTs are fabricated using a combination of electron beam lithography (EBL) [58] and photolithography. Although direct-write EBL is a low throughput exposure process, it facilitates accurate definition and alignment of sub-micron geometries while providing flexibility and fast turnaround for design iterations. EBL is widely and routinely used to produce gate dimensions of less than 0.25 μm [59–61]. To improve the wafer exposure throughput, optical lithography is used for the coarse features (>1 μm or 1000 nm), and the direct EBL is used only for the very short gates. This hybrid lithography has the advantage of both the high throughput of optical lithography and the high resolution and accuracy of EBL. The hybrid lithography scheme is also used for HEMT-based MMICs requiring sub-micron gates.

5.5.2.1 Frontside Processing—Device Isolation or Mesa Formation. The devices are isolated from each other by selective etching [62] of doped layers down to the buffer layer or all the way to the substrate. The isolation process involves a number of steps that include resist deposition, photolithographic exposure, development, semiconductor etching, and resist strip and cleaning. The etching results in formation of isolated islands of conducting epitaxial layers or mesas that are surrounded by semi-insulating buffer or substrate material. Device isolation is checked with a simple direct current measurement. This process also reduces the parasitic capacitances and back-gating, and it provides an insulating surface for MMIC passive components.

5.5.2.2 Ohmic Contact Formation. Since HEMTs are large-current and small-voltage devices, the saturation voltage and transconductance are very sensitive to the contact resistance. It is essential that extremely low contact resistances be formed to the 2-DEG that is situated approximately 300–1000 Å (30–100 nm) below the surface to fully utilize the HEMT's potential.

After mesa formation, the ohmic contact areas are patterned using lithographic techniques and then metallized. Ohmic contact areas are either alloyed or non-alloyed [63]. The goal is to dope the surface of the semiconductor sufficiently high to assure that the dominant conduction mechanism between the contact metal and the semiconductor is field emission [64].

For non-alloyed contacts, a metal can be deposited directly on the ohmic area. To obtain good non-alloyed ohmic contacts, a very heavily doped Ge layer or low-band-gap material such as InGaAs is usually used for the capping layer.

The most commonly used elements for alloyed contacts are a judicious combination of gold (Au), germanium (Ge), and nickel (Ni). These are evaporated onto the patterned HEMT wafer to form ohmics for the source and drain electrodes. After removing the unwanted metal through a lift-off process (whereby the unwanted metal is lifted off by dissolving the underlying resist), the wafer is then thermally alloyed in an inert nitrogen (N_2) atmosphere at a temperature between 320 and 450 deg C to form low-resistance ohmic contacts. A rapid thermal annealing (RTA) or furnace annealing technique is typically used for ohmic alloy [65]. The optimum ohmic surface morphology depends on the metallization composition, thicknesses, and alloy cycle. A good ohmic morphology also provides clean, sharply defined ohmic contact edges in the device channel so that the gate can be placed close to the source to minimize the source resistance.

5.5.2.3 Gate Formation. In addition to the short gate length, a small gate resistance is essential to the fabrication of HEMTs for high-gain, low-noise, and high-power applications. The most widely used gate cross-sectional structure is the T-shaped or mushroom-shaped gate formed using a multi-layer resist technique with E-beam lithography [66–69]. In this structure, the small footprint or bottom of the T defines the gate length, and the wider top of the T provides a low resistance.

A trilayer resist system, PMMA/P(MMA,MAA)/PMMA (PMMA is polymethylmethacrylate, and the other materials are copolymers of PMMA) is used to define T-gates [70]. The least sensitive resist is first deposited on the wafer. Then a sensitive resist is deposited, and finally, a thin, relatively insensitive resist is used to define a good lift-off mask. In addition, the trilayer resist system gives good control of the recess slot width.

Following gate lithography and resist development, the exposed HEMT channel area is recessed to achieve the desired channel current and threshold voltage prior to the gate metallization. The recess etching is performed using either a wet chemical etch or a reactive ion etching (RIE) [71,72] technique. The depth to which the gate is recessed is a critical parameter to the HEMT performance. The etching is discontinued when a target source to drain current

is achieved. Figure 5-6 shows the T-shaped resist cavity using the trilayer resist system.

After the recess, the wafer is then metallized, and the lift-off process is performed to form the metal gates. The metal used to create a Schottky barrier must adhere to the semiconductor and possess thermal stability. The gate metal is typically composed of several metal layers to decrease resistivity. The most commonly used gate metallization layers are titanium/platinum/gold (Ti/Pt/Au), with titanium/palladium/gold (Ti/Pd/Au), titanium/molybdenum/gold (Ti/Mo/Au) and chromium//palladium/gold (Cr/Pd/Au) as possible alternative compositions. Because the gates are very small, a scanning electron microscope (SEM) is used before and after gate metallization to measure the gate dimension and inspect for defects produced during the gate-formation process.

Submicron T-gates fabricated using this technique have demonstrated excellent mechanical stability, and they also exhibit extremely low gate resistances. However, for 0.15-μm or less gates, the T-gate resistance rapidly increases [73] and becomes significant compared to the source resistance. Thus for extremely short gate lengths, a trade-off between the gate resistance and the gate length must be made. An overview of the gate formation process for FETs can be found in Weitzel's review article [74]. Figure 5-7 shows the lifted-off T-shaped gate using the PMMA/P(MMA,MAA)/PMMA trilayer resist system.

To increase the contact conductivity, to simplify bonding to the device and to add MMIC components, such as, inductors, capacitors, and transmission lines, it is necessary to add more metal. Addition of thicker metallization and higher levels of metallization requires additional processing steps such as, resist deposition, photolithographic exposure, development, thick metallization evaporation, lift-off, etc. Au is usually used because of its good conductivity in combination with another metal like Ti or Cr. The Ti or Cr layer, usually 200–1000 Å (20–100 nm) thick, is applied first to provide good adhesion to the III-V semiconductor.

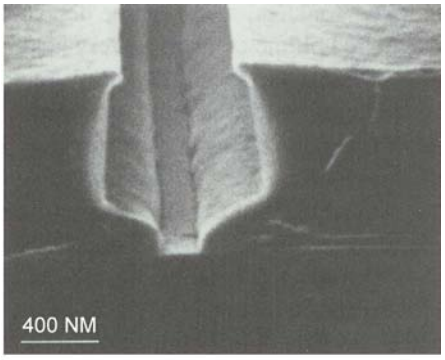


Fig. 5-6. T-shaped resist cavity using the trilayer resist system.

Fig. 5-7. T-shaped gate structure formed using the trilayer resist system.

5.5.2.4 Dielectric Deposition and HEMT Passivation. The device channel area is susceptible to surface damage, chemical and mechanical. Long-term degradation can occur through oxidation or particulate contamination and/or damage during handling and probing. Dielectric films such as polyimide, silicon nitride (Si_3N_4) and silicon dioxide (SiO_2) are commonly used to seal and protect the surface, keeping humidity, chemicals, gases, and particles away from sensitive areas of the device. Device protection or passivation (as it is named), requires a continuous, uniform, low-loss dielectric film. The polyimide film can be spun on the wafer, while Si_3N_4 and SiO_2 can be deposited using plasma-enhanced chemical vapor deposition (CVD).

For capacitors in a HEMT MMIC, a certain dielectric film thickness is needed to achieve the desired capacitance and to improve device and circuit reliability. A thick passivation improves reliability by eliminating pin-holes but reduces device performance by introducing extra parasitic input and feedback capacitance between the gate and drain. A thick passivation degrades the device noise figure, gain, and possibly power-added efficiency, especially at millimeter-wave frequencies. The final dielectric film thickness for HEMTs or MMICs is determined by a trade-off among the device and circuit reliability, performance, and capacitor requirements for the application.

5.5.2.5 Backside Processing. The last steps in HEMT fabrication are wafer thinning, via-hole formation, and dicing. The substrate is thinned to reduce its thermal impedance, to improve its mechanical handling, and to facilitate

transmission line and via-hole formation. A final substrate thickness of 3 to 4 mils (0.0762–0.1016 mm) is typically used for microwave low-noise HEMTs and MMICs. Uniform wafer thinning is the key to a high-yield via-hole process. Via-holes provide low-inductance source grounding which is critically important for high-frequency power HEMTs and low-noise and power MMICs.

The wafer thinning requires mounting of the wafer, usually with the wax frontside down, on a carrier such as glass, quartz, or silicon. Thinning the wafer is accomplished by mechanically lapping the backside of the wafer with a slurry of water and grit, usually silicon carbide (SiC), between the wafer and a flat plate, usually glass. To obtain a smoother surface, the wafer can either be further lapped with a finer grit or chemically polished.

After the wafer has been thinned and polished, the backside is patterned to open holes corresponding to the desired via locations. The via-hole pattern is defined with photolithography using infrared light for aligning the backside pattern to the frontside pattern. (Most III–V semiconductors are transparent to infrared light.) Via-holes can be formed with a wet-chemical etch or with RIE techniques. Compared to the wet-chemical process, the RIE via-hole process is less sensitive to the uniformity of the final substrate thickness and also provides smaller vias with controlled etch profiles. Figure 5-8 shows typical via-holes etched through a 4-mil-thick (10-mm) substrate in a HEMT using the wet-chemical approach. After the via-hole formation, the backside of the wafer is metallized. The chips are then separated through a wafer sawing or a scribe-and-break technique. Figure 5-9 shows a typical HEMT chip after scribe-and-break. Chips with mechanical damage or surface contamination are screened out using both optical microscope and SEM inspections. The HEMT chips are finally electrically tested for dc and RF performance.

To summarize, in order to obtain uniform, high-yield, and reproducible HEMTs, the following are the most critical fabrication issues:

1) Uniform and low-defect density HEMT epilayers
2) Formation of uniform submicron gates on large-size wafers
3) Uniformity of the HEMT gate recess etch.

Each of these areas represents a significant technical challenge and requires a substantial investment in equipment and technology development to insure future HEMT device performance improvements.

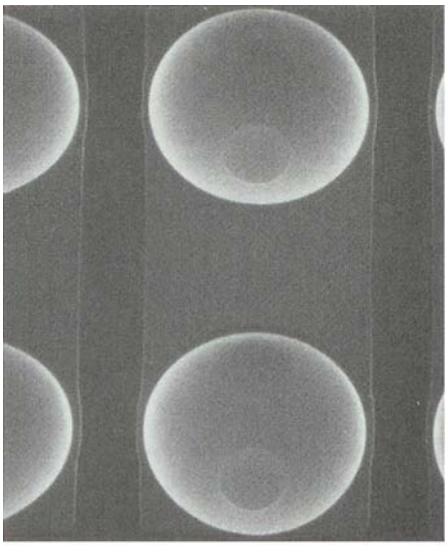


Fig. 5-8. Via-holes etched through a 4-mil-thick (10-µm) GaAs substrate using a wet etch.

Fig. 5-9. Typical HEMT chip after scribe and break. (Note: all chips are inscribed with data identifying position on the wafer, gate number, and gate width.

5.6 HEMT Noise Modeling

To develop ultra-low noise-microwave amplifiers for cryogenic applications, one must have a reliable, reproducible source of state-of-the-art cryogenic devices and the capacity to accurately characterize and model them at the wafer or device level at cryogenic temperatures. Currently, small-signal, semi-empirical, circuit models have proven to be the best way to simulate both the noise and scattering parameters of low-noise HEMTs and FETs. Although this approach is quite successful for the circuit design of cryogenic, low-noise amplifiers, it provides only modest feedback data required for iterative HEMT materials and device optimization.

5.6.1 Noisy Linear Two Port Model

For purposes of circuit modeling and device characterization, a noisy linear two-port device can be represented as a noiseless linear two-port device with the noise sources at the input and/or output [75,76]. Depending upon the utility of the representation, the internal (voltage and/or current) noise sources can be placed at the input or output port of the noiseless network. Figure 5-10 shows a convenient representation that leads to four noise parameters ($T_{\min}$, R_{opt}, X_{opt}, and R_n) that can be determined from the measurement of noise temperature as a function of input match, Z_g. It consists of a series noise voltage (e_n) and shunt noise current (i_n) sources at the input [77]. In this representation, the two-port device's noise parameters are given by the equivalent noise resistance (R_n), noise conduction (g_n), and correlation coefficient (r),

$$R_n = \frac{\langle e_n e_n^* \rangle}{4kT_0B} \tag{5.6-1}$$

$$g_n = \frac{\langle i_n i_n^* \rangle}{4kT_0B} \tag{5.6-2}$$

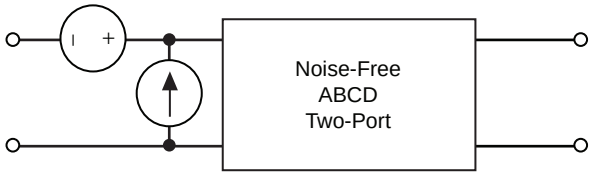


Fig. 5-10. ABCD representation of noisy two-port using voltage and current noise sources at the input.

and

$$r = \frac{\langle e_n i_n \rangle}{\sqrt{\langle e_n e_n^* \rangle \langle i_n i_n^* \rangle}} \tag{5.6-3}$$

where $T_0 = 290$ K, k is Boltzmann's constant, and B is the noise bandwidth.

The noise temperature (T_n) of the two-port device is driven by a generator impedance (Z_g) and is given by the expression

$$T_n = T_{\min} + \frac{T_0 g_n}{R_g \left| Z_g - Z_{\text{opt}} \right|^2} \tag{5.6-4}$$

where Z_{opt} is the optimal generator impedance that yields a minimum noise temperature and $Z_g = R_g + jX_g$ is the generator impedance (where R_g is the real component and jX_g is the imaginary component). The relationship between the first set of noise parameters and those in the above expression is given by the following equations

$$X_{\text{opt}} = \frac{\text{Im}(C)}{g_n} \tag{5.6-5}$$

$$R_{\text{opt}} = \sqrt{R_n / g_n - X_{\text{opt}}^2} \tag{5.6-6}$$

and

$$T_{\min} = 2T_0 \left[g_n R_{\text{opt}} + \text{Re}(C) \right] \tag{5.6-7}$$

where

$$C = r\sqrt{R_n g_n} \tag{5.6-8}$$

In principle, the above noise parameters (Z_{opt}, $T_{\min}$, and R_n) for FET and HEMT devices can be determined by measuring the noise temperature for four or more different known source impedances at a given frequency. In practice, however, since there are errors associated with the source impedance and the noise temperature measurements, additional measurements are usually taken to improve the statistics.

Next, the above equations are rewritten in terms of parameters (A, B, C, and D) that linearize the noise figure or noise temperature expression [78]. The noise parameters are then computed by linear regression analysis of the data. The noise parameters, along with the scattering parameters, can then be utilized for the optimum design of an amplifier circuit.

5.6.2 Semi-Empirical Small Signal Noise Models

In principle, the measured small signal noise and scattering parameters at relevant bias settings as well dc characteristics can be tabulated and utilized for the low-noise amplifier design. The main advantage of a tabular model is that no equivalent circuit needs to be extracted or optimized. However, this approach has some clear disadvantages. First, it may require a large amount of memory for multiple bias settings. Secondly, tabulated data cannot be extrapolated to higher frequencies where measurements may not be possible, and thirdly, tabular models cannot be scaled.

In contrast, semi-empirical models require some experimentally determined fitting factors but do not have these limitations. The goal of semi-empirical circuit models is to find the minimum number of idealized, frequency independent circuit components that reasonably represent complex physical processes so that the equivalent circuit accurately models noise and scattering parameters.

Analytical models that consider fundamental semiconductor steady-state transport properties usually only treat thermal noise within the channel, and thus only model the intrinsic HEMT. These models are progressively more complex treatments of van der Ziel's original work [79,80]. The numerical noise model approach taken by Cappy et al. [81] takes into account electron dynamics and adequately explains noise temperature results at room temperature. Joshin's analytical one-dimensional electron transport noise model [82] also reasonably explains noise temperature results at room temperature and suggests that drain noise current is nearly canceled by induced-gate-noise current due to the asymmetric distribution of noise generation along the HEMT channel. However, the dependence of the measured noise temperature on device parasitics (pad capacitances, inductances, and resistances) and input circuit impedance complicates the full evaluation of numerical and analytical models.

5.6.2.1 PRC Model. The PRC model, based on the work of Pucel et al., is a current source model, where the current sources are connected to the input and output ports of the intrinsic transistor [83,84]. The current sources are correlated and the correlation is imaginary. Figure 5-11 shows the intrinsic HEMT equivalent circuit model with the PRC model current noise sources. The advantages of the PRC model are its close connection to the physical processes

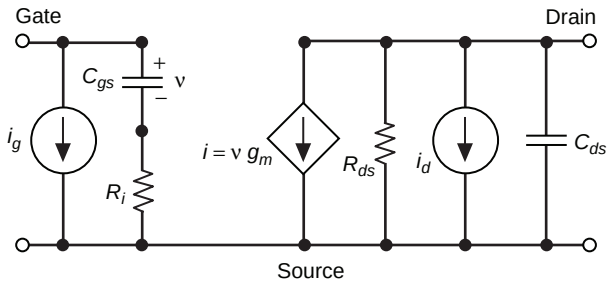


Fig. 5-11. Equivalent circuit of intrinsic HEMT showing input and output current noise sources for PRC model.

in the device and simplification of the model extraction process. The model is named after the parameter names *P*, *R*, and *C* and are given below along with the minimum noise figure, $F_{\min}$:

$$P = \frac{\left\langle i_d i_d^* \right\rangle}{4kT_0 B g_m} \tag{5.6-9}$$

$$R = \frac{\left\langle i_g i_g^* \right\rangle g_m}{4kT_0 B\left(\omega C_{gs}\right)^2} \tag{5.6-10}$$

$$C = -j\frac{\left\langle i_g i_d^* \right\rangle}{\sqrt{\left\langle i_g i_g^* \right\rangle \left\langle i_d i_d^* \right\rangle}} \tag{5.6-11}$$

and

$$F_{\min} = 1 + 2\sqrt{PR\left(1 - C^2\right)}\frac{f}{f_T} + 2g_m R_i P\left(1 - C\sqrt{\frac{P}{R}}\right)\left(\frac{f}{f_T}\right)^2 \tag{5.6-12}$$

5.6.2.2 Fukui Model. The most well established model for device and circuit optimization is the semi-empirical one developed by Fukui [85]. In the Fukui model, the noise parameters are simple frequency dependent functions of the equivalent small-signal intrinsic circuit elements (transconductance, g_m, gate-to-source capacitance, C_{gs}, and source and gate resistances, R_s and R_g). These circuit elements are (in turn) analytic functions of the device's geometrical (e.g., gate and channel dimensions) and material (e.g., doping

concentration and active channel thickness) parameters. The semi-empirical approach of Fukui yields the following expressions for the noise parameters

$$T_{\min} = \frac{k_1 \omega T_0 C_{gs}}{2\pi} \sqrt{\frac{\left(R_g + R_s\right)}{g_m}} \tag{5.6-13}$$

$$R_n = \frac{k_2}{g_m} \tag{5.6-14}$$

$$R_{\text{opt}} = k_3 \left(\frac{1}{4 g_m} + R_g + R_s \right) \tag{5.6-15}$$

and

$$X_{\text{opt}} = \frac{2\pi k_4}{\omega C_{gs}} \tag{5.6-16}$$

where k_1, k_2, k_3, and k_4 are fitting factors that are determined experimentally. Although the Fukui model is widely used by device designers and served to guide the development of the first cryogenic HEMT device for the Voyager 2 encounter with Neptune, the model is equivalent-circuit dependent and provides little insight into the physics of noise in HEMTs.

5.6.2.3 Pospieszalski Noise Model. The Pospieszalski model uses small signal circuit elements that yield closed-form expressions for the noise parameters [86]. This model introduces frequency-independent equivalent temperatures T_g and T_d for the intrinsic gate resistance R_i and drain resistance R_{ds}, respectively. The noise processes are modeled by R_{ds}, and T_d is the only free parameter, while T_g is taken to be the ambient temperature of the device. The equivalent noise model for the intrinsic HEMT is shown in Fig. 5-12 while the noise correlation relations are given below [87]

$$R_{ds} = \frac{\left\langle i_d i_d^* \right\rangle}{4 k T_d B} \tag{5.6-17}$$

$$R_i = \frac{\left\langle i_g i_g^* \right\rangle}{4 k T_g B} \tag{5.6-18}$$

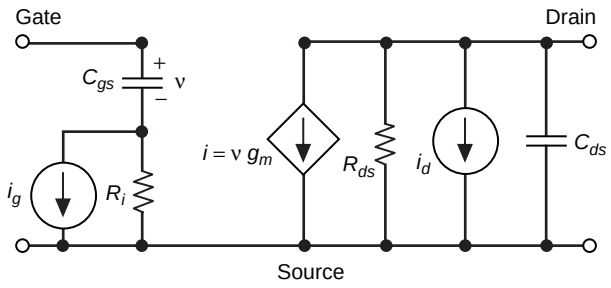


Fig. 5-12. Equivalent circuit of intrinsic HEMT for Pospieszalski noise model.

$$0 = \left\langle i_d i_g^* \right\rangle \tag{5.6-19}$$

The noise parameters for the above intrinsic circuit are given by the following expressions

$$R_{\text{opt}} = \sqrt{\frac{T_g R_i R_{ds}}{T_d}\left(\frac{g_m}{\omega C_{gs}}\right)^2 + R_i^2} \tag{5.6-20}$$

$$X_{\text{opt}} = \frac{1}{\omega C_{gs}} \tag{5.6-21}$$

$$T_{\min} = 2\frac{\omega C_{gs}}{g_m}\sqrt{\frac{T_d T_g R_i}{R_{ds}} + \left(\frac{T_d R_i \omega C_{gs}}{R_{ds} g_m}\right)^2} + 2\frac{T_d R_i}{R_{ds}}\left(\frac{\omega C_{gs}}{g_m}\right)^2 \tag{5.6-22}$$

and

$$R_n = \frac{T_{\min}}{T_0} R_i + \frac{T_d}{T_0}\frac{1}{R_{ds} g_m^2}\left(1 + \left(\omega C_{gs} R_i\right)^2\right) \tag{5.6-23}$$

The utility of this model is that it allows prediction of the noise parameters for a broad frequency range from a single frequency noise-parameter measurement at a given temperature. Although the Pospieszalski model only considers thermal noise sources and does not take into account the correlated noise between the gate and the drain, it is an accurate model for high-quality devices operated at low-noise bias. Additionally, for devices operated at higher drain currents, T_g

is taken to be higher than the ambient temperature; and gate leakage current can be modeled with a resistor across C_{gs} and R_i at an elevated temperature [88].

5.6.2.4 Monte Carlo HEMT Noise Model. Although considerable research has been conducted on the noise performance and theory of HEMTs, a noise model that is useful for cryogenic-device optimization and circuit design is not yet available. However, Monte Carlo techniques using microscopic theory have recently been successfully applied to calculate and predict the noise parameters of HEMTs at room temperature [89].

The most valuable feature of Monte Carlo techniques is tha t these techniques enable investigation of the physical origins of noise in semiconductor devices and allow differentiation between noise temperature and electron temperature [90]. The essential feature of the Monte Carlo noise temperature model is to follow the evolutionary motion of the charge carriers in time domain while taking into account all of the important microscopic scattering mechanisms associated with the semiconductor material (ionized impurities, inter-valley transitions, phonons, alloy, electron-electron, etc.) [91]. The Monte Carlo simulator calculates (1) instantaneous velocities and energies, (2) mean velocity and mean energy from the instantaneous values, (3) instantaneous velocity fluctuations from (1) and (2), and then (4) the spectral density of velocity fluctuations. The intrinsic noise temperature, T_n, is calculated from the spectral density of velocity fluctuations, $S_v(f)$, and electron temperature, T_e, is c alculated from the average energy using the equipartition principle as shown below

$$T_n(f) = \frac{qS_v(f)}{4k\mu_D(f)} \tag{5.6-24}$$

and

$$T_e = \frac{2}{3k}\langle E \rangle \tag{5.6-25}$$

$$S_v(f) = 4\int_0^\infty C(t)\cos(2\pi ft)dt \tag{5.6-26}$$

$$C(t) = \langle \delta v(t')\delta v(t'+t) \rangle \tag{5.6-27}$$

where μ_D is the differential mobility, $C(t)$ is the autocorrelation function of velocity fluctuations and $\delta v(t) = v(t) - < v >$. In fact, Pantoja et al. [90]

demonstrated very good a greement between measured and modeled noise temperature (T_n) at 300 K and 77 K for GaAs at 10 GHz.

Furthermore, Mateos et al. [89] performed a complete Monte Carlo analysis on a low-noise 0.1-μm T-gate AlInAs/GaInAs HEMT. The modeled dc and rf properties showed exceptionally good a greement with measurements. The model included effects such as degeneracy, surface charges, T-shape of the gate, presence of dielectrics, and contact resistances. Moreover, the extrinsic parameters of the device were added to the intrinsic small-signal equivalent circuit, allowing a realistic calculation of the dc characteristics and the noise and scattering parameters. Although, the simulations take hours on a personal computer, the reliability allowed by this M onte Carlo simulator will enable faster and cheaper optimization of HEMT devices.

5.7 LNA Development

The development and demonstration of cryogenic, InP HEMT-based front-end low-noise amplifiers for the DSN also require accurate LNA component characterization and modeling from 1 to 100 GHz at physical temperatures down to 12 K. The characterization and modeling starts with the individual HEMT chip, RF and DC bias components, proceeds to the multi-stage HEMT LNA module, and it culminates with the complete cryogenic front-end receiver package for the antenna.

5.7.1 Device Characterization—Cryogenic Probe Station

The development of a complete on-wafer cryogenic microwave measurement system has been primarily driven by the need for

1) Greater understanding of the device physics in a dvanced high spe ed transistor technologies
2) Continued advancement of cryogenic LNAs with noise temperatures less than five times the quantum limit ($T_n < 5hf / k$) for ground-based and space-based applications
3) Hybrid and monolithic microwave integrated-circuit (MMIC) semiconductor-superconductor circuits

The cryogenic microwave system uses coplanar waveguide probes in a vacuum station coupled to a vector network analyzer for scattering parameter measurements, and a noise meter and noise test set with a noise system for microwave noise parameter measurements. The cryogenic probe measurement system, shown schematically in F ig. 5-13, contains ports for RF cables, thermometers, vacuum pumps, dry nitrogen back-fill lines, coplanar probes with manipulators, and a closed-cycle refrigerator cold head. The probe body rests on a copper block attached to a fiberglass post. The fiberglass reduces the

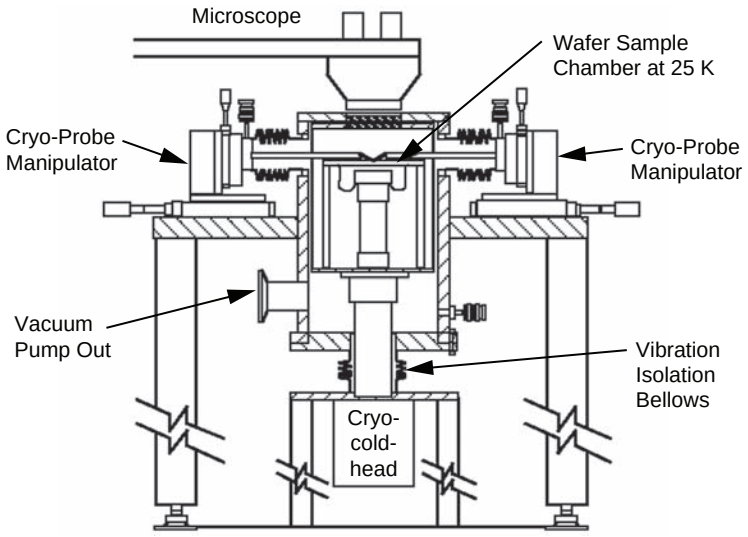


Fig. 5-13. Schematic of the cryogenic probe measurement system [106].

thermal load, and copper braiding from the cold head thermally anchors the probe to the 12-K cold station, assuring sample temperatures of 12 to 20 K. The mechanical and thermal stability of the wafer stage is established by supporting it on fiberglass posts above the cold head and thermally anchoring it to the cold station with flexible copper braids.

The most important feature of this design is the incorporation of a closed-cycle helium refrigeration system. The first successful designs of on-wafer cryogenic systems used open-cycle cooling to reduce start-up costs and avoid mechanical vibrations. However, for long term use at the rate of one cool down per week, a closed-cycle system is significantly less expensive.

Decoupling and damping of the vibrations from the cold head to the probe station are accomplished with a two-dimensional bellows and vibration mounts. This system allows small-signal microwave measurements from dc to 40 GHz over a physical temperature range of 16 to 300 K. Since the microwave hardware is insulated by vacuum, there is no frost buildup or large thermal gradients, resulting in a system that is accurate, reliable, and flexible (active and passive discrete devices, as well as MMICs, can be measured). Figure 5-14 is a photograph of the cryogenic probe test chamber showing the input (at left) and the output (at right) coplanar microwave probes with a calibration standard and test HEMT devices epoxied to an alumina substrate (near the center).

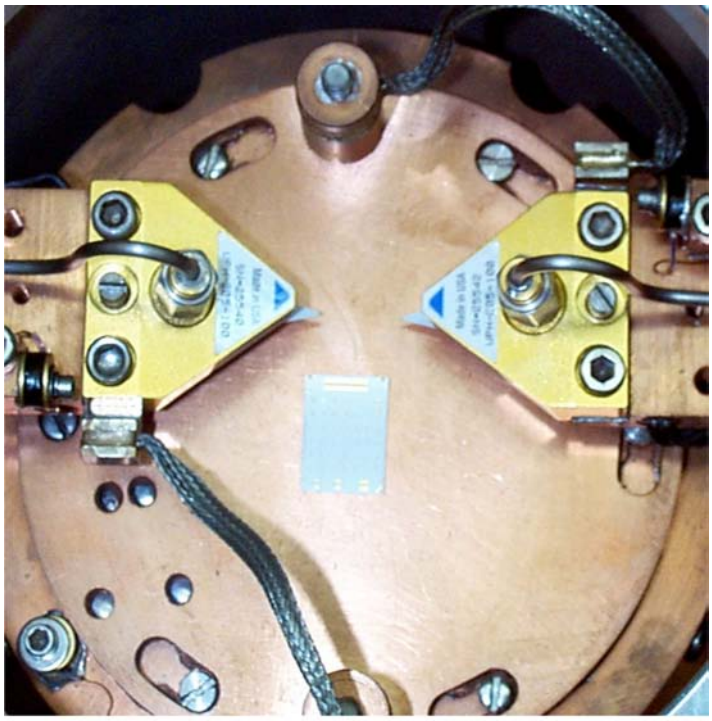


Fig. 5-14. Photograph of the cryogenic probe station showing probes and calibration standard.

5.7.2 Device Characterization—Cryogenic Probe Station Calibration

The key to accurate on-wafer microwave and millimeter-wave measurements is proper establishment of the electrical reference plane. The reference plane can be determined with either the line-reflect-match (LRM) or the short-open-load-thru (SOLT) calibration method utilizing an impedance standard substrate (ISS), which is available from Cascade Microtech [92]. The LRM method requires fewer standards, and the reflect standards need not be well known. In addition, experience has shown that the LRM calibration is slightly better in accuracy than is the SOLT at cryogenic temperatures. In the SOLT method, the short standard introduces uncertainty in the reference plane location because of sensitivity to probe tip placement. The LRM method obviates this problem by replacing the short with an open and by having the probe tip held approximately 10 mils (0.25 mm) above the substrate during the calibration sequence.

The measurement accuracy is also directly related to the calibration conditions. Thermal gradients across the gold-plated ceramic probe tips and

coax-to-coplanar transitions [93] alter the electrical characteristics of the measurement lines. The process of cooling the sample in the laboratory can also produce changes in the calibration. The network analyzer typically requires a new calibration if the ambient laboratory temperature varies by greater than 1 deg C. The proximity of the cooled probe system to the network analyzer can change the ambient environment (both temperature and humidity). The combined results of these effects are appreciable errors in cryogenic temperature measurements. For example, room-temperature calibrations have been shown to introduce as much as a 20-percent [94] error in the cryogenic measurement. In addition, measured results have been reported with room-temperature calibrations with moding effects (deviations from one-pole roll-off) [93–96].

Since the microwave probe offers a significant thermal load to the device under test (DUT), it is evident that the chuck and device temperatures are different. For example, with a chuck temperature of 20 K and the probes contacting the DUT, a DUT temperature as high as 50 K has been observed. This large temperature differential affects calibration and skews the interpretation of data collected at different device temperatures.

The solution to maintaining calibration integrity and achieving low sample temperatures is to thermally anchor the probe body and perform cryogenic calibrations. By thermally anchoring the probe to the cold head at 12 K, the thermal load to the DUT is minimized. The remaining microwave hardware (connectors, cables, and input to the automatic network analyzer (ANA)) are thermally isolated via vacuum and stainless steel hardware. The thermally anchored probe can be calibrated at specific temperatures during the measurement cycle. This eliminates the problem of moding and allows accurate correlation of DUT temperature and measured characteristics [97,98].

For this initial investigation of on-wafer noise parameter measurements at cryogenic temperatures, only the probe tips are cooled while the impedance state generator and solid-state noise source (both commercially available) are kept at room temperature (several wavelengths away from the DUT). In this configuration, the input losses introduce noise comparable to or greater than the noise of the DUT and reduce the range of available impedance states. For example, in the frequency range of 2 to 18 GHz for cryogenic temperatures, the worst-case noise temperature error is ±25 K, while device noise temperatures are typically under 10 K. Although this configuration does not provide accurate single-frequency noise parameter measurements, it does provide for fast and efficient broadband (2- to 18-GHz) on-wafer measurements [99].

The most accurate and repeatable method of measuring noise parameters at cryogenic temperatures is to place the impedance generator within a wavelength of the DUT input. The equivalent noise temperature of the noise source must also be comparable to the DUT noise temperature. This approach

would, however, require development of a cryogenic noise generator and noise source.

5.7.3 Device Characterization Measurements and Models

The InP HEMTs developed by TRW for the DSN are grown by molecular beam epitaxy on 3-in. (7.6-cm) semi-insulating InP wafers. The cross-section of this device is shown in Fig. 5-15. First, to inhibit impurity diffusion into the active region and to improve carrier confinement, two buffer layers (one of InAlAs and the other InP) are grown on the semi-insulating InP wafer. Then the active undoped $In_{0.65}Ga_{0.35}As$ (65% indium concentration) layer is grown followed by another spacer layer of InAlAs to further improve carrier confinement and to reduce donor ion scattering. The donor Si atoms are added in a single atomic layer in the next undoped InAlAs Schottky layer. The last two layers are n-type InGaAs cap layer and a heavily doped InGaAs layer to provide ohmic contacts. The devices are then passivated with a thin SiN layer.

A key step in the process of developing InP HEMT hybrid and MMIC LNA modules is the accurate, broadband, cryogenic characterization and modeling of active and passive circuit components. From 0.05 to 40 GHz, device and component measurements covering room to cryogenic temperatures are made with a cryogenic, coplanar waveguide probe station. To enhance and

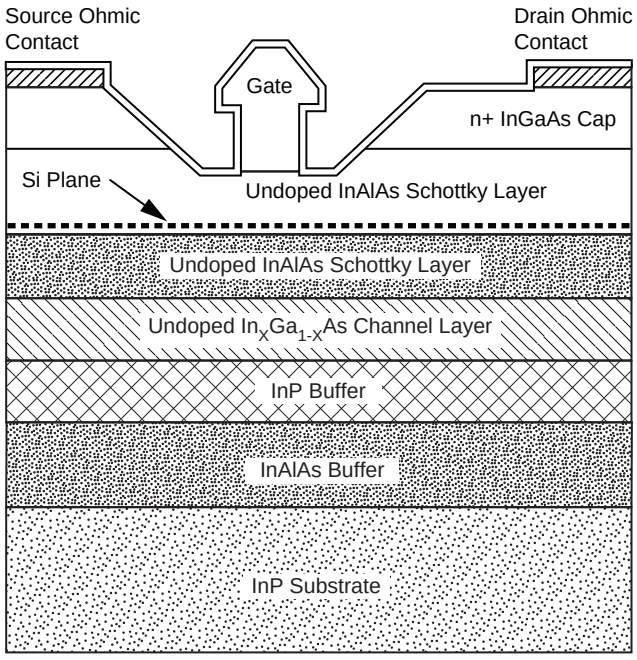


Fig. 5-15. Cross section of TRW Cryo-03 InGaAs/InAlAs/InP HEMT device [13].

extrapolate component models up to 100 GHz, a 2½ dimensional electromagnetic (2½ DEM) simulator was used.

There are three main methods for small-signal parameter extraction:

1) Complete parameter extraction from calculations on the measured S-parameters [100]
2) Parameter extraction based upon iterative computer optimization routines
3) A combination of (1) and (2) where as many elements as possible are quickly identified, providing constraints to the software optimization routine that serves as the validity check.

All of the above methods can be applied to the standard HEMT small-circuit model, the hybrid π-circuit topology [101]. The small-signal elements in Fig. 5-16 are broken down into the intrinsic and extrinsic elements. The intrinsic elements are the transconductance (g_m), output r esistance, (R_{ds}), gate-source capacitance, (C_{gs}), gate-drain capacitance, (C_{gd}), drain-source capacitance, (C_{ds}), gate-source resistance, (R_i), and delay time (τ). These elements are b ias dependent and important to the understanding of device behavior. The extrinsic elements are independent of bias and include the three terminal inductances (L_g, L_d, and L_s), the contact resistances for the three terminals (R_g, R_d, and R_s), and the parasitic pad capacitances (C_{pg} and C_{pd}).

The simplest and most straight-forward method is the complete parameter extraction method. It has been shown that, with a sequence of microwave and dc measurements, each term can be uniquely determined [100]. A so-called

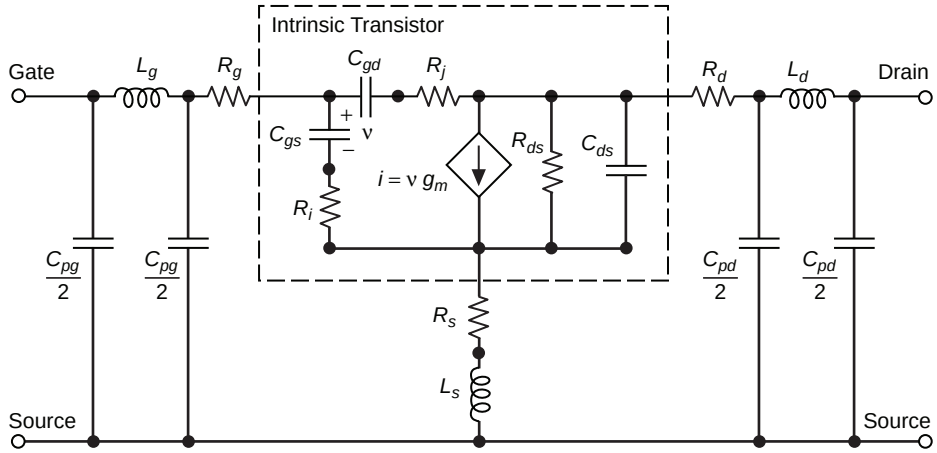


Fig. 5-16. Pospieszalski HEMT circuit model showing both extrinsic and intrinsic circuit elements.

shell technique can be applied to remove the inductances, the pad parasitics, and the contact resistances. For this technique, the S-parameters of the device are measured for a variety of drain-to-source voltage conditions. By using the simplifying assumptions that $(\omega C_{gs} R_i)2 < 0.01$ and that $\omega\tau << 1$, one can determine the parasitic elements uniquely, leaving only the intrinsic circuit to be determined. The drawbacks to this method are that several measurements must be made under different bias conditions and simplifying assumptions must be made, increasing the uncertainty of the extracted parameters.

The second method of extraction is to simply fit the measured S-parameters to the equivalent circuit by iteratively solving for the individual elements. The disadvantage to this technique is that, with such a large parameter space, it is very difficult to uniquely identify each element. This frequently produces unrealistic values for a number of the circuit elements. The most desirable and accurate method is to apply the techniques from the complete parameter extraction method coupled with small-signal circuit optimization routines that serve as a validity check. The important parasitics are estimated by measuring a simple test structure and by independent dc measurements.

Although a test structure for the parasitic elements was not available, the method used is based on the methods developed by [102–104]. Extraction starts with determination of the parasitic pad capacitances with the HEMT biased under ColdFET ($V_{ds} = V_{gs} = 0$) and pinch off conditions (i.e., $V_{ds} = 0$ and $V_{gs} < V_{pinch\text{-}off}$). Next, the device is biased under ColdFET and full channel conditions (i.e., $V_{ds} = 0$ and $V_{gs} > 0$). The gate resistance, sheet resistivity, gate width, and gate length are used to estimate the initial values of parasitic resistances. This is sufficient to extract the rest of the extrinsic components. The intrinsic components of the active HEMT are calculated by fitting the measured S-parameters to the HEMT model. Finally, starting values of T_g and T_d are estimated from measured room temperature device noise parameters and than recalculated based on room temperature and cryogenic LNA module noise and gain measurements [105].

The widely used Pospieszalski HEMT noise model [86], shown in Fig. 5-16, possesses physically realizable elements, and it is symmetric and sufficiently broadband for this work. Figure 5-17 shows the agreement between the measured and modeled scattering parameters at 18 K with (+) illumination and without (–) illumination at low noise bias, $V_{ds} = 0.8$ v, and $I_{ds} = 2.0$ mA, and $V_{gs} = 0.08$ v.

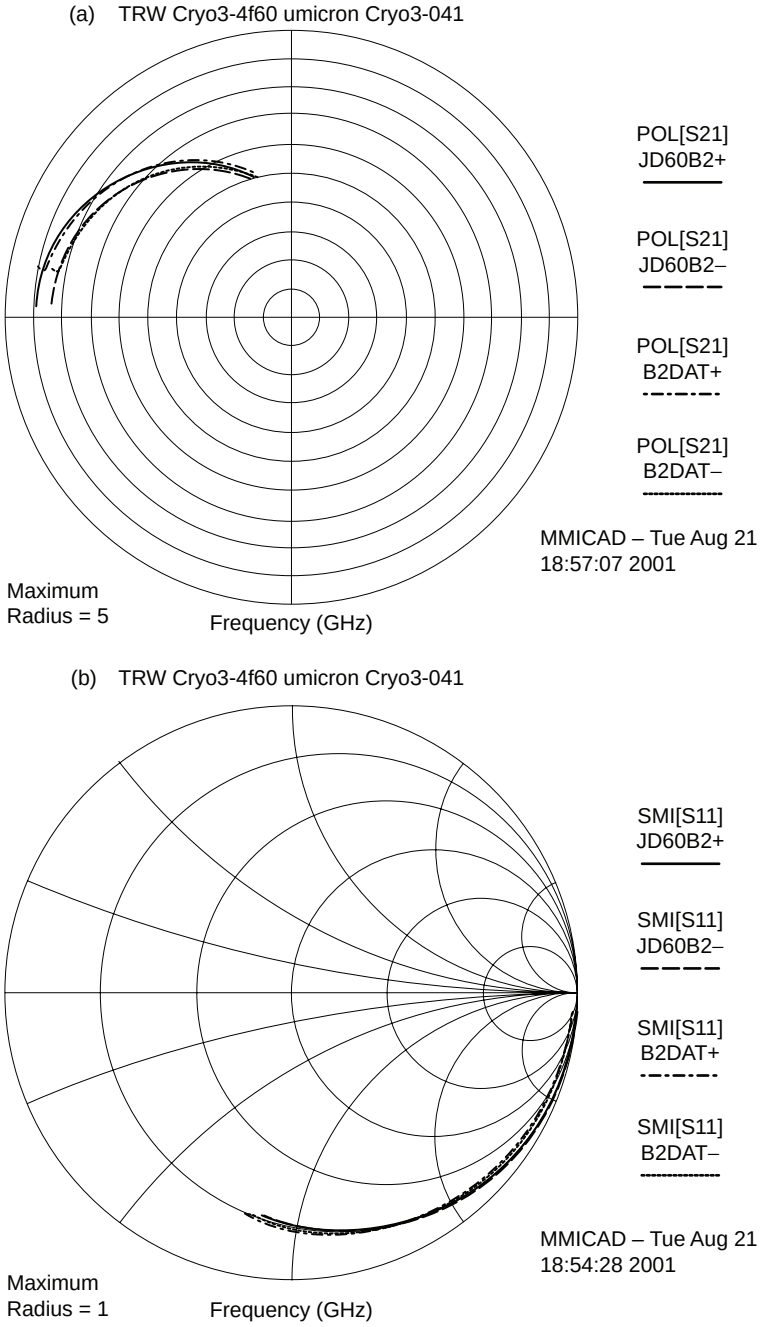


Fig. 5-17. Models versus measured cryogenic (a) gain and (b) reflection of 32-GHz HEMT at low noise bias with a frequency of 1 to 35 GHz (Note: Model and measurement are in excellent agreement).

5.7.4 Passive Component Characterization Measurements and Models

From 1 to 40 GHz, the 2½ DEM simulator was used to verify our cryogenic measurements on passive components and to refine and extend our user-defined cryogenic models employed in M MICAD (Optotek, Ltd.) to 100 GHz. Although measurements to 40 GHz are sufficient for the 8.4-GHz LNA module modeling and design, it is inadequate for the 32-GHz LNA module modeling and design. The maximum frequency of operation, f_{max}, for the 32-GHz device approaches 150 GHz. Thus, at 32 GHz, especially in regards to stability, it is critical to have accurate component models up to 100 GHz. Figure 5-18 shows the measured and modeled response of a 1-kΩ bias circuit resistor at cryogenic temperatures.

5.8 LNA Modeling and Characterization

The design approach and fabrication process for both the 8.4-GHz and the 32-GHz LNA modules are essentially the same with some practical differences.

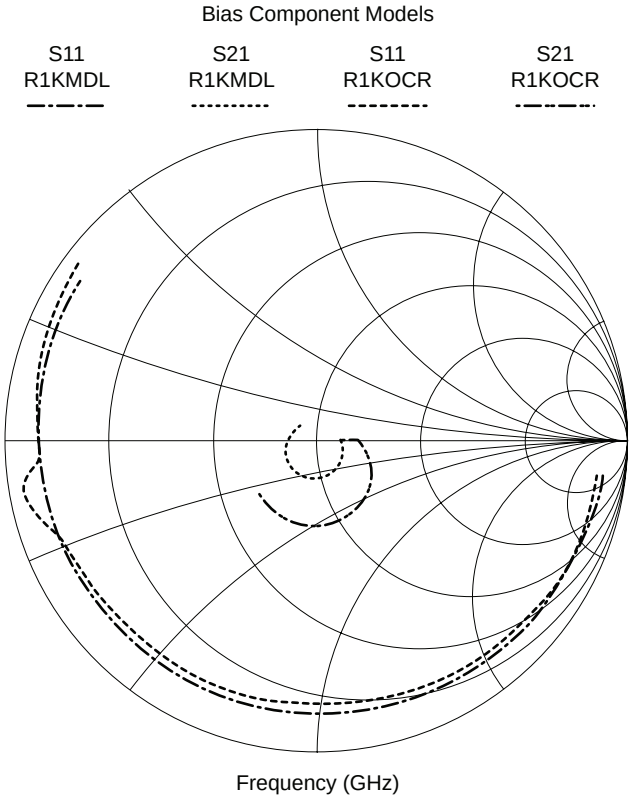


Fig. 5-18. Measured and modeled S-Parameters of 1-kΩ Bias Resistor at 18K.

At 8.4 GHz only three stages are needed, since there is sufficient gain per stage, while the 32-GHz module requires four stages. The 8.4-GHz module uses TRW four-finger 300-μm gate width InP HEMTs in all three stages, while the 32-GHz module uses TRW four-finger 80-μm gate width InP HEMTs in all four stages. Additionally, WR-28 wave guide input and outputs are used at 32 GHz instead of coaxial k-connectors to further reduce RF losses.

The LNA module design goal is to minimize the noise temperature at the DSN band of operation while at the same time maintaining unconditional stability both inside and outside the module's bandwidth. The 8.4-GHz LNA is designed to be unconditionally stable from 0 to 40 GHz, while the 32-GHz LNA is designed to be unconditionally stable from 0 to 100 GHz.

Since all of the HEMT devices used for the LNA designs are unstable (i.e., μ-factor <1 [106] over their usable gain bandwidth), the first step is to stabilize the device at the LNA module band of operation without significantly increasing the device noise temperature. For example, for the 8.4-GHz LNA module design, the first-stage device is first stabilized near 10 GHz by a judicious choice of gate, drain, and source bond wire lengths. Next, the device gate and drain bias networks are used to load the device and control the stability below 10 GHz and from 10 to 40 GHz, respectively. The rest of the stages are similarly optimized. The loaded biased devices then serve as the fundamental circuit building block. A similar procedure is used for the 32-GHz LNA module design. Figure 5-19 shows how the device bond wires and bias network affect and help control the device stability, while Fig. 5-20 shows the trade-off, an associated increase in noise temperature for the 8.4-GHz first-stage device.

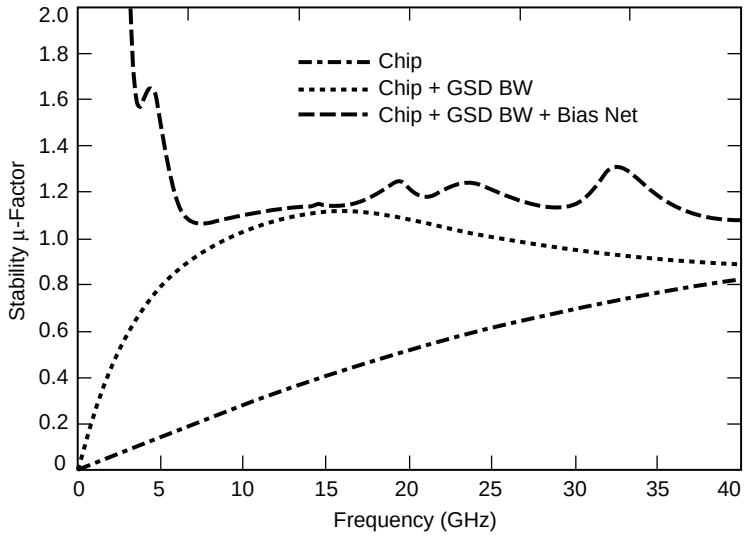


Fig. 5-19. Bondwire and bias circuit stability (μ-factor) affects for 8.4-GHz HEMT at physical temperature of 16K.

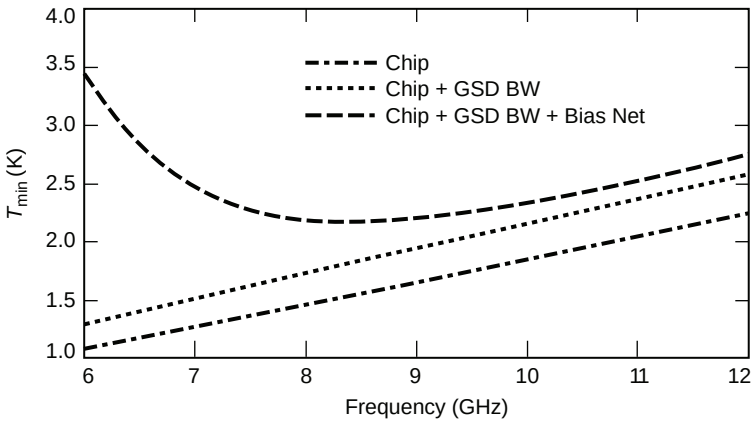


Fig. 5-20. Bondwire and bias circuit minimum noise temperature (T_{min}) trade-off for 8.4-GHz TRW HEMT at physical temperature of 16 K.

Then, each of the stages is iteratively matched to its optimum source and load impedance. At 8.4 GHz the first and second stage are optimized for noise performance while the third stage is optimized for gain and output match. Following initial optimization, the LNA model is loaded at the input and the output with the string of passive microwave components that will be implemented for field use, and the bond wire lengths re-optimized.

A photograph of the 8.4-GHz LNA module is shown in Fig. 5-21. The module carrier is gold-plated brass, the input, inter-stage and output matching circuits are etched on Cuflon, and the dc blocking and bias circuitry use surface-mount thin-film resistors and capacitors.

The completed LNA modules are then characterized in a cryogenic testbed at a physical temperature of 12 K using the cold attenuator method [107]. A photograph of the 32-GHz testbed is shown in Fig. 5-22. In this technique, a 20-dB attenuator connected to the LNA module input is cooled along with the module. The cooled attenuator serves as the cold noise source when the hot-noise source, noise diode, is turned off and eliminates impedance-match errors associated with the noise diode on–off states. The noise and gain are automatically measured using a commercial noise diode and noise figure meter. At 8.4 GHz, the gain measurement error is approximately ±0.1 dB, and the noise error is ±0.3 K, while at 32 GHz the errors are about five times higher. Figure 5-23 shows the measured and modeled noise and gain performance of the 8.4-GHz LNA, while Fig. 5-24 shows a similar plot for the 32-GHz LNA module. The module is subsequently cooled without the attenuator, and the output is monitored for oscillations or instabilities with a spectrum analyzer as the input impedance is varied.

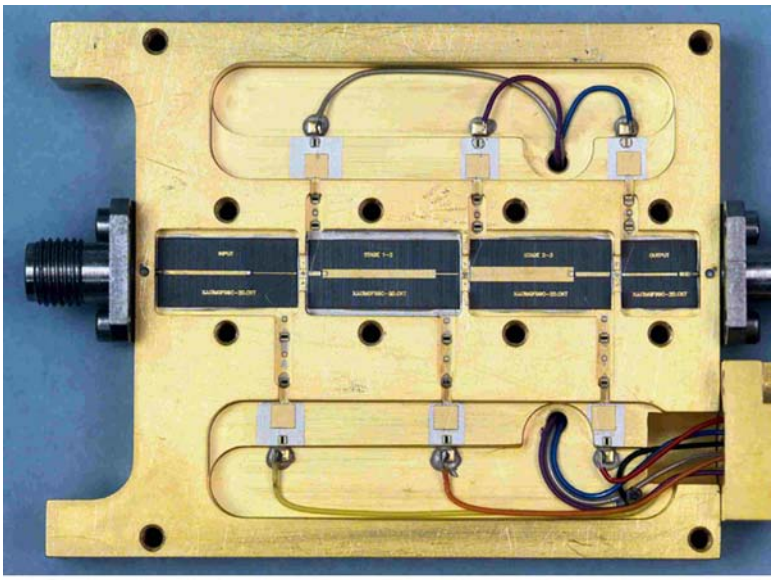


Fig. 5-21. Photograph of three-stage 8.4 GHz InP LNA module (JPL part number JPL 2000-10).

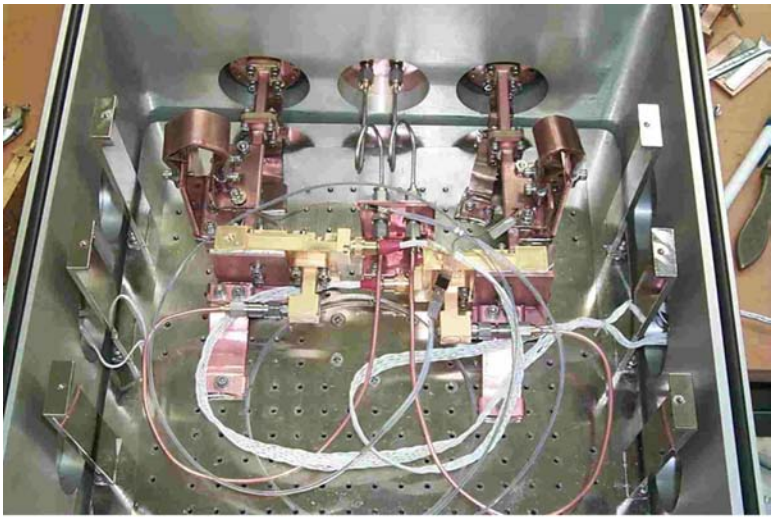


Fig. 5-22. Photograph of the 32-GHz LNA cryogenic testbed.

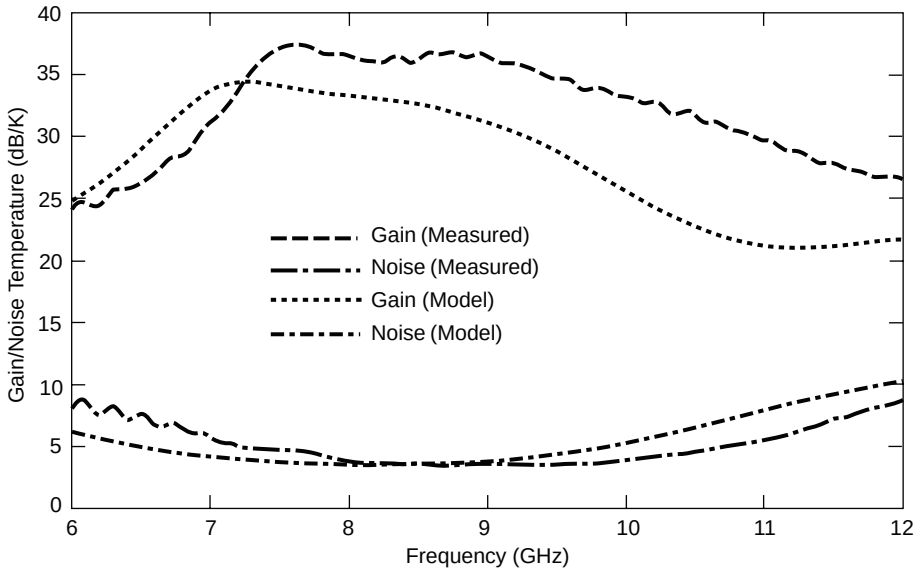


Fig. 5-23. Measured and modeled noise and gain of an 8.4-GHz LNA module (LNA module 1 noise/gain at $T_{physical}$ = 12.5 K model versus measurement).

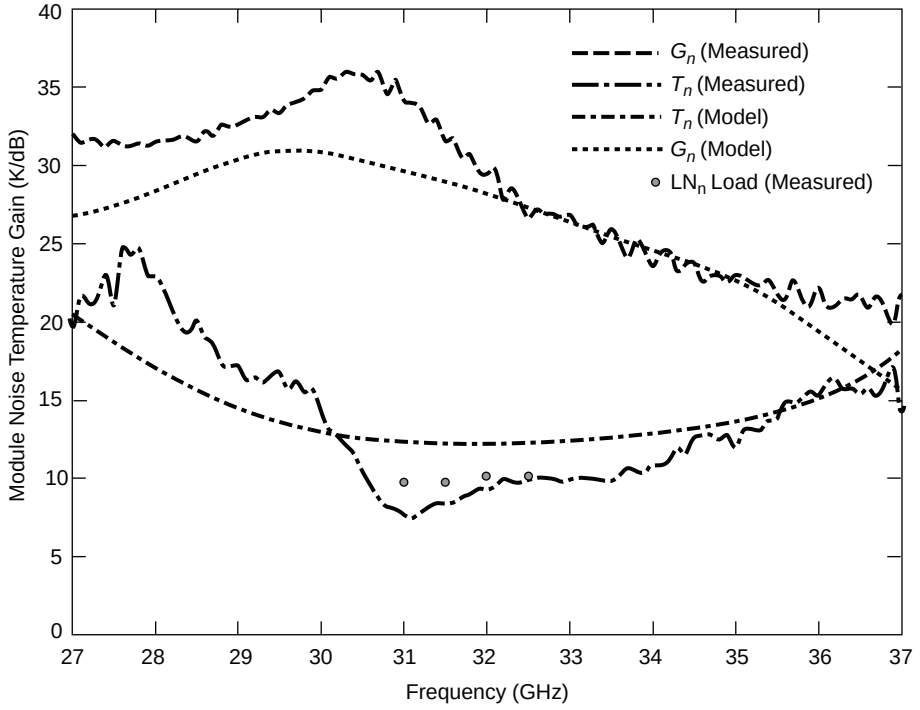


Fig. 5-24. Measured and modeled noise temperature and gain of a 32-GHz LNA module.

5.9 Subsystem Measurements

Additional passive microwave components (filters, isolators, adapters, couplers, and polarizers) are required to implement an LNA module in a DSN antenna due to the deleterious effects of radio-frequency interference (RFI), the need for calibration signals, and the need for a redundant receive capability. Since significant effort is devoted to minimizing the LNA module's noise temperature, an equivalent effort must be expended to minimize the noise temperature contribution of these components to meet the DSN low-noise receiver specifications.

A useful expression to determine the noise temperature contribution of passive two-port networks can be derived from the noise temperature function of the noise wave matrix representation for a passive network [108]. When the network is placed at the input of an LNA, the effective input-noise temperature, T_e, of the cascaded pair is given by the following expression [109]:

$$T_e = \frac{[(L-1)+\Gamma_L{}^2]T_L + LT_{\mathrm{LNA}}}{(1-\Gamma_L{}^2)} \tag{5.9-1}$$

where L = loss ratio, Γ_L = reflection coefficient, and T_L = physical temperature of the passive network, while T_{LNA} = LNA module noise temperature. Hence, in order to minimize the noise temperature contribution of the passive network, it must be well matched, be low-loss, and be kept at the lowest physical temperature possible.

Following complete characterization, the LNA module is integrated and measured with the necessary filters, isolators, and adapters. This cascaded network is placed in a larger testbed and characterized once again using the cold attenuator method described above.

A photograph of the cascaded LNA network is shown inside the testbed in Fig. 5-25. Noise temperature and gain measurement results are shown in Fig. 5-26.

Finally, the LNA module and components are integrated into the closed cycle refrigerator (CCR) package and characterized using the DSN's ambient load/cold sky method [110]. A photograph of the LNA CCR package is shown in Fig. 5-27. (Current systems use a CCR built by Sumitomo Heavy Industries Ltd., which provides 1.5 W of cooling capacity at a physical temperature of 4.2 K.) This method uses the sky as the cold noise source and ambient load as the hot noise source placed over a calibrated feed horn [111] to determine the LNA CCR-package noise temperature. Noise temperature measurements (right-hand circularly polarized (RCP) and left-hand circularly polarized (LCP)) of the DSN LNA CCR package are shown in Fig. 5-28. These measurements are referenced to the room temperature waveguide input flange.

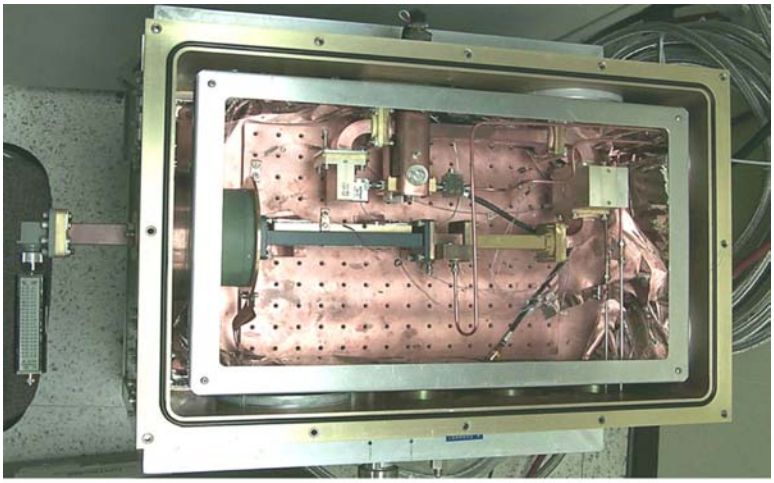


Fig. 5-25. Photograph of the cascaded 8.4-GHz LNA assembly in testbed.

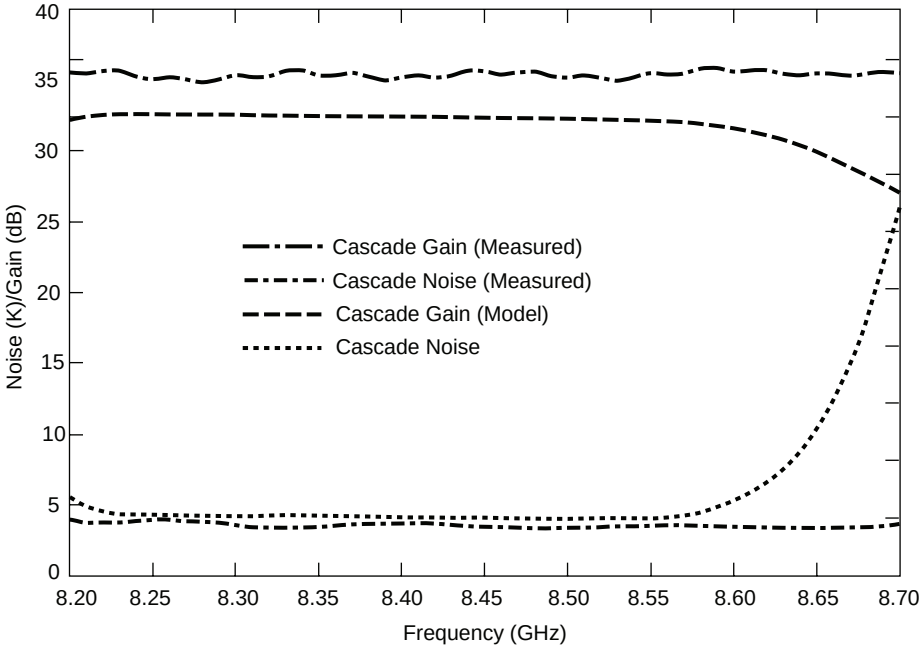


Fig. 5-26. Cascaded LNA assembly noise and gain (model versus measured) at 9.6-K physical temperature.

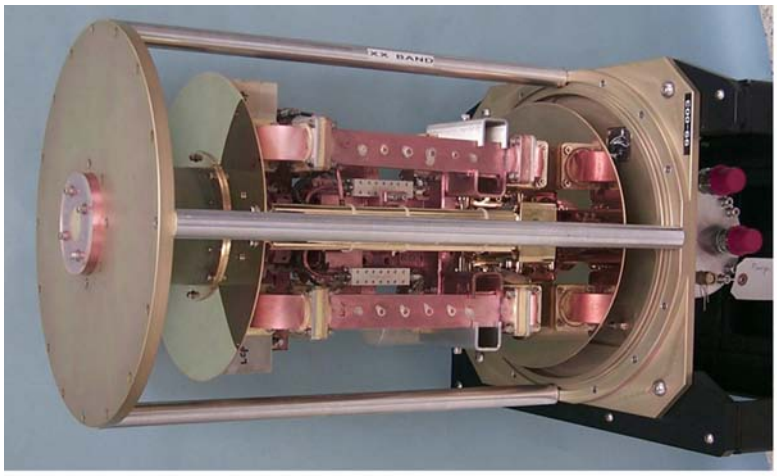


Fig. 5-27. DSN dual-channel InP HEMT/CCR LNA package.

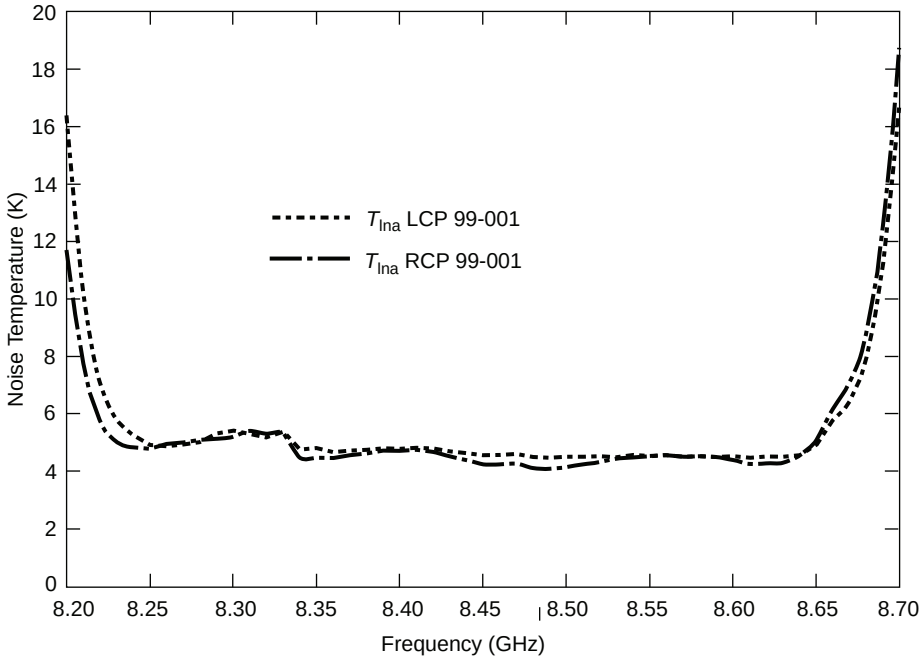


Fig. 5-28. Measured noise temperature of DSN dual-channel InP HEMT/CCR LNA package (referenced to room temperature).

Once implemented in the field these front-end receivers demonstrated operating system noise temperatures of 17 K at 8.4 GHz on a 70-m Cassegrain antenna and 39.4 K at 32 GHz on a 34-m beam waveguide antenna, both at zenith [112]. (To date, three more 8.4-GHz, two more 32-GHz comparable

front-end receivers, and six X/X/Ka (X-transmit, X-receive, and Ka-receive) HEMT/CCRs systems have been delivered to the DSN.) At 8.4 GHz, the T_{op} for these units has been within 1 K of the predicted value, while at 32-GHz it has been within 2.5 K. This close agreement between predicted and measured performance is a testament to the measurement and modeling accuracy required to successfully develop these ultra-low noise cryogenic InP-HEMT-based CCR receive systems.

5.10 Conclusion

The work reported here is the by-product of a program at JPL to develop cryogenic InP HEMTs and MMICs for both ground-based and spaceborne radiometers and receivers. This work is thoroughly investigating the device parameters that will yield the best InP HEMTs. This study investigated a variety of indium concentrations, dopant concentrations and profiles, spacer and buffer layer thicknesses and compositions, and device geometry variations (shorter length gates and/or multiple gates, sources, and drains). In summary, the device types successfully fabricated to date are of five different indium concentrations (53, 60, 65, 70, and 80 percent) with two different gate dimensions (0.1 and 0.07 μm) [113].

In order for the noise temperature of a cryogenic, 32-GHz HEMT LNA module to drop to 3 K, the HEMT maximum frequency of operation, f_{max}, should exceed 500 GHz. The f_{max} of the InP device used for the 32-GHz modules was 150 GHz. The best TRW InP devices that are currently being tested have f_{max}'s of 250 to 300 GHz. TRW device research is aimed at pushing f_{max} beyond 300 GHz.

Although the emphasis of this chapter is technical, it must be pointed out that the partnership among the Jet Propulsion Laboratory (JPL), the Georgia Institute of Technology (GIT), and TRW is really the corner stone of this work. This partnership was sustained for several years and provided the right technical mix to develop all of the key components required for the successful development of 8.4- and 32-GHz state-of-the-art, ultra-low-noise cryogenic, InP HEMT-based CCR receive systems for the DSN. In summary, the key contributions of each of the partners were cryogenic, LNA module design and characterization (JPL), cryogenic, on-wafer noise parameter measurements and HEMT modeling (GIT), and state-of-the-art cryogenic InP HEMTs (TRW). Organizations seeking to upgrade their HEMT state-of-the-art need to have this complete a suite of capabilities and a comparable long-term commitment to the development process.

References

[1] C. Kittel, *Introduction to Solid State Physics*, 2nd Ed., John Wiley and Sons, Hoboken, New Jersey, pp. 360–363, 1978.

[2] G. H. Dohler, "Solid-State Superlattices," *Scientific American*, vol. 249, pp.144–151, November 1983.

[3] S. M. Sze, *Physics of Semiconductor Devices*, 2nd Ed., John Wiley and Sons, Hoboken, New Jersey, pp. 122–129, 1985.

[4] J. C. M. Hwang, A. Kastalsky, H. L. Stormer, and V. G. Kermidas, "Transport Properties of Selectively Doped GaAs-(AlGa)As Heterostructures Grown by Molecular Beam Epitaxy," *Applied Physics Letters*, vol. 44, pp. 802–804, April 15, 1988.

[5] L. Esaki and R. Tsu, "Superlattice and Negative Differential Conductivity in Semiconductors," *IBM Journal of Research and Development*, vol.14, pp. 61–65, 1970.

[6] R. Dingle, H. L. Stormer, A. C. Gossard, and W. Wiegmann, "Electron Mobilities in M odulation-doped Semiconductor Heterojunction Superlattices," *Applied Physics Letters*, vol. 33, pp. 665–667, October 1978.

[7] H. L. Stormer, R. Dingle, A. C. Gossard, W. Wiegmann, and M. D. Sturge, "Two-Dimensional Electron Gas at a Semiconductor-Semiconductor Interface," *Solid State Communications*, vol. 29, pp. 705–709, 1979.

[8] N. T. Linh, M. Laviron, P. Delescluse, P. N. Tung, D. Delage-beaudeuf, F. Diamond, and J. Chevrier, "Low Noise Performance of Two-Dimensional Electron GaAs FETs," *Proceedings of IEEE/Cornell Conference on High-Speed Semiconductor Devices and Circuits*, pp. 187–193, August 15–17, 1983.

[9] K. Joshin, T. Mimura, Y. Yamashita, K. Kosemura, and J. Saito, "Noise performance of Microwave HEMT," *Proceedings of IEEE-MTT-S International Microwave Symposium*, pp. 563–565, June 1983.

[10] M. W. Pospieszalski, "Noise Parameters and Light Sensitivity of Low-Noise High-Electron-Mobility-Transistors," *IEEE Transactions on Electron Devices*, vol. ED-33, pp. 218–223, 1986.

[11] M. W. Pospieszalski and W. J. Llakatosh, "Millimeter-Wave, Cryogenically-Coolable Amplifiers Using AlInAs/GaInAs/InP HEMTs," *Proceedings of the 1993 IEEE MTT-S International Symposium*, Atlanta, Georgia, pp. 515–518, June 1993.

[12] S. Weinreb, R. Lai, N. Erickson, T. Gaier, and J. Wielgus, "W-Band InP Wideband MMIC LNA with 30 K Noise Temperature," *IEEE MTT-S Digest*, pp. 101–104, 1999.

[13] J. J. Bautista, J. G. Bowen, N. E. Fernandez, B. Fujiwara, J. Loreman, S. Petty, J. L. Prater, R. Grunbacher, R. Lai, M. Nishimoto, M. R. Murti, and J. Laskar, "Cryogenic, X-band and Ka-band InP HEMT based LNAs for the Deep Space Network," *Aerospace Conference, 2001, IEEE Proceedings*, vol. 2, pp. 829–842, 2001.

[14] A. Y. Cho and J. R. Arthur, *Progress in Solid State Chemistry*, Vol. 10, G. Somorjai and J. McCaldin, eds., Pergamon Press, New York, New York, p. 157, 1975.

[15] E. H. C. Parker, ed., *The Technology and Physics of Molecular Beam Epitaxy*, Plenum Press, New York, New York, 1985.

[16] L. Pavesi, "MBE Growth of GaAs," *Properties of Gallium Arsenide*, 2nd Ed., INSPEC, The Institute of Electrical Engineers, London, United Kingdom, and New York, New York, pp. 640–650, 1990.

[17] J. J. Coleman and P. D. Dapkus, "Metalorganic Chemical Vapor Deposition," *Gallium Arsenide Technology*, D. K. Ferry, ed., Howard W. Sams, Indianapolis, Indiana, p. 79, 1985.

[18] A. G. Thompson, "MOVPE Growth of GaAs," *Properties of Gallium Arsenide*, 2nd Ed., INSPEC, The Institute of Electrical Engineers, London, United Kingdom, and New York, New York, pp. 627–639, 1990.

[19] J. W. Mathews and A. E. Blakeslee, "Defects in Epitaxial Multi-Layers, I. Misfit Dislocations," *Journal of Crystal Growth,* vol. 27, pp. 118–123, 1974.

[20] Z. Dobrovolskis, K. Grigoras, and A. Krotkus, "Measurement of the Hot-Electron Conductivity in S emiconductors Using Ultrafast Electric Pulses," *Applied Physics, A: Solids and Surfaces*, vol. 48, pp. 245–249, 1989.

[21] B. R. Bennett, B. P. Tinkham, J. B. Boos, M. D. Lange, and R. Tsai, "Materials Growth for InAs High Electron Mobility Transistors and Circuits," *Journal of Vacuum Science and Technology B*, vol. 22, no. 2, pp. 688–694, March/April 2004.

[22] J. B. Hacker, J. Bergman, G. Nagy, G. Sullivan, C. Kadow, H.-K.Lin, A. C. Gossard, M. Rodwell, and B. Brar, "An Ultra-Low Power InAs/AsSb HEMT W-Band Low-Noise Amplifier," *Proceedings of IEEE-MTT-S International Microwave Symposium*, pp. 1–4, June 2005.

[23] T. Sugeta, A. Majerfeld, A. K. Saxera, P. N. Robson, and G. Hill, "High Field Transport Properties of $Ga_{1-x}Al_xAs$," *Proceedings of the IEEE/Cornell Conference on Microwave Devices*, Cornell University, Ithaca, New York, p. 45, 1977.

[24] H. Morkoc, W. F. Kopp, T. J. Drummond, S. L. Su, R. E. Thorne, and R. Fisher, "Submicron gate GaAs/$Al_{0.3}Ga_{0.7}As$ MESFETs with Extremely Sharp Interface (40 Angstroms)," *IEEE Transactions on Electron Devices*, vol. ED-29, no. 6, pp. 1013–1018, 1982.

[25] R. Fisher, T. Masselink, Y. L. Sun, T. J. Drummond, Y. C. Chang, M. V. Klein, and H. Morkoc, "Improvement of the Inverted GaAs/AlGaAs Heterointerface," *Journal of Vacuum Science and Technology*, vol. B-2, pp. 170–174, April 1984.

[26] K. H. G. Duh, W. F. Kopp, P. Ho, P-C. Chao, M-Y. Ko, P. M. Smith, J. M. Ballingall, J. J. Bautista, and G. G. Ortiz, "32-GHz Cryogenically Cooled HEMT Low-Noise Amplifiers," *IEEE Transactions on Electron Devices*, vol. 36, no. 8, pp. 1528–1535, August 1989

[27] H. Lee, *Growth of Optimization GaAs/(Al,Ga)As Modulation-Doped Heterostructures by MBE for FET Applications*, Ph.D. Thesis, Cornell University, Ithaca, New York, 1985.

[28] M. Hueschen, N. Moll, E. Gowen, and J. Miller, "Pulse Doped MODFETs," *IEEE IEDM Technical Digest*, pp. 348–351, San Francisco, California, 1984.

[29] T. W. Hickmott, P. M. Solomon, R. Fischer, and H. Morkoc, "Negative Charge, Barrier Heights, and the Conduction-Band Discontinuity in $Al_xGa_{1-x}As$ Capacitors," *Journal of Applied Phys*ics, vol. 57, pp. 2844–2853, 1985.

[30] L. Camnitz, "Design Principles and Performance of Modulation-Doped Field Effect Transistors for Low-Noise Microwave Amplification," Ph.D. Thesis, Cornell University, Ithaca, New York, 1986.

[31] P. M. Mooney, "Deep Donor Levels (DX Centers) in III-V Semiconductors," *Journal of Applied Physics,* vol. 67, no. 3, pp. R1–R26, 1 February 1990.

[32] H. L. Stormer, R. Dingle, A. C. Gossard, W. Wiegmann, and M. D. Sturge, "Two-Dimensional Electron Gas at a Semiconductor-Semiconductor Interface," *Solid State Communications*, vol. 29, pp. 705–709, March 1979.

[33] T. J. Drummond, R. J. Fischer, W. F. Kopp, H. Morkoc, K. Lee, and M. Shur, "Bias Dependence and Light Sensitivity of (Al,Ga)As/GaAs MODFETs at 77 K," *IEEE Transactions on Electron Devices*, vol. ED-30, pp. 1806–1811, December 1983.

[34] W. E. Hoke, P. S. Lyman, W. H. Labossier, J. C. Huang, M. Zaitlin, H. Hendriks, and G. Flynn, "Molecular-Beam Epitaxial Growth of Pulse-Doped Pseudomorphic GaAlAs/GaInAs Transistors with High Gain and Low Noise Properties," *Journal of Vacuum Science Technology*, vol. B 8, pp. 397–401, 1990.

[35] K. L. Tan, "94-GHz 0.1-μm T-gate Low-Noise Pseudomorphic InGaAs HEMTs," *IEEE Electron Device Letters*, vol. 11, pp. 585–587, December 1990.

[36] K. L. Tan, "High Performance W-Band Low-Noise Pseudomorphic InGaAs HEMT MMIC Amplifiers," *Electronics Letters*, vol. 27, pp. 1166–1167, 1990.

[37] T. Griem, M. Nathan, G. W. Wicks, J. Huang, P. M. Capani, and L. F. Eastman, "High Conductance and Low Persistent Photoconductivity in $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.52}As$ Modulation-Doped Structures with Pinchoff Capabilities," *Journal of Vacuum Science and Technology B,* vol. B2, pp. 655–656, March 1985.

[38] P. Ho, M. Y. Kao, P. C. Chao, K. H. G. Duh, J. M. Ballingall, S. T. Allen, A. J. Tessmer, and P. M. Smith, "Extremely High Gain 0.15 μm Gate-Length InAlAs/InGaAs/InP HEMTs," *Electronics Letters*, vol. 27, pp. 325–327, 1991.

[39] D. C. Streit, K. L. Tan, R. M. Dia, A. C. Han, P. H. Liu, H. C. Yen, P. D. Chow, "High Performance W-Band InAlAs-InGaAs-InP HEMTs," *Electronics Letters*, vol. 279, pp. 1149–1150, June 1991.

[40] U. K. Mishra, A. S. Brown, and F. E. Rosenbaum, "DC and RF Performance of 0.1 μm Gate Length $Al_{0.48}In_{0.52}As/Ga_{0.38}In_{0.62}As$ HEMT with Optimized Transport Parameters," *IEEE Technology Digest of International Electron Devices Meeting (IEDM*), San Francisco, California, pp. 180–183, December 1988.

[41] C. C. Eugster, T. P. E. Broekaert, J. A. del Alamo, and C. G. Fonstad, "An InAlAs/InAs MODFET," *IEEE Electron Device Letters*, vol. 12, pp. 707–709, December 1991.

[42] T. Akazaki, T. Enoki, K. Arai, Y. Umeda, and Y. Ishii, "High-Frequency Performance for Sub-0.1 μm Gate InAs-Inserted-Channel InA1As/InGaAs HEMT," *Electronics Letters*, vol. 28, pp. 1230–1231, June 1992.

[43] D. Yang, Y. C. Chen, T. Brock, and P. K. Bhattacharya, "DC and Microwave Performance of a 0 .1 μm Gate InAs/ $In_{0.52}Al_{0.48}As$ MODFET," *IEEE Electron Device Letters*, vol. 13, pp. 350–352, June 1992.

[44] K. Yoh, T. Moriuchi, and M. Inoue, "An InAs Channel Heterojunction Field Effect Transistor with High Transconductance," *IEEE Electron Device Letters*, vol. 11, pp. 526–528, November 1990.

[45] J. Kolodzey, J. Laskar, S. Boor, S. Agarwala, S. Caracci, A. A. Ketterson, I. Adesida, K. C. Hsieh, D. Sivco, and A. Y. Cho, "Direct-Current and Radio-Frequency Properties of InAlAs/InGaAs Pseudomorphic Modulation Doped Field Effect Transistors with Graded Channels," *Journal of Vacuum Science Technology B*, vol. B8, pp. 360–363, March 1990.

[46] M. Wojtowicz, R. Lai, D. C. Streit, G. I. Ng, T. R. Block, K. L. Tan, P. H. Liu, A. K. Freudenthal, and R. M. Dia, "0.10 μm Graded InGaAs Channel InP HEMT with 305 GH z fT and 340 GHz fmax," *IEEE Electron Device Letters*, vol. 15, pp. 477–479, November 1994.

[47] L. D. Nguyen, "650 A Self-Aligned Gate Pseudomorphic $Al_{0.48}In_{0.52}As/Ga_{0.20}In_{0.80}As$ High Electron Mobility Transistors," *IEEE Electron Device Letters*, vol. 13, pp. 143–145, March 1992.

[48] N. Rorsman, C. Karlsson, H. Zirath, S. M. Wang, and T. G. Andersson, "Characterization of InAlAs/InGaAs HFETs with High Indium Content in the Channel Grown on GaAs Substrate," *ESSDERC 2003, Proceedings of the 33rd European Solid-State Device Research Conference* (Estoril, Portugal, September 16–18, 2003), pp. 765–768, 2003.

[49] K. Inoue, J. C. Harmand, and T. Matsuno, "High-Quality InxGa1-xAs/InAlAs Modulation-Doped Heterostructures Grown Lattice-Mismatched on GaAs Substrates," *Journal of Crystal Growth*, vol. 111, pp. 313–317, 1991.

[50] P. Win, Y. Druelle, P. Legry, S. Lepilliet, A. Cappy, Y. Cordier, and J. Favre, "Microwave Performance of 0.4 μm Gate Metamorphic $In_{0.29}Al_{0.71}As/\ In_{0.30}Ga_{0.70}As$ HEMT on GaAs Substrate," *Electronics Letters*, vol. 29, pp. 169–170, 1993.

[51] K. Higuchi, M. Kudo, M. Mori, and T. Mishima, "First High Performance InAlAs/InGaAs HEMTs on GaAs Exceeding That on InP," *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 891–894, 1994.

[52] L. D. Nguyen, L. E. Larson, and U. K. Mishra, "Ultra-High-Speed Modulation-Doped Field-Effect Transistors: A Tutorial Review," *Proceedings of the IEEE*, vol. 80, no. 4, pp. 494–518, April 1992.

[53] M. Berg, T. Hackbarth, and J. Dickmann, "80–100 GHz Broadband Amplifier MMIC Utilizing CPWs and Quarter Micron InP-Based HEMTs," *Proceedings on Indium Phosphide and Related Materials (IPRM)* (Cape Cod, Massachusetts), TuA4, pp. 245–248, 1997.

[54] R. Lai, H. Wang, Y. C. Chen, T. Block, P. H. Liu, D. C. Streit, D. Tran, P. Siegel, M. Barsky, W. Jones, and T. Gaier, "D-Band MMIC LNAs with 12 dB Gain at 155 GHz Fabricated on a High Yield InP HEMT MMIC Production Process," *International Conference on Indium Phosphide and Related Materials (ICIPRM)* (Cape Cod, Massachusetts), TuA3, pp. 241–244, 1997.

[55] R. Lai, M. Barsky, R. Grundbacher, L. Tran, T. Block, T. P. Chin, V. Medvedev, E. Sabin, H. Rogers, P. H. Liu, Y. C. Chen, R. Tsai, and D. Streit, "0.1 μm InGaAs/InAlAs/InP HEMT Production Process for High Performance and High Volume MMW Applications," *Proceedings of the GaAs International Conference on Gallium Arsenide Manufacturing Technology (MANTECH*), (Vancouver, Canada), pp. 249–252, April 1999.

[56] R. Grundbacher, R. Lai, M. Barsky, R. Tsai, T. Gaier, S. Weinreb, D. Dawson, J. J. Bautista, J. F. Davis, N. Erickson, T. Block, and A. Oki, "0.1 μm InP HEMT Devices and MMICs for Cryogenic Low Noise Amplifiers from X-Band to W-Band," *Proceedings of the 14th Indium Phosphide and Related Materials Conference (ICPIRM)*, B6-5, pp. 455–458, 2002.

[57] K. van der Zanden, M. Behet, and G. Borghs, "Comparison of Metamorphic InGaAs/InAlAs HEMT's on GaAs with InP Based LM HEMT's," *Proceedings of the GaAs International Conference on Gallium Arsenide Manufacturing Technology (MANTECH)* (Vancouver, Canada), pp. 1–4, April 1999.

[58] F. J. Hohn, "Electron Beam Lithography: Its applications," *Journal of Vacuum Science Technology B,* vol. B7, pp. 1405–1411, November/ December 1989.

[59] R. C. Henderson, "Device Fabrication by Electron-Beam Lithography," *Electron-Beam Technology in Microelectronics Fabrication*, G. R. Brewer, ed., pp. 215–217, Academic Press, New York, New York, 1980.

[60] P. C. Chao, P. M. Smith, S. Wanuga, W. H. Perkins, and E. D. Wolf, "Channel-Length Effects in Quarter-Micrometer Gate-Length GaAs MESFETs," *IEEE Electron Device Letters*, vol. EDL-4, p. 326, September 1983.

[61] W. Patrick, W. S. Mackie, S. P. Beaumont, C. D. W. Wilkinson, and C. H. Oxley, "Very Short Gate-Length GaAs MESFETs," *IEEE Electron Device Letters*, vol. EDL-6, pp. 471–472, September1985.

[62] R. Williams, *Modern GaAs Processing Methods*, 2nd Ed., Artech House, Boston, Massachusetts, and London, United Kingdom, 1990.

[63] S. Kuroda, N. Harda, T. Katakami, T. Mimura, and M. Abe, "A New Fabrication Technology for AlGaAs/GaAs HEMT LSIs Using InGaAs Nonalloyed Ohmic Contacts," *IEEE Transactions on Electron Devices*, vol. ED-36, pp. 2196–2203, 1989.

[64] V. L. Rideout, "A Review of the Theory and Technology for Ohmic Contacts to Gr oup III–V Compound Semiconductors," *Solid-State Electronics*, vol. 18, pp. 541–550, 1975.

[65] L. P. Sadwick, "Device and Material Properties of Pseudomorphic HEMT Structures Subjected to R apid Thermal Annealing," *IEEE Transactions on Electron Devices*, vol. 39, pp. 50–55, January 1992.

[66] P. C. Chao, M. Shur, R. C. Tiberio, K. H. G. Duh, P. M. Smith, J. M. Ballingall, P. Ho, and A. A. Jabra, "DC and Microwave Characteristics of Sub-0.1 μm Gate-Length Planar-Doped Pseudomorphic HEMTs," *IEEE Transactions on Electron Devices*, vol. ED-36, p. 461, 1989.

[67] P. C. Chao, W. H. Ku, P. M. Smith, and W. H. Perkins, "0.2 μm Length T-Shaped Gate Fabrication Using Angle-Evaporation," *IEEE Electron Device Letters*, vol. EDL-4, pp. 122–124, 1983.

[68] M. Matsumura, K. Tsutsui, and Y. Naruke, "Submicrometer Lift-off Line with T-Shaped Cross-Sectional Form," *Electronics Letters*, vol. 12, pp. 429–430, 1981.

[69] S. G. Bandy, Y. G. Chai, R. Chow, C. K. Nishimoto, and G. Zdasiuk, "Submicrometer GaAs Microwave FETs with Low Parasitic Gate and Source Resistances," *IEEE Electron Device Letters*, vo1. EDL-4, pp. 42–44, 1983.

[70] P. C. Chao, P. M. Smith, S. C. Palmateer, and J. C. M. Hwang, "Electron Beam Fabrication of GaAs Low-Noise MESFETs Using a New Tri-Layer Resist Technique," *IEEE Transactions on Electron Devices*, vol. ED-32, pp. 1042–1046, June 1985.

[71] J. J. Lepore, "An Improved Technique for Selective Etching of GaAs and $Al_xGa_{l-x}As$," *Journal of Applied Physics*, vol. 51, pp. 6441–6442, December 1980.

[72] K. Hikosaka, T. Mimura, and K. Joshin, "Selective Dry Etching of AlGaAs/GaAs Heterojunction," *Japanese Journal of Applied Physics*, vol. 20, pp. L847–L850, November 1981.

[73] P. C. Chao, A. Tessmer, K. H. G. Duh, P. Ho, M. Kao, P. M. Smith, J. M. Ballingall, S. M. Liu, and A. A. Jabra, "W-Band Low-Noise InAlAs/InGaAs/InP Lattice-Matched HEMTs," *IEEE Electron Device Letters*, vol. EDL-II, pp. 59–62, 1990.

[74] C. E. Weitzel and D. A. Doane, "A Review of GaAs MESFET Gate Electrode Fabrication Technologies," *Journal of the Electrochemical Society*, vol. 133, pp. 409C–416C, October 1986.

[75] IRE Subcommittee 1.9 on Noise, "Noise Representation of Noise in Linear Two Ports," *Proceedings of the Institute of Radio Engineers*, vol. 48, pp. 60–68, January 1960.

[76] H. Rothe and W. Dahlke, "Theory of Noisy Fourpoles," *Proceedings of the IRE, Institute of Radio Engineers*, vol. 44, pp. 811–818, June 1956.

[77] P. Penfield, "Wave Representation of Amplifier Noise," *IRE Circuit Theory*, vol. CT-9, pp. 84–86, March 1962.

[78] R. Q. Lane, "The Determination of Device Noise Parameters," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1461–1462, August 1969.

[79] A. van der Ziel, "Thermal Noise in Field-Effect Transistors," *Proceedings of the Institute of Radio Engineers*, vol. 50, pp. 1808–1812, 1962.

[80] A. van der Ziel, "Gate Noise in Field-Effect Transistors at Moderately High Frequencies," *Proceedings of the Institute of Radio Engineers*, vol. 51, pp. 461–467, March 1963.

[81] A. Cappy, A. Vanoverschelde, M. Schortgen, C. Versnaeyen, and G. Salmer, "Noise Modeling Submicrometer-Gate, Two-Dimensional, Electron-Gas, Field Effect Transistors," *IEEE Transactions on Electron Devices*, vol. Ed-32, pp. 2787–22796, December 1985.

[82] K. Joshin, S. Asai, Y. Hirachi, and M. Abe, "Experimental and Theoretical Noise Analysis of Microwave HEMTs," *IEEE Transactions on Electron Devices*, vol. 36, no. 10, pp. 2274–2280, October 1989.

[83] H. Statz, H. Haus, and R. A. Pucel, "Noise Characteristics of Gallium Arsenide Field Effect Transistors," *IEEE Transactions on Electron Devices*, vol. ED-21, pp. 549–562, September 1974.

[84] R. A. Pucel, H. Haus, and H. Statz, "Signal and Noise Properties of GaAs FETs," *Advances in Electronics and Electron Physics*, vol. 38, pp. 549–561, September 1974.

[85] H. Fukui, "Design of Microwave GaAs MESFETs for Broadband, Low-Noise Amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-27, pp. 643–650, July 1979.

[86] M. Pospieszalski, "Modeling of Noise Parameters of MESFETs and MODFETs and Their Frequency and Temperature Dependence," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-37, no. 9, pp. 1340–1350, September 1989.

[87] J. Stenarson, Thesis: *Microwave Transistor Noise Model Extraction Methods and A Non-contacting Scattering Parameter Measurement Method*, Microwave Electronics Laboratory, Department of Microelectronics, Chalmers University of Technology, Gotenborg, Sweden, 2001.

[88] R. Reuter, S. van Waasen, and F. J. Tegude, "A New Noise Model of HFET with Special Emphasis on Gate-Leakage," *IEEE Electron Device Letters*, vol. 16, no. 2, pp.74–76, 1995.

[89] J. Mateos, T. González, D. Pardo, V. Hoël, and A. Cappy, "Monte Carlo Simulator for the Design Optimization of Low-Noise HEMTs," *IEEE Transactions on Electron Devices*, vol. 47, no. 10, pp. 1950–1956, October 2000.

[90] J. M. M. Pantoja, C.-I. Lin, M. Shaalan, J. L. Sebastian, and H. L. Hartnagel, "Monte Carlo Simulation of Microwave Noise Temperature in Cooled GaAs and InP," *IEEE Transactions on Microwave Theory and Techniques*, vol. 48, no. 7, pp. 1275–1279, July 2000.

[91] J. Mateos, T. González, D. Pardo, V. Hoël, H. Happy, and A. Cappy, "Improved Monte Carlo Algorithm for the Simulation of δ-Doped AlInAs/GaInAs HEMT's," *IEEE Transactions on Electron Devices*, vol. 47, no. 1, pp. 250–253, January 2000.

[92] *Cascade Microtech Model 42D User's Manual*, Cascade Microtech, Beaverton, Oregon, 1987.

[93] H. Meschde, R. Reuter, J. Albers, J. Kraus, D. Peters, W. Brockerhoff, F. Tegude, M. Bode, J. Schubert, and W. Zander, "On-Wafer Microwave Measurement Setup for Investigations on HEMT's and High Tc Superconductors at Cryogenic Temperatures Down to 20 K ," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, pp. 2325–2331, December 1992.

[94] J. Laskar and J. Kolodzey, "Vacuum Cryogenic High Frequency Probe Station," *Journal of Vacuum Science and Technology B*, vol. 8, pp. 1161–1165, September–October 1990.

[95] Y. Kwark, P. Solomon, and D. Tulipe, "S-Parameter Characterization of GaAs Gate SISFETs at Liquid Nitrogen Temperatures," *Proceedings of IEEE/Cornell University Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits*, Ithaca, New York, pp. 208–217, 1989.

[96] Kolodzey, H. Laskar, S. Boor, S. Reis, A. Ketterson, I. Adesida, D. Sivco, R. Fischer, and A. Y. Cho, "Cryogenic Temperature Performance of Modulation Doped Field Effect Transistors," *Electonics Letters*, vol. 25, pp. 777–778, 1988.

[97] J. J. Bautista, J. Laskar, and P. Szydlik, "On-Wafer, Cryogenic Characterization of Ultra-Low Noise HEMT Devices," *The Telecommunications and Data Acquisition Progress Report 42-120, October–December 1994*, Jet Propulsion Laboratory, Pasadena, California, pp. 104–120, February 15, 1995. http://ipnpr.jpl.nasa.gov/progress_report/

[98] J. Laskar, J. J. Bautista, M. Nishimoto, M. Hamai, and R. Lai, "Development of Accurate On-Wafer, Cryogenic Characterization Techniques," *IEEE Transactions on Microwave Theory and Technology*, vol. 44, pp. 1178–1183, July 1996.

[99] M. R. Murti, , J. Laskar, S. Nuttinck, S. Yoo, A. Raghavan, J. I. Bergman, J. J. Bautista, R. Lai, R. Grundbacher, M. Barsky, P. Chin, P. H. Liu, "Temperature Dependent Small-Signal and Noise Parameter Measurements and Modeling on I nP HEMTs," *IEEE Transactions on Microwave Theory and Techniques*, IEEE MTT-S International Symposium, Boston, Massachusetts, vol. 28, pp. 2579–2587, December 2000.

[100] G. Dambrine, A. Cappy, F. Heliodor, and E. Playez, "A New Method for Determining the FET Small-Signal Equivalent Circuit," *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, pp. 1151–1159, July 1988.

[101] C. A. Liechti, "Microwave Field-Effect Transistors - 1976," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-24, pp. 279–300, June 1976.

[102] M. Berroth and R. Bosch, "Broad-Band Determination of the FET Small-Signal Equivalent Circuit," *IEEE Transactions on Microwave Theory and Techniques*, vol. 38, pp. 891–895, 1990.

[103] N. Rorsman, M. Garcia, C. Karlson, and H. Zirath, "Accurate Small-Signal Modeling of HFETs for Millimeter-Wave Applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, pp. 432–437, March 1996.

[104] M. Garcia, N. Rorsman, K. Yhland, H. Zirath, and I. Angelov, "Fast, Automatic and Accurate HFET Small-Signal Characterization," *Microwave Journal*, vol. 40, pp. 102–117, July 1997.

[105] M. W. Pospieszalski and A. C. Niedwiecki, "FET Noise Model and On-Wafer Measurement of Noise Parameters," *IEEE MTT-S International Microwave Symposium Digest*, Boston, Massachusetts, Vol. 3, pp. 1117–1122, June 1991.

[106] M. Edwards and J. H. Sinsky, "New Criterion for Linear 2-Port Stability Using a Single Geometrically Derived Parameter," *IEEE Transactions on Microwave Theory and Techniques*, vol. 40, no. 12, pp. 2303–2311, December 1992.

[107] J. E. Fernandez, "A Noise-Temperature Measurement System Using a Cryogenic Attenuator," *The Telecommunications and Mission Operations Progress Report 42-135, July–September 1998*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–9, November 15, 1998. http://ipnpr.jpl.nasa.gov/progress_report/

[108] J. J. Bautista, "Cryogenic, Low-Noise High Electron Mobility Transistor Amplifiers for the Deep Space Network," *The Telecommunications and Data Acquisition Progress Report 42-115, July–September 1993*, Jet Propulsion Laboratory, Pasadena, California, pp. 65–80, November 15, 1993. http://ipnpr.jpl.nasa.gov/progress_report/

[109] S. Wedge and D. Rutledge, "Noise Waves and Passive Linear Multiports," *IEEE Microwave Guided Wave Letters*, vol. MGL-1, pp. 117–119, May 1991.

[110] C. T. Stelzreid, "Operating Noise Temperature Calibrations of Low Noise Receiving Systems," *Microwave Journal*, vol.14, no. 6, pp. 41–48, June 1971.

[111] R. C. Clauss, M. S. Reid, and H. F. Reilly, "Low Noise Receivers: Microwave MASER Development," *Space Program Summary 37–62, The Deep Space Network, January 1 to February 28, 1970*, Vol. II, Jet Propulsion Laboratory, Pasadena, California, pp. 74–78, March 31, 1970.

[112] C. T. Stelzreid, *The Deep Space Network–Noise Temperature Concepts, Measurements, and Performance*, JPL Publication 82-33, Jet Propulsion Laboratory, Pasadena, California, pp. 11.1–12.8, September 15, 1982.

[113] R. Lai, "Two Approaches to Improvement of InGaAs/InAlAs/InP HEMTs," *NASA Tech Briefs*, vol. 25, no. 2, Item NPO-20341, February 2000.

Chapter 6
Atmosphere Attenuation and Noise Temperature at Microwave Frequencies

Shervin Shambayati

6.1 Introduction

This chapter is concerned with the effects of atmospherics on attenuation and noise temperature. These effects are universal for low-noise ground stations. Other atmospheric effects (such as those caused by the wind, which causes both defocusing and mispointing of the antenna) are much more antenna specific and are not covered here.

In general, all models and activities attempt to simplify the effects of the atmosphere into that of a homogeneous attenuator with attenuation L and physical temperature T_p. For such an attenuator, the effective noise temperature at the output of the attenuator, T_E, relates to L and T_p according to the following equation:

$$T_E = \left(1 - \frac{1}{L}\right) T_p \tag{6.1-1}$$

where L is unitless and T_p and T_E are in kelvins.

The actual atmosphere is not homogenous. Nevertheless, Eq. (6.1-1) forms the basis for calculating the atmospheric loss and the equivalent atmospheric noise temperature.

The loss L for a homogenous attenuator of length l with absorption coefficient α is given by

$$L = \exp(\alpha l) \tag{6.1-2}$$

Now consider a column of air above a station. Assuming that at any height h the atmospheric absorption in nepers per kilometer for frequency f is given by $\alpha(h, f)$; then the atmospheric loss for that frequency at height h_0 is given by:

$$L_{\text{atm}}(h_0, f) = \exp\left(\int_{h_0}^{\infty} \alpha(h, f)\, dh \right) \tag{6.1-3}$$

If the physical temperature at height h is given by $T_p(h)$, then for a layer of infinitesimal thickness dh at height h, the equivalent noise temperature at the bottom of the layer is given by:

$$\begin{aligned} T_E(h, f) &= \left(1 - \exp\left[-\alpha(h, f)\, dh\right]\right) T_p(h) \\ &= \left(1 - 1 + \alpha(h, f)\, dh\right) T_p(h) \\ &= \alpha(h, f) T_p(h)\, dh \end{aligned} \tag{6.1-4}$$

The noise temperature contribution of a layer at height h with infinitesimal thickness dh to the equivalent atmospheric noise temperature at height $h_0 < h$, $T_{\text{atm}}(h_0, f)$, is given by:

$$T_{\text{atm},h}(h_0, f) = \exp\left(-\int_{h_0}^{h} \alpha(h', f)\, dh' \right) \alpha(h, f) T_p(h)\, dh \tag{6.1-5}$$

Integrating Eq. (6.1-5) over h, yields the equation for the equivalent atmospheric noise temperature at height h_0, $T_{\text{atm}}(h_0, f)$:

$$T_{\text{atm}}(h_0, f) = \int_{h_0}^{\infty} \exp\left(-\int_{h_0}^{h} \alpha(h', f)\, dh' \right) \alpha(h, f) T_p(h)\, dh \tag{6.1-6}$$

If the observations take place at an angle θ, Eqs. (6.1-3) and (6.1-6) take the forms of

$$L_{\text{atm}}^{(\theta)}(h_0, f) = \exp\left(\int_{h_0}^{\infty} \frac{\alpha(h, f)}{\sin\theta}\, dh \right) \tag{6.1-7}$$

and

$$T_{\text{atm}}^{(\theta)}(h_0, f) = \int_{h_0}^{\infty} \exp\left(-\int_{h_0}^{h} \frac{\alpha(h', f)}{\sin\theta} dh' \right) \frac{\alpha(h, f)}{\sin\theta} T_p(h)\, dh \qquad (6.1\text{-}8)$$

respectively, assuming a flat-Earth model. It should be noted that due to the relative thinness of the atmosphere compared to the radius of Earth, the flat-Earth model is a very good approximation.

Note that in Eqs. (6.1-2) through (6.1-8), the absorption $\alpha(h, f)$ is expressed in nepers per unit length. In practice, $\alpha(h, f)$ is usually expressed in terms of decibels per unit length. Using decibels for the absorption $\alpha(h, f)$, Eqs. (6.1-7) and (6.1-8) could be written as

$$L_{\text{atm}}^{(\theta)}(h_0, f) = 10^{\int_{h_0}^{\infty} (\alpha(h, f)/10\sin\theta) dh} \qquad (6.1\text{-}9)$$

and

$$T_{\text{atm}}^{(\theta)}(h_0, f) = \int_{h_0}^{\infty} 10^{-\int_{h_0}^{\infty} (\alpha(h', f)/10\sin\theta) dh'} \frac{\alpha(h, f)}{4.343\sin\theta} T_p(h)\, dh \qquad (6.1\text{-}10)$$

respectively. In order to maintain consistency with the literature, through the rest of this chapter, the unit of decibels per kilometer (dB/km) is used for $\alpha(h, f)$.

In order to calculate Eqs. (6.1-9) and (6.1-10), absorption and temperature profiles $\alpha(h, f)$ and $T_p(h)$ are needed. Lacking these, direct radiometer measurements of the sky brightness temperature could be used to calculate $L_{\text{atm}}(h_0, f)$ and $T_{\text{atm}}(h_0, f)$. At the National Aeronautics and Space Administration (NASA) Deep Space Network (DSN) both direct radiometer measurements and models for $\alpha(h, f)$ and $T_p(h)$ are used to estimate $L_{\text{atm}}(h_0, f)$ and $T_{\text{atm}}(h_0, f)$. In addition, meteorological forecasts are also used to calculate $\alpha(h, f)$ and $T_p(h)$, from which $L_{\text{atm}}(h_0, f)$ and $T_{\text{atm}}(h_0, f)$ could be calculated. Section 6.2 introduces the Jet Propulsion Laboratory (JPL) standard surface model. Section 6.3 discusses the processing of the data obtained from a water vapor radiometer (WVR) and an advanced water vapor radiometer (AWVR). Section 6.4 addresses weather forecasting, and Section 6.5 provides concluding remarks.

6.2 Surface Weather Model

The surface weather model is an attempt to use ground-level meteorological measurements to calculate $\alpha(h, f)$ and $T_p(h)$ and from them obtain $L_{\text{atm}}(h_0, f)$ and $T_{\text{atm}}(h_0, f)$ for a given frequency *f*. This topic has been extensively covered in texts and papers covering remote sensing (for example, see [1]); therefore, a detailed treatment of this topic is not considered useful at this time. Hence, in this section, we only address the surface model that is used by JPL and NASA's DSN for antenna calibrations. This JPL/DSN model has been developed by Dr. Stephen D. Slobin of JPL, and it uses variations on the weather models already available for calculation of $\alpha(h, f)$ and $T_p(h)$ for the first 30 km of atmosphere above the surface height from surface meteorological measurements.

6.2.1 Calculation of $T_p(h)$

The JPL/DSN surface weather model equation for $T_p(h)$ is a variation of that for the U.S. Standard Atmosphere, 1962, model (see [1]). For the first 2 km above the surface, the temperature is calculated through a linear interpolation of the surface temperature and the temperature given by the U.S. standard atmosphere for $T_p(h)$ at a height of 2 km above the surface. Furthermore, at heights above 20 km the temperature is assumed to be a constant 217 kelvin (K). To be exact, let *h* be the height above sea level in kilometers and h_0 be the height above the sea level for the station. Then if $T_{\text{US}}(h)$ is the temperature at height according to the U.S. standard atmosphere and T_{h_0} is the measured temperature at the surface, then $T_p(h)$ is given by

$$T_p(h) = \begin{cases} T_{h_0} + \dfrac{(h-h_0)}{2}\left(T_{\text{US}}(h_0+2) - T_{h_0}\right), & h_0 \le h \le h_0 + 2 \\ T_{\text{US}}(h), & h_0 + 2 < h \le 20 \\ 217, & 20 < h \le 30 \end{cases} \qquad (6.2\text{-}1)$$

$T_{\text{US}}(h)$ for $h < 32$ km is given by [1]

$$T_{\text{US}}(h) = \begin{cases} 288.16 - 6.5h, & 288.16 - 6.5h > 217 \\ 217, & 288.16 - 6.5h \le 217,\ h < 20 \\ 197 + h, & \text{otherwise} \end{cases} \qquad (6.2\text{-}2)$$

Figure 6-1 shows a comparison of the U.S. standard model and the JPL/DSN model for $h_0 = 1$ km and $T_{h_0} = 295$ K.

6.2.2 Calculation of $\alpha(h, f)$

Radio frequency (RF) absorption characteristics of a medium are related to the resonance frequencies of that medium and how close the RF frequency is to those frequencies. In a mixture such as the atmosphere, the resonance frequencies of all molecules that comprise the medium should be taken into account. For the atmosphere, the absorption is primarily due to four factors: oxygen, water vapor, clouds (liquid water), and rain. Therefore, for a given frequency f, atmospheric absorption for that frequency at height h, $\alpha(h, f)$ could be written as:

$$\alpha(h, f) = \alpha_{\text{ox}}(h, f) + \alpha_{\text{wv}}(h, f) + \alpha_{\text{cloud}}(h, f) + \alpha_{\text{rain}}(h, f) \quad (6.2\text{-}3)$$

The oxygen contribution $\alpha_{\text{ox}}(h, f)$ depends on the oxygen content of the atmosphere at height h, which is a function of pressure and temperature at that height. The pressure model is based on a weighted curve-fit of the pressure profile in the U.S. standard atmosphere. This weighted curve-fit attempts to model the pressure at lower heights more accurately as this would indirectly provide us with a more accurate measure of the oxygen content of the

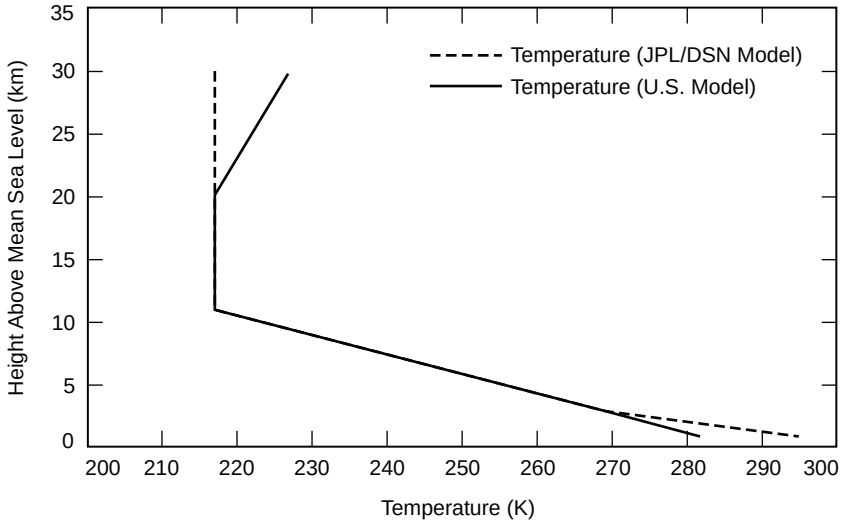


Fig. 6-1. JPL/DSN model and U.S. standard atmospheric temperature model for T_{h_0} = 295 K and h_0 = 1 km.

atmosphere. If P_{h_0} is the pressure measured at height h_0 km above the sea level, then the pressure profile for the atmosphere is calculated as:

$$P(h) = P_{h_0} \exp\left(\frac{8.387(h_0 - h)}{(8.387 - 0.0887h_0)(8.387 - 0.0887h)} \right) \qquad (6.2\text{-}4)$$

where h is height in kilometers above the mean sea level. Figure 6-2 shows a pressure profile of the atmosphere calculated from Eq. (6.2-4) for a P_{h_0} of 900 mbar at a height h_0 of 1 km.

The resonance frequencies of oxygen consist of a series of frequencies in the 50–70 GHz range (called the 60-GHz complex, see [1]) with an additional absorption line at 118.75 GHz. Due to interaction of oxygen molecules with each other and other molecules in the atmosphere, these absorption frequencies manifest themselves as two high-absorption bands around 60 GHz and 118.75 GHz. While there is a generalized model for oxygen absorption available over all frequencies, this model is very complicated. Therefore, due to the fact that as of this writing, NASA deep space missions are using only RF frequencies up to 32 GHz, the JPL/DSN surface weather model uses a simplified model that provides excellent agreement up to 45 GHz. This model is a slight modification of that presented in [1]. The absorption factor for oxygen as a function of temperature and pressure is given by

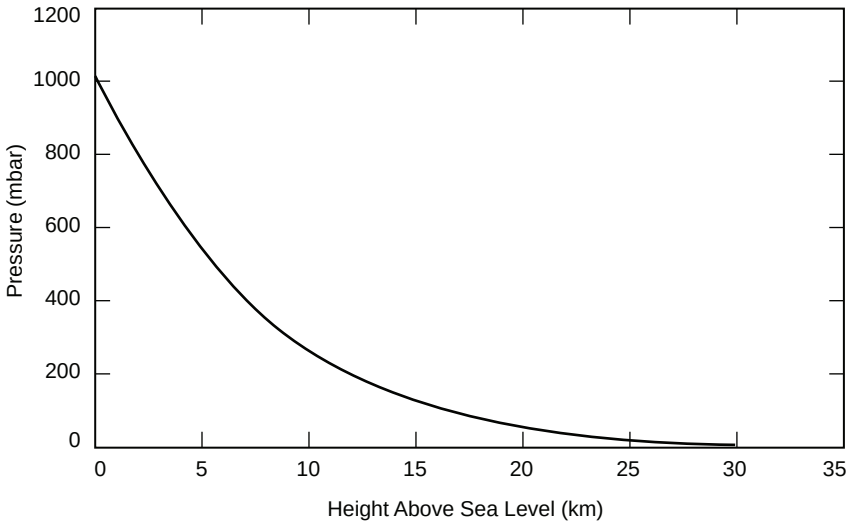


Fig. 6-2. JPL/DSN atmospheric pressure profile for P = 900 mbar (9.00 10^4 Pa) **at a height of 1 km above mean sea level.**

$$\alpha_{\text{ox}}(h,f) = C(f)\gamma_0(h) f^2 \left(\frac{P(h)}{1013}\right)^2 \left(\frac{300}{T_p(h)}\right)^{2.85} \cdot \left(\frac{1}{(f-60)^2+\gamma(h)^2} + \frac{1}{f^2+\gamma(h)^2}\right) \text{dB/km} \tag{6.2-5}$$

where f is frequency in gigahertz, h is the height in kilometers, $P(h)$ is the pressure in millibars at height h, and $T_p(h)$ is the temperature in kelvins at height h; $\gamma(h)$ is given by

$$\gamma(h) = \gamma_0(h)\left(\frac{P(h)}{1013}\right)\left(\frac{300}{T_p(h)}\right)^{0.85} \tag{6.2-6}$$

and $\gamma_0(h)$ is given by

$$\gamma_0(h) = \begin{cases} 0.59, & P(h) > 333 \\ 0.59\left(1 + 0.0031\left(333 - P(h)\right)\right), & 25 < P(h) \le 333 \\ 1.18, & P(h) \le 25 \end{cases} \tag{6.2-7}$$

Equation (6.2-5) differs from the same equation in [1] in that the term $C(f)$ in [1] is a constant and does not depend on the frequency. However, as illustrated in [1], this constant underestimates the oxygen absorption. The JPL/DSN model uses a frequency-dependent model for $C(f)$ based on a fourth-order curve fit. Using this model, $C(f)$ is given by:

$$C(f) = 0.011\left(7.13 \cdot 10^{-7} f^4 - 9.2051 \times 10^{-5} f^3 + 3.280422 \cdot 10^{-3} f^2 - 0.01906468 f + 1.110303146\right) \tag{6.2-8}$$

Using Eq. (6.2-5) very good agreement is obtained with the exact model for frequencies less than 45 GHz. (see Fig. 6-3).

The water vapor attenuation and in-band radiation is dominated by two absorption bands, one at 22.2 GHz and one at 180 GHz. Again, while there is a generalized model that covers water vapor absorption for all frequencies, the DSN is only using frequencies up to 32 GHz. Therefore, the JPL/DSN model uses a simplified model applicable up to 100 GHz in the surface weather model.

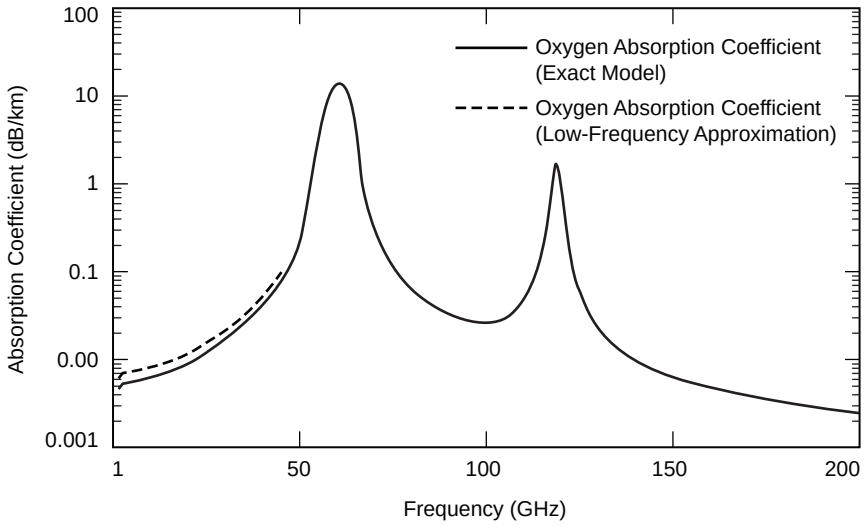


Fig. 6-3. Oxygen absorption coefficient versus frequency, T = 300 K, P = 1013 mbar (1.013 10^5 Pa)**.**

The water vapor contribution $\alpha_{\text{wv}}(h, f)$ is dependent on the absolute humidity at height *h*. Given the absolute humidity at surface $\rho_{\text{wv}}(h_0)$, the absolute humidity at height *h* is given by

$$\rho_{\text{wv}}(h) = \rho_{\text{wv}}(h_0)\exp\left(\frac{h}{H_{\text{wv}}}\right) \quad \frac{\text{g}}{\text{m}^3} \tag{6.2-9}$$

where H_{wv} is the scale height for water vapor, and it is set to 2 km. The absolute humidity at the surface, $\rho_{\text{wv}}(h_0)$, is calculated from the relative humidity at the surface r_{h_0} from the following formula:

$$\rho_{\text{wv}}(h_0) = \frac{1320.65}{T_{h_0}} r_{h_0} \bullet 10^{7.4475(T_{h_0} - 273.14)/T_{h_0} - 39.44} \quad \frac{\text{g}}{\text{m}^3} \tag{6.2-10}$$

Given $\rho(h)$, $\alpha_{\text{wv}}(h, f)$ is given by

$$\alpha_{\text{wv}}(h, f) = k_{\text{wv}}(h, f)\left(a_{\text{wv}}(h, f) + 1.2 \times 10^{-6}\right) \text{ dB/km} \tag{6.2-11}$$

where *f* is frequency in gigahertz and

$$k_{\mathrm{wv}}(h, f) = 2 f^2 \rho_{\mathrm{wv}}(h) \left(\frac{300}{T_p(h)} \right)^{1.5} \gamma_1(h) \tag{6.2-12}$$

and

$$a_{\mathrm{wv}}(h, f) = \left(\frac{300}{T_p(h) d_{\mathrm{wv}}(h, f)} \right) \exp\left(\frac{-644}{T_p(h)} \right) \tag{6.2-13}$$

and

$$d_{\mathrm{wv}}(h, f) = \left(22.2^2 - f^2 \right)^2 + 4 f^2 \gamma_1(h)^2 \tag{6.2-14}$$

and $\gamma_1(h)$ in gigahertz, the line width parameter of the water vapor, is given by

$$\gamma_1(h) = 2.85 \left(\frac{P(h)}{1013} \right) \left(\frac{300}{T_p(h)} \right)^{0.626} \left(1 + 0.018 \frac{\rho_{\mathrm{wv}}(h) T_p(h)}{P(h)} \right) \tag{6.2-15}$$

Figure 6-4 shows the $\alpha_{\mathrm{wv}}(h, f)$ for $P(h) = 1013$, $T_p(h) = 300$, and $\rho_{h_0} = 0.25$.

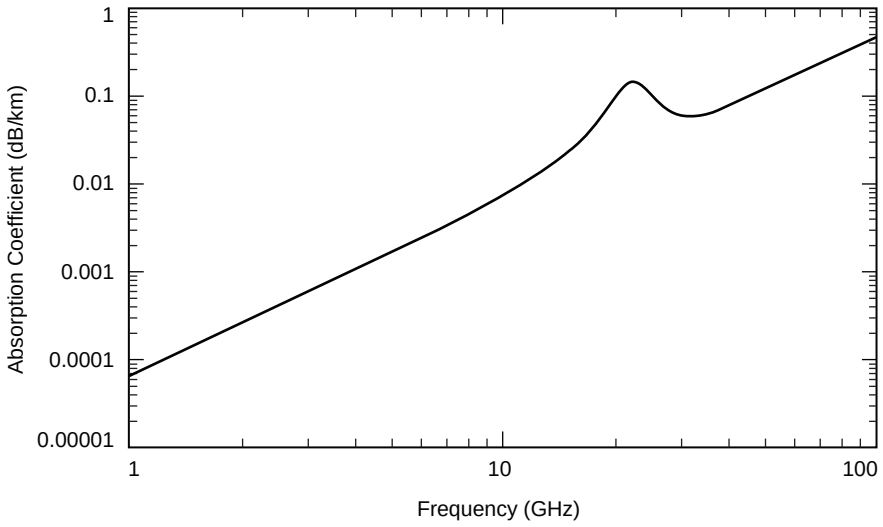


Fig. 6-4. Water vapor absorption coefficient versus frequency, T = 300 K, P = 1013 mbar (1.013 10^5 Pa)**, relative humidity = 25 percent.**

For calculating the contributions of the clouds to the atmospheric absorption, liquid water content profile of the atmosphere is needed. As this profile cannot be derived from the ground meteorological measurements, the liquid water content of the atmosphere could only be guessed at. This is the major failing of the surface weather model because, in the presence of the clouds, the surface weather model cannot give accurate modeling of the atmospheric effects. However, if the liquid water content (LWC) profile of the atmosphere $\rho_{\text{lwc}}(h)$ in grams per cubic meter (g/m^3) is known, then the cloud contribution in decibels per kilometer to $\alpha(h, f)$ is given by (see [1])

$$\alpha_{\text{cloud}}(h, f) = \rho_{\text{lwc}}(h) f^{1.95} \exp\left(1.5735 - 0.0309 T_p(h)\right) \qquad (6.2\text{-}16)$$

Figure 6-5 illustrates the $\alpha_{\text{cloud}}(h, f)$ for $\rho_{\text{lwc}}(h)$ of 0.1 g/m^3.

The rain models used in the JPL/DSN surface weather model are those developed by Olsen et al. [2]. In this model, the attenuation is calculated as a function of the rain rate. Given the rain rate in mm/h at height h, $r(h)$, the rain absorption, $\alpha_{\text{rain}}(h, f)$ is given by:

$$\alpha_{\text{rain}}(h, f) = a_{\text{rain}}(f) r(h)^{b_{\text{rain}}(f)} \qquad (6.2\text{-}17)$$

where

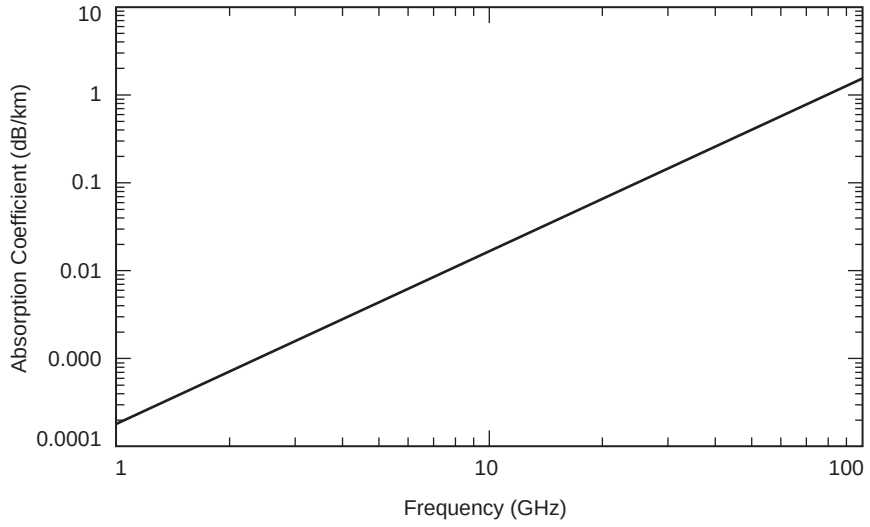


Fig. 6-5. Cloud absorption coefficient versus frequency, T = 275 K, LWC = 0.2 g/m^3**.**

$$a_{\text{rain}}(f) = \begin{cases} 6.39 \bullet 10^{-5} f^{2.03}, & f \le 2.9 \text{ GHz} \\ 4.21 \bullet 10^{-5} f^{2.42}, & 2.9 \text{ GHz} < f \le 54 \text{ GHz} \\ 4.9 \bullet 10^{-2} f^{0.699}, & 54 \text{ GHz} < f < 180 \text{ GHz} \end{cases} \quad (6.2\text{-}18)$$

and

$$b_{\text{rain}}(f) = \begin{cases} 0.851 f^{0.158}, & f \le 8.5 \text{ GHz} \\ 1.41 f^{-0.0779}, & 8.5 \text{ GHz} < f \le 25 \text{ GHz} \\ 2.65 f^{-0.272}, & 25 \text{ GHz} < f < 164 \text{ GHz} \end{cases} \quad (6.2\text{-}19)$$

Figure 6-6 shows $\alpha_{\text{rain}}(h, f)$ for r(h) = 2 mm/hr.

Again, note that in Eq. (6.2-17) the rain rate at height h, $r(h)$, needs to be known. Since the surface weather measurements cannot provide information about the rain rate at heights above the ground other than through guess work, the surface weather model provides no more than an educated guess about the effects of rain on the channel. Because of the inability of the surface weather model to deal with rain and clouds, direct radiometric measurements of the atmosphere are needed. For this purpose, JPL uses several water vapor radiometers (WVRs) [5] and advanced water vapor radiometers (AWVRs) [3,4]. In the next section, use of these instruments for characterization of atmospheric noise and atmospheric loss at different frequencies is discussed.

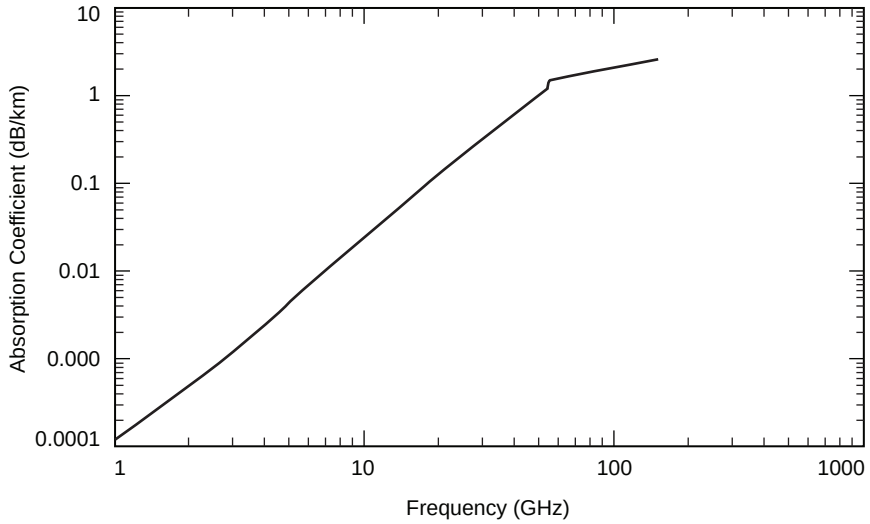


Fig. 6-6. Rain absorption coefficient versus frequency, rain rate = 2 mm/h**.**

6.3 Water Vapor Radiometer Data

6.3.1 Overview of Water Vapor Radiometer Operations and Data Processing Approach

JPL uses several types of WVRs to measure the sky brightness temperature at different elevations. The data thus recorded are processed so that the brightness temperature is referenced to zenith. Then these brightness temperature values are used to derive statistics for the zenith atmospheric noise temperature.

JPL uses several different WVRs: JT unit, D2 unit, R unit and two AWVR units. All these radiometers are Dicke radiometers. The JT unit measures the sky brightness temperatures at 20.7 GHz, 22.2 GHz, and 31.4 GHz; the D2 unit measures the sky brightness temperatures at 20.7 GHz and 31.4 GHz; and the R unit measures the sky brightness temperature at 23.8 GHz and 31.4 GHz [5]. AWVR measures the sky brightness at 22.2 GHz, 22.8 GHz, and 31.4 GHz [3]. The aggregate measurements are recorded every 3.5 minutes and checked against the radiometers health status to weed out any data that were affected by instrument malfunction. These instruments are located at the three DSN sites at Goldstone, California; Madrid, Spain; and Canberra, Australia. As of this writing (July 2005), we have 108 months of data for Goldstone, 147 months of data for Madrid, and 69 months of data for Canberra.

In standard radiometric measurements, the frequencies around 22 GHz are used to obtain the water vapor content of the atmosphere, while the 31.4-GHz frequency is used to obtain the liquid water content of the atmosphere. However, from a telecommunications point of view, we are interested in the total in-band radiation/absorption of the atmosphere. For this we use the 31.4-GHz measurements to obtain the zenith atmospheric noise temperature and then convert the zenith atmospheric noise temperature to that for the frequencies of interest (2.295 GHz [S-band], 8.42 GHz [X-band], and 32 GHz [Ka-band]).

6.3.2 Calculation of Atmospheric Noise Temperature from Sky Brightness Measurements at 31.4 GHz

Calculation of the atmospheric noise temperature, T_{atm}, from the sky brightness temperature measurements, T_B, by using the following relationship:

$$T_B = T_{\text{atm}} + \frac{T_{\text{cosmic}}}{L_{\text{atm}}} \tag{6.3-1}$$

where T_{cosmic} is the cosmic background noise temperature equal to 2.725 K, and L_{atm} is derived by substituting T_{atm} for T_E in Eq. (6.1-1), thus resulting in:

$$L_{\text{atm}} = \frac{T_p}{T_p - T_{\text{atm}}} \tag{6.3-2}$$

Substituting Eq. (6.3-2) in Eq. (6.3-1) and solving for T_{atm} results in

$$T_{\text{atm}} = T_p \left(\frac{T_B - T_{\text{cosmic}}}{T_p - T_{\text{cosmic}}} \right) \tag{6.3-3}$$

The problem with Eq. (6.3-3) is that the physical temperature of the atmosphere T_p is not known. However, practice has shown that a value of 275 K for T_p is a reasonable approximation; therefore, a physical temperature of 275 K is assumed for all calculations converting WVR and AWVR sky brightness temperature measurements to atmospheric noise temperature measurements. The AWVR and WVR data files provided for these calculations provide the sky brightness temperature measurements at zenith; therefore, the T_{atm} values that are calculated from them are also referenced to zenith.

The conversion formulas that are used to convert zenith atmospheric noise temperature values at 31.4 GHz to noise temperature values at deep space frequencies have been developed by Dr. Stephen D. Slobin of the Jet Propulsion Laboratory. The general formula for conversion of 31.4 GHz T_{atm} to 32 GHz T_{atm} is of the form

$$T_{\text{atm}}^{(32)} = T_{\text{atm}}^{(31.4)} + 5\left(1 - \exp\left(-0.008 T_{\text{atm}}^{(31.4)}\right)\right) \tag{6.3-4}$$

The conversion formulas for frequencies below 12 GHz are derived from the observation that for these frequencies the loss due to the water content of the atmosphere as well as for frequencies from 32 GHz to 45 GHz are approximately proportional to the square of the frequency (see Eqs. (6.2-11), (6.2-16), and (6.2-17)). Using this observation, we can then convert losses (in decibels) due to water vapor from 32 GHz to frequencies below 12 GHz by using a square of frequencies ratio:

$$L_{\text{H}_2\text{O}}^{(\text{dB})}(f) = L_{\text{H}_2\text{O}}^{(\text{dB})}(32)\left(\frac{f}{32}\right)^2, \qquad f < 12 \tag{6.3-5}$$

We then note that the total atmospheric loss in de cibels is the sum of the losses due to water and the losses due to oxygen:

$$L_{\text{atm}}^{(\text{dB})} = L_{\text{H}_2\text{O}}^{(\text{dB})} + L_{\text{O}_2}^{(\text{dB})} \tag{6.3-6}$$

Therefore, if we know the losses due to oxyge n at 32 GHz and at the frequencies of interest below 12 GHz, we can easily calculate the atmospheric losses for these frequencies from T_{atm} at 32 GHz using Eqs. (6.3-2), (6.3-5), and (6.3-6). Fortunately, the oxygen content of the atmosphere at a given site remains relatively constant and is not affected much by the weather; therefore, the surface weather model is used to calculate losses due to oxygen (see Section 6.2). These losses for deep-space S-band, X-band and Ka-band for different DSN sites are shown in Table 6-1.

Using the losses in T able 6-1 along with Eq. (6.3-2) can be used to calculate the 0-percent atmospheric noise temperature, T_{O_2} . This is the atmospheric noise temperature that would be observed if the atmosphere had no water content (vapor, liquid, or rain). These values are shown in Table 6-2.

Using values of T_{O_2} in Table 6-2, the equation for calculating the atmospheric loss for frequency $f < 12$ GHz , $L_{\text{atm}}(f)$, is given by

$$L_{\text{atm}}(f) = \left(\frac{275}{275 - T_{\text{O}_2}(f)}\right)\left(\frac{275 - T_{\text{O}_2}(32)}{275 - T_{\text{atm}}(32)}\right)^{(f/32)^2} \tag{6.3-7}$$

By using Eq. (6.3-7) in Eq. (6.1-1) we obtain the atmospheric noise temperature at frequency $f < 12$ GHz , $T_{\text{atm}}(f)$

$$T_{\text{atm}}(f) = 275\left(1 - \frac{1}{L_{\text{atm}}(f)}\right) \tag{6.3-8}$$

Table 6-1. Oxygen losses in decibels at different DSN sites for S-band (2.295 GHz), X-band (8.42 GHz), and Ka-band (32 GHz).

Site	Losses for Each Band (dB)		
	S-band	X-band	Ka-band
Goldstone	0.031	0.034	0.108
Madrid	0.032	0.036	0.114
Canberra	0.033	0.037	0.116

Table 6-2. T_{O_2} values (K) for deep space S-band, X-band, and Ka-band at different sites.

Site	0-Percent Atmospheric Noise Temperature for Each Band (K)		
	S-band	X-band	Ka-band
Goldstone	1.935	2.156	6.758
Madrid	2.038	2.273	7.122
Canberra	2.081	2.323	7.277

While Eqs. (6.3-7) and (6.3-8) are applicable to the DSN current frequencies of interest, in the future the DSN may be asked to support other frequencies. Among these are the near-Earth Ka-band (26.5 GHz), Ka-band frequencies for manned missions (37.25 GHz), and W-band (90 GHz). For these frequencies the simple frequency squared approach does not work, and different models are needed. Stephen Keihm of JPL has developed simple regression models for the sky brightness temperature. These regression models could easily be translated to regression models for the atmospheric noise temperature [6]. For W-band, the Goldstone regression formula is given by

$$T_{\text{atm}}^{(90)} = -10.81 + 4.225T_{\text{atm}}^{(31.4)} - 0.01842\left(T_{\text{atm}}^{(31.4)}\right)^2 \tag{6.3-9}$$

and for Madrid and Canberra the regression formula is given by

$$T_{\text{atm}}^{(90)} = -15.69 + 4.660T_{\text{atm}}^{(31.4)} - 0.02198\left(T_{\text{atm}}^{(31.4)}\right)^2 \tag{6.3-10}$$

For the 37.25 GHz frequency band, the regression formula for Goldstone is given by

$$T_{\text{atm}}^{(37.25)} = 1.1314 + 1.2386T_{\text{atm}}^{(31.4)} \tag{6.3-11}$$

and for Madrid and Canberra is given by

$$T_{\text{atm}}^{(37.25)} = 1.1885 + 1.241T_{\text{atm}}^{(31.4)} \tag{6.3-12}$$

Matters are slightly different for the 26.5-GHz frequency. While there is a straight linear interpolation from 31.4 GHz to 26.5 GHz, this interpolation is not very accurate due to the proximity of the 26.5 GHz frequency to the water vapor 22.2-GHz absorption band. Therefore, a formula using both the 20.7-GHz and 31.4-GHz T_{atm} values is derived. Unfortunately, not all of the radiometers measure the sky brightness in the 20.7-GHz band. Therefore,

equations both with and without 20.7-GHz T_{atm} values are presented. The equations for T_{atm} at 26.5 GHz that do not include the 20.7-GHz value of T_{atm} are

$$T_{atm}^{(26.5)} = 4.035 + 0.8147T_{atm}^{(31.4)} \tag{6.3-13}$$

for Goldstone and

$$T_{atm}^{(26.5)} = 3.4519 + 0.8597T_{atm}^{(31.4)} \tag{6.3-14}$$

for Madrid and Canberra.

The equations for T_{atm} that use the 20.7-GHz value are

$$T_{atm}^{(26.5)} = -0.11725 + 0.3847T_{atm}^{(20.7)} + 0.5727T_{atm}^{(31.4)} \tag{6.3-15}$$

for Goldstone and

$$T_{atm}^{(26.5)} = -0.09853 + 0.4121T_{atm}^{(20.7)} + 0.5521T_{atm}^{(31.4)} \tag{6.3-16}$$

for Madrid and Canberra.

6.3.3 DSN Atmospheric Noise Temperature Statistics Based On WVR Measurements

As mentioned before, the WVR data are used to generate zenith atmospheric noise temperature statistics for the DSN. These statistics are then used for link design. The cumulative statistics for S-band, X-band, and Ka-band are shown in Figs. 6-7 through 6-9 for Goldstone, Madrid, and Canberra, respectively. As seen from these figures, the zenith atmospheric noise temperature is much lower for S-band and X-band than for Ka-band. Furthermore, Ka-band frequencies have a much larger range of possible temperature values than either X-band or S-band. Also note that as the 26.5-GHz Ka-band absorption is dominated by the water vapor resonance line at 22.2 GHz, the lower percentile weather values for 26.5 GHz are actually higher than those for the 32-GHz and 31.4-GHz bands. However, for the higher percentile values, the 26.5-GHz Ka-band has lower T_z values.

Figure 6-10 shows the T_z variation observed from complex to complex for the 32-GHz Ka-band. As indicated from the T_z values, Goldstone has better weather than either Madrid or Canberra. Canberra has slightly worse weather

than Madrid. However, as the distributions shown are aggregate, this figure does not tell the whole story. At higher frequencies, seasonal variations also play a role.

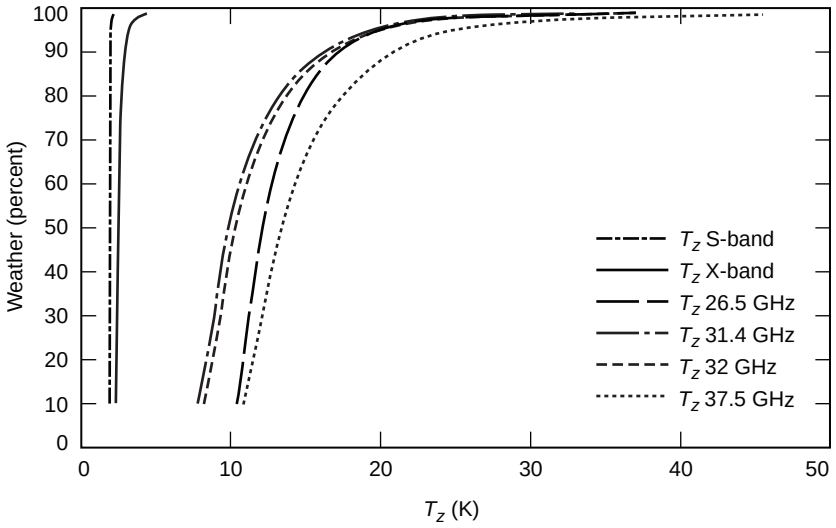


Fig. 6-7. Zenith atmospheric noise temperature (T_z) distributions for S-band, X-band, and Ka-band at the Goldstone, California, DSCC.

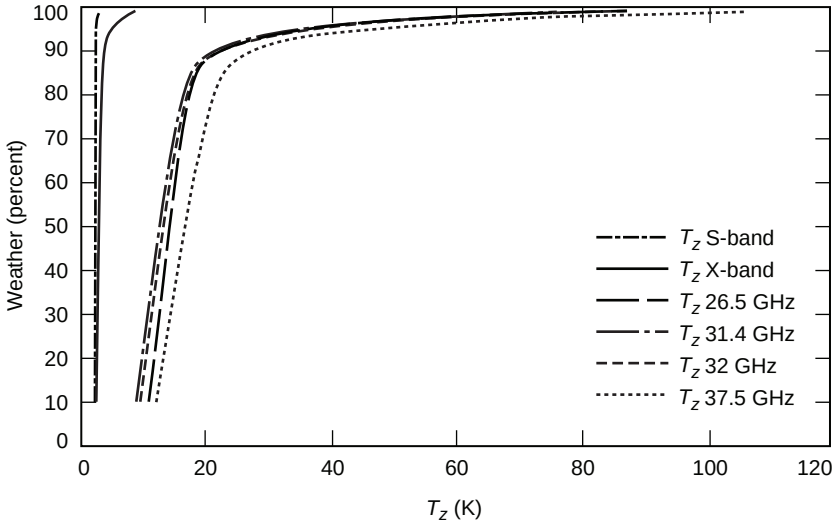


Fig. 6-8. Zenith atmospheric noise temperature distributions (T_z) for S-band, X-band, and Ka-band at the Madrid, Spain, DSCC.

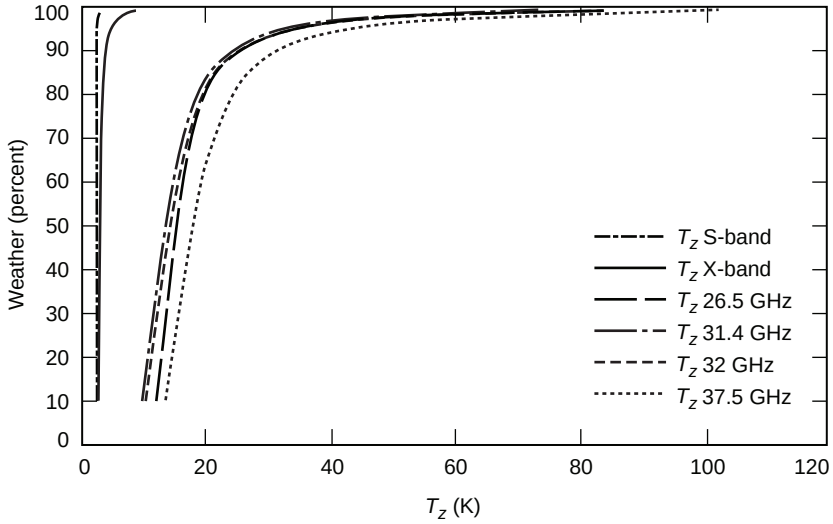


Fig. 6-9. Zenith atmospheric noise temperature distributions (T_z) for S-band, X-band, and Ka-band at the Canberra, Australia, DSCC.

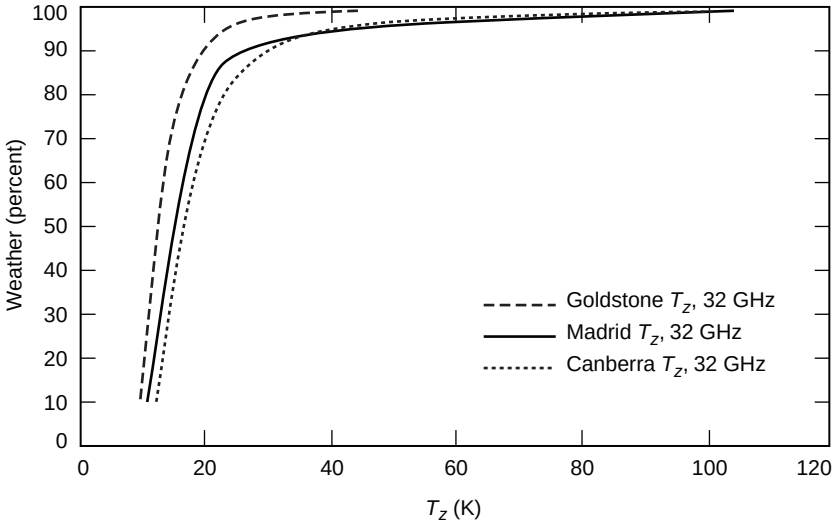


Fig. 6-10. Comparison of Goldstone, Madrid, and Canberra zenith atmospheric noise temperature distributions for the 32-GHz Ka-band.

Taking the monthly 90-percentile T_z value as an indicator of monthly T_z distributions, the weather effects could vary significantly from month to month for the higher frequencies. As seen in Fig. 6-11, Goldstone does not display much seasonal variation; however, both Madrid and Canberra have large

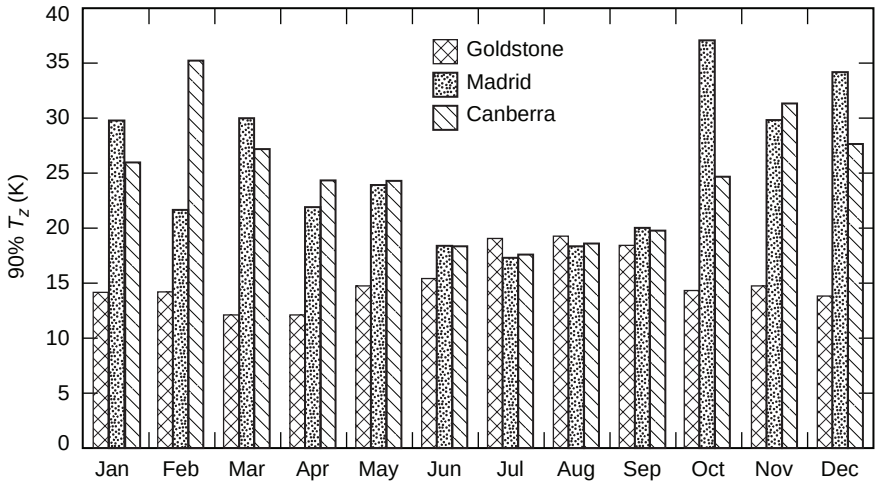


Fig. 6-11. 32-GHz Ka-band 90-percentile monthly noise temperature (T_z) values for Goldstone, Madrid, and Canberra.

weather variations during the year. For both these complexes the 90-percentile T_z for the best month is nearly half the 90-percentile T_z for the worst month.

6.4 Weather Forecasting

As seen in the previous section, weather effects at higher frequencies manifest themselves in large fluctuations in T_z. As T_z values are used in the link design, this means that for higher frequencies, the cost of reliability becomes higher in terms of margin that must be carried on the link in order to maintain a given reliability. Therefore, using just the long-term statistics means that either the link has to operated with a relatively low data rate most of the time in order achieve relatively high reliability or that the link can operate at higher data rates with lower reliability. However, if the weather effects could be predicted, then an algorithm which could adjust the data rate on the link according to weather conditions could be used in order to both maximize the link reliability and its data return capacity.

Before discussing the means of weather forecasting, one has to consider how a link is designed. The standard link equation is as follows:

$$R \bullet \left(\frac{E_b}{N_0} \right)_{(th)} = \frac{P_{SC} G_{SC}}{L_{\text{space}}} \bullet \frac{1}{L_{\text{atm}}} \bullet \frac{G_G}{kT_{\text{op}}} \qquad (6.4\text{-}1)$$

where R is the supportable data rate; $(E_b / N_0)_{(th)}$ is the required bit signal to noise ratio; P_{SC} is the spacecraft transmitted power; G_{SC} is the spacecraft antenna gain[1]; L_{space} is the space loss; L_{atm} is the atmospheric loss; G_G is the ground antenna gain; T_{op} is the system noise temperature and k is Boltzman's constant.

In Eq. (6.4-1), the term $P_{SC}G_{SC} / L_{space}$ is deterministic, depending only on the spacecraft telecommunications hardware and the distance between the spacecraft and Earth. Similarly, $(E_b / N_0)_{(th)}$ is determined by the type of channel coding that is available onboard the spacecraft, and G_G is a deterministic value. However, both L_{atm} and T_{op} are dependent on $T_z \bullet L_{atm}$, which is given by

$$L_{atm} = \frac{T_p}{T_p - T_{atm}} \tag{6.4-2}$$

where T_{atm} at elevation θ is given by

$$T_{atm} = \left(1 - \left(\frac{T_p - T_z}{T_p}\right)^{1/\sin\theta}\right) T_p \tag{6.4-3}$$

Let T_{mw} be the combined microwave noise temperature of the physical hardware that is used to track the spacecraft downlink signal. This includes the noise temperature of the antenna and the low noise amplifier (LNA). Then T_{op} is given by

$$\begin{aligned} T_{op} &= T_{mw} + T_B \\ &= T_{mw} + T_{atm} + \frac{T_{cosmic}}{L_{atm}} \end{aligned} \tag{6.4-4}$$

Given Eqs. (6.4-2) through (6.4-4), it is clear that the data rate selected in Eq. (6.4-1) is a random variable because T_z is a random variable and L_{atm} and T_{op} are functions of T_z. As in Eq. (6.4-1), a constant value for T_z must be assumed, and the selection of this value is based on the link design approach

[1] The product $P_{SC}G_{SC}$ is referred to as equivalent isotropic radiated power or EIRP.

that is taken and the distribution of T_z, $F_{T_z}(T) = \Pr\{T_z < T\}$. To put this mathematically, let $\theta(t)$ be the elevation profile of the pass for which the link is designed. Then the data rate profile for this pass, $R(t)$, is a function of $F_{T_z}(T)$ and $\theta(t)$:

$$R(t) = \gamma\left(\theta(t); F_{T_z}\right) \tag{6.4-5}$$

If T_z is relatively constant (as is the case for X-band and S-band), then the dependence of $R(t)$ on $F_{T_z}(T)$ is relatively minor. However, if like Ka-band and W-band, T_z can take a wide range of values, then dependence of $R(t)$ on $F_{T_z}(T)$ is significant as L_{atm} and T_{op} could significantly vary over time as a function of the weather, and care must be taken in selecting the proper $R(t)$ in order to take into account the uncertainty caused by variation of T_z.

In order to make the performance of the link more predictable, weather forecasting could be used to reduce the range of values that T_z could take. Let $w(t)$ be the predicted weather according to a weather forecasting algorithm. Then we can define a conditional distribution for T_z based on the weather forecast, $w(t)$:

$$F_{T_z|w(t)}(T) = \Pr\{T_z < T \mid w(t)\} \tag{6.4-6}$$

Given $F_{T_z|w(t)}(T)$, Eq. (6.4-5) is rewritten as

$$R(t) = \gamma\left(\theta(t); F_{T_z|w(t)}\right) \tag{6.4-7}$$

Note that $w(t)$ could take many forms. It could be something as simple as the date and time of the pass or as complicated as a detailed multi-layer meteorological description of the atmosphere provided by sophisticated mesonet models.

Currently, DSN is exploring the use of forecasts generated by the Spaceflight Meteorology Group at Johnson Space Flight Center for Ka-band link design. These forecasts, originally intended for use by NASA's Space Shuttle program, give a detailed multi-layer meteorological description of the atmosphere including details such as pressure, temperature, dew point, absolute

humidity, and liquid water content every 6 hours from 12 to 120 hours into the future. Therefore, each forecast set includes 19 different forecast types. (A forecast 30 hours into the future is of a different type than a forecast 36 hours into the future). Each forecast is valid for a single point in time; however, for our purposes, they could be taken as representative of the 6-hour period centered around them. The values of each forecast type were categorized, and for each category, a $F_{T_z|w(t)}(T)$ was obtained from AWVR sky brightness temperature measurements.

These forecasts were used as part of a study [7] where the values of T_z for a Ka-band link were selected according to a particular link design approach from these distributions and compared to T_z values obtained from the sky brightness temperature measurements made by the AWVR as well as with T_z values derived from monthly statistics in a "blind" test. The results are shown in Fig. 6-12. As seen from this figure, the T_z values derived from the weather forecasts follow very closely the T_z values derived from the AWVR measurements. This indicates that these forecasts could be used for adaptive link design. For a more complete treatment of this topic see [7].

6.5 Concluding Remarks/Future Directions

6.5.1 Current State

Currently most space missions use primarily X-band for their science data return; therefore, very little thought has been given to the effects of the weather on the telecommunications link performance. All but the severest weather events have very little effect on the performance of the X-band link. However, as future space missions start to use Ka-band, understanding weather effects on the link performance will become a priority.

6.5.2 Ka-Band Near-Term Development

As of this writing, several NASA and European Space Agency (ESA) spacecraft are slated to have telemetry downlink capability at Ka-band. In addition, NASA is augmenting its ground receiving capability to process both 32-GHz and 26.5-GHz Ka-band. Furthermore, ESA is building a series of 35-m beam waveguide (BWG) stations that will be Ka-band capable.

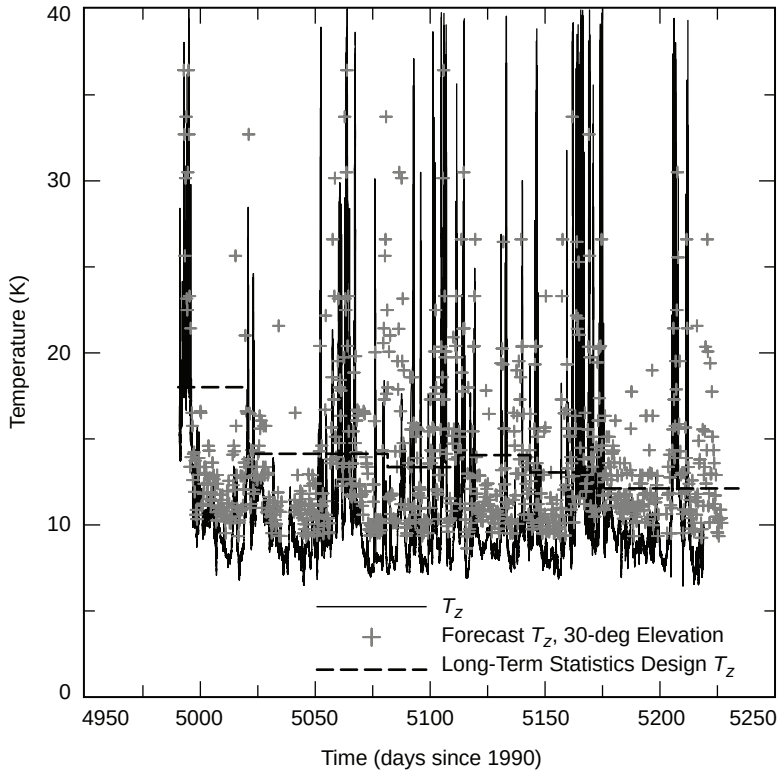


Fig. 6-12. Forecast T_z values, monthly T_z and AWVR-derived T_z measurements for Goldstone, September 2003 through April 2004.

NASA's Mars Reconnaissance Orbiter (MRO) has a fully functioning 32-GHz Ka-band downlink capability that is used for demonstration purposes. If this capability is proven reliable during the course of the demonstration, MRO will be using Ka-band to augment its science return. NASA's Kepler spacecraft and Space Interferometry Mission (SIM) spacecraft will use 32-GHz Ka-band for their primary science downlink. Lunar Reconnaissance Orbiter (LRO) and James Webb Space Telescope (JWST) will use 26.5-GHz Ka-band for their high-rate science return.

NASA has been implementing 32-GHz Ka-band support at its Deep Space Network (DSN). At this time (July 2006) the DSN has four 34-m BWG antennas capable of tracking Ka-band downlink. These are DSS 25 and DSS 26 at Goldstone Deep Space Communication Complex (DSCC) at Goldstone, California; DSS 34 at Canberra DSCC near Canberra, Australia; and DSS 55 at Madrid DSCC near Madrid, Spain. In the near future (2007–2008) plans are in place to upgrade one additional BWG antenna at Goldstone and another at Madrid to support Ka-band.

NASA is also considering implementing 37-GHz Ka-band at the DSN antennas for support of future manned missions to Mars. However, as of this writing, there are no plans in place to implement this capability.

6.5.3 Arraying

Currently, the DSN is considering replacing its monolithic 34-m and 70-m antennas with large arrays of 12-m parabolic antennas. At this time it is not clear whether or not these arrays will be located at the current DSCCs or that new complexes specifically for these antennas will be built. These antennas will be co-located and therefore, will observe the same atmospheric affects. Mathematically, let G_i be the gain of the ith antenna in the array and $T_{\text{op}} = T_{\text{mw}}^{(i)} + T_B$ be the system noise temperature of that antenna where $T_{\text{mw}}^{(i)}$ is the microwave noise temperature of ground equipment and T_B is sky brightness temperature observed in the line of site for the antenna. Then, assuming no combining loss, for an array of l antennas the signal at the receiver is

$$r = \sqrt{s}\sum_{i=1}^{l}\alpha_i\sqrt{G_i} + \sum_{i=1}^{l}\alpha_i n_i \tag{6.5-1}$$

where s is flux density of the signal observed at the antenna site, α_i is the optimum combining weight for the ith antenna (see [10]), and n_i is the noise at the ith antenna with a one-sided spectral density of $N_0^{(i)}$. The received signal power at the output of the combiner is $s\left(\sum_{i=1}^{l}\alpha_i\sqrt{G_i}\right)^2$. Assuming that the noise processes for different antennas in the array are independent of each other, then the signal to noise ratio at the output of the combiner is given by:

$$\frac{P}{N_0^{(\text{combined})}} = \frac{s\left(\sum_{i=1}^{l}\alpha_i\sqrt{G_i}\right)^2}{\sum_{i=1}^{l}\alpha_i^2 N_0^{(i)}} \tag{6.5-2}$$

$N_0^{(i)}$ is defined as

$$N_0^{(i)} = \mathrm{k}\left(T_{\mathrm{mw}}^{(\mathrm{i})} + T_B\right) \tag{6.5-3}$$

where k is the Boltzman constant.

Substituting Eq. (6.5-3) in Eq. (6.5-2) we obtain

$$\frac{P}{N_0^{(\mathrm{combined})}} = \frac{s\left(\sum_{i=1}^{l} \alpha_i \sqrt{G_i}\right)^2}{\sum_{i=1}^{l} \alpha_i^2 k\left(T_{\mathrm{mw}}^{(i)} + T_B\right)} \tag{6.5-4}$$

If the system noise temperature is defined by

$$T_{\mathrm{op}}^{(\mathrm{combined})} \triangleq \frac{N_0^{(\mathrm{combined})}}{\mathrm{k}} \tag{6.5-5}$$

then

$$T_{\mathrm{op}}^{(\mathrm{combined})} = \sum_{i=1}^{l} \alpha_i^2 \left(T_{\mathrm{mw}}^{(i)} + T_B\right) \tag{6.5-6}$$

Note that, as Eq. (6.5-2) indicates, the optimum combining weights, α_i, are not unique and if an l-tuple, $(\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_l)$, is a set of optimum weights, then so is $c(\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_l)$, where c is a positive constant. However, proper calculation of $T_{\mathrm{op}}^{(\mathrm{combined})}$ must be of the form

$$T_{\mathrm{op}}^{(\mathrm{combined})} = T_{\mathrm{mw}}^{(\mathrm{combined})} + T_B \tag{6.5-7}$$

Therefore, Eqs. (6.5-6) and (6.5-7) imply that

$$\sum_{i=1}^{l} \alpha_i^2 = 1 \tag{6.5-8}$$

and

$$T_{\mathrm{mw}}^{(\mathrm{combined})} = \sum_{i=1}^{l} \alpha_i^2 T_{\mathrm{mw}}^{(i)} \tag{6.5-9}$$

6.5.4 Optical

NASA is considering using optical frequencies for transmission of high-rate science data from its deep space probes in the future. A detailed description of the weather effects on the optical channel is beyond the scope of this document since these effects are substantially different than those on the RF link. These differences arise from the fact that optical channels operate at the quantum level; therefore, standard analog equations used for the RF channel do not apply. For a better treatment of this topic the reader is referred to [8,9].

6.5.5 Space-Based Repeaters

Since the weather effects become more severe at higher RF frequencies, one option that has been seriously considered is that of space-based repeaters for deep space missions. However, technological challenges and cost issues presented by such repeaters usually result in preference for ground-based antennas for tracking of deep-space missions.

Most of the technological challenges arise from the fact that the capacity of a multi-hop link (such as the one formed through the use of a repeater) is limited to that of its minimum-capacity hop. Since most space-based repeaters under consideration are Earth-orbiting, the minimum capacity hop is usually the probe-to-repeater link. Therefore, in order for a repeater-based link to compete with an Earth-based system, the capacity of the probe-to-repeater hop must be at least equal the capacity of the direct probe-to-Earth link.

For any RF receiving system, the capacity of the link is proportional to the gain-to temperature ratio (G/T). At Ka-band, for example, the G/T of a 34-m BWG for 90-percent weather at 30-deg elevation is around 60 dB with an antenna gain of about 78 dB and a system noise temperature of about 70 K. For a space-based repeater, the system noise temperature (SNT) is usually around 300 K to 450 K due to lack of cryogenically cooled LNAs. Therefore, the repeater already has a 6-dB disadvantage over the ground-based system because of the SNT. To compensate for this 6 dB, either innovative reliable space-based cryogenic technologies must be developed to reduce the receiver noise temperature on the repeater, or the gain of the antenna on the repeater must be increased.

Assuming that space-based antennas could be made as efficient as the 34-m BWG antenna regardless of size, a 68-m antenna is required to provide the 84-dB gain needed on the repeater. Needless to say, station-keeping and pointing for such a large antenna in Earth orbit requires sophisticated and advanced technologies, which are quite costly.

In addition to all these challenges, there also is the question of maintenance of such a space-based repeater network. Because these repeaters are in space, if these repeaters fail, their repair would be extremely difficult and expensive. Therefore, unless they are made reliable, space-based repeaters are of limited value.

References

[1] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive, Vol. I: Microwave Remote Sensing Fundamentals and Radiometry,* Chapter 5, Artech House, Norwood Massachusetts, pp. 256–343, 1981.

[2] R. L. Olsen, D. V. Rogers, and D. B. Hodge, "The aR^b Relation in the Calculation of Rain Attenuation," *IEEE Transactions on Antennas and Propagation,* vol. AP-26, no. 2, pp. 318–329, March 1978.

[3] A. Tanner and A. Riley, "Design and Performance of a High Stability Water Vapor Radiometer," *Radio Science*, vol. 38, no. 3, pp.15.1–15.12, March 2003.

[4] A. Tanner, "Development of a High Stability Water Vapor Radiometer," *Radio Science*, vol. 33, no. 2, pp. 449–462, March–April 1998.

[5] S. J. Keihm, "Final Report, Water Vapor radiometer Intercomparison Experiment: Platteville, Colorado, March 1–13, 1991," JPL D-8898 (JPL internal document), Jet Propulsion Laboratory, Pasadena, California, July 1991.

[6] S. Shambayati, "On the Use of W-Band for Deep-Space Communications," *The Interplanetary Network Progress Report 42-154, April–June 2003*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–28, August 15, 2003.
http://ipnpr.jpl.nasa.gov/progress_report/

[7] S. Shambayati, "Weather Related Continuity and Completeness on Deep Space Ka-band Links: Statistics and Forecasting," *IEEE Aerospace Conference,* Big Sky Montana, March 5–10, 2006.

[8] R. M. Gagliardi and S. Karp, *Optical Communications*, 2nd ed., John Wiley and Sons, New York, New York, 1995.

[9] H. Hemmati, ed., *Deep Space Optical Communications*, John Wiley and Sons, New York, New York, April 2006.

[10] D. H. Rogstad, A. Mileant, and T. T. Pham, *Antenna Arraying Technique in the Deep Space Network*, John Wiley and Sons, New York, New York, January 2003.

Chapter 7
Antenna Calibration

David J. Rochblatt

7.1 Introduction

The United States National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL) Deep Space Network (DSN) of large, dual reflector, Earth-based antennas is subject to continuing demands for improved performance, performance evaluation, and reliability as a result of escalating requirements for communications, control, and radio science requirements in supporting future missions.

The DSN provides the communications links with many spacecraft of the nation's unmanned space exploration program. In order to satisfactorily perform this mission, each antenna must undergo various calibrations to insure that it is operating as efficiently as possible, and hence delivering maximum information at minimum cost.

As part of the strategy of improving the overall performance capability of the DSN, there has been a steady increase in the operating frequency of these antennas over the years, going from S-band (frequency ~2.3 gigahertz (GHz)), to X-band (frequency ~8.4 GHz), and most recently to Ka-band (frequency ~32 GHz).

One can gain a better appreciation of the implications of these frequency increases for antenna calibration by considering the corresponding wavelengths (λ) of the radiation. Thus, for S-band, λ ~13 centimeters (cm), for X-band λ ~3.6 cm, and for Ka-band λ ~0.9 cm. The essential performance characteristics of an antenna, such as pointing capability and aperture efficiency, are strongly dependent on the wavelength of the radiation being detected. That is why the large (14:1) decrease in wavelength has resulted in the need for much greater

precision of such parameters as reflector surface figure, azimuth track smoothness, and subreflector and beam waveguide (BWG) mirror alignments.

This work describes the development of antenna-calibration instrumentation. The purpose of the antenna calibration instrumentation is to provide reliable tools f or the assessment, calibration, and performance improvement of the large number of antenna systems in the DSN. Utilizing the principles of noise temperature measurements, the instrumentation measures and derives the antenna gain, systematic pointing corrections, subreflector focus, as well as the calibration of radio stars[1] used in the assessment of antenna gain. The alignments of the antenna main reflector panels and antenna stability are best measured by coherent holographic techniques (described in Chapter 8 of this book).

As a consequence, a new generation of 34-meter (m) BWG antennas retrofitted with X-band (8.42 GHz) and Ka-band (32 GHz) transmitting and receiving systems is being added to the existing complement of 34-m and 70-m Cassegrain-like antennas (shaped reflectors).

As a result of a systematic analysis of the entire measurement procedure, with particular attention to the noise characteristics of the total-power radiometer (TPR) (Figs. 7-1, 7-2, and 7-3), plus the tropospherically induced radiometer fluctuations and the implementation of new techniques for data acquisition and reduction, it has been possible to obtain measurement precision yielding as much as an order of magnitude improvement over previous methods in the determination of antenna aperture efficiency, and factors of five or more in the determination of pointing errors and antenna beamwidth. This improvement has been achieved by performing continuous, rapid raster scans of both extended and point radio sources. Use of such scans is termed: on-the-fly (OTF) mapping.

The advantages of OTF mapping over the traditional or boresight approach to antenna calibration are that it

1) Removes a major source of error in de termining antenna efficiency by eliminating the need for independently derived-source size-correction factors.
2) Takes into account the actual spectrum of the TPR noise fluctuations in determining the optimum integration time during continuous-scan measurements. This includes radiometer flicker noise and tropospheric turbulence effects.
3) Provides a direct comparison of the flux density of one radio source with that of another, thus enabling the accurate calibration of many radio sources for future antenna calibrations.

---

[1] The term “radio star” refers to any natural, compact source of radiation, which, for example, may actually be a remote galaxy of stars, a region of bright microwave emission in our own Milky Way galaxy, or a planet in our Solar System.

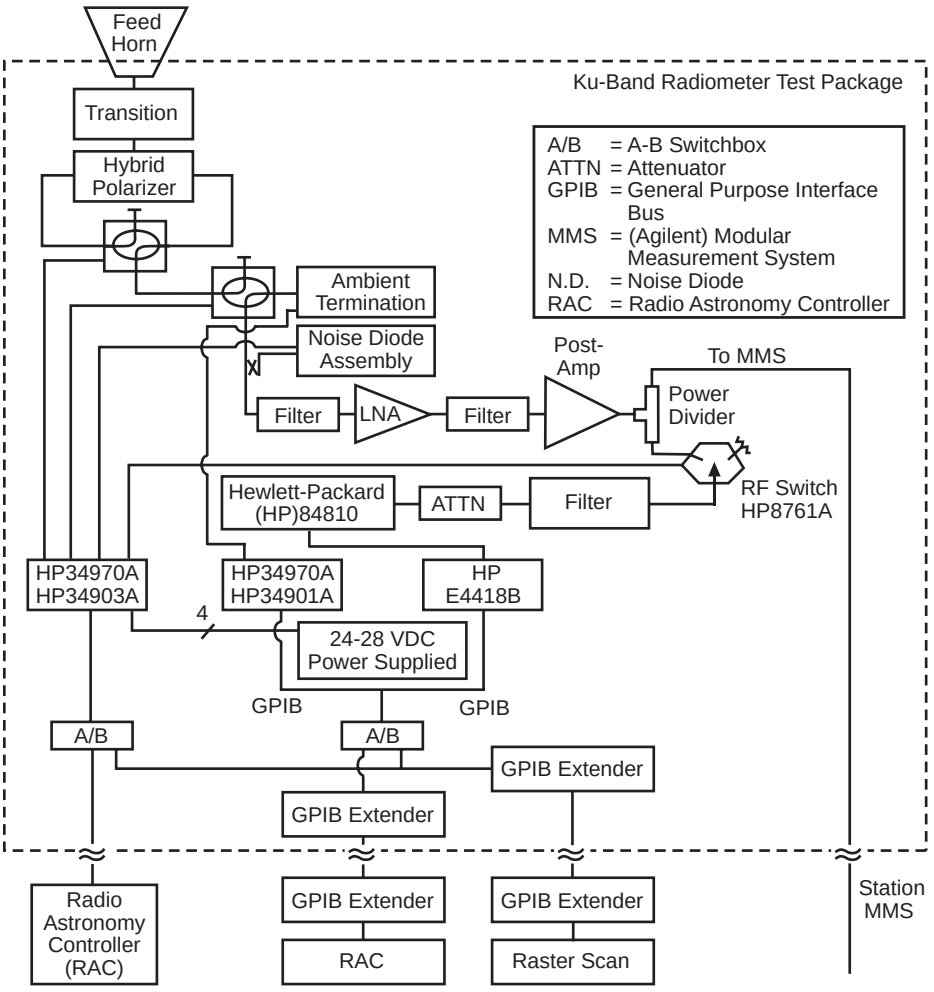


Fig. 7-1. Block diagram of a Ku-band TPR designed for operation with Deep Space Station 13 (DSS-13), 34 m-diameter BWG antenna at 13.8 GHz.

4) Reduces cost and improves the reliability of antenna performance through use of a highly accurate, repeatable, and fully automated system.

While other calibration techniques have advantages for certain applications, OTF mapping accurately provides all of the required calibration data, in the shortest measurement period, including the beam shape as well as the source flux density. The significance of improvements in antenna calibration, performance, or performance evaluation in the DSN can be put in perspective by recognizing that each decibel (dB) of improvement in the quantity gain over noise temperature (G/T) is estimated to be worth about U.S. $160M/dB in terms of mission support capability [1].

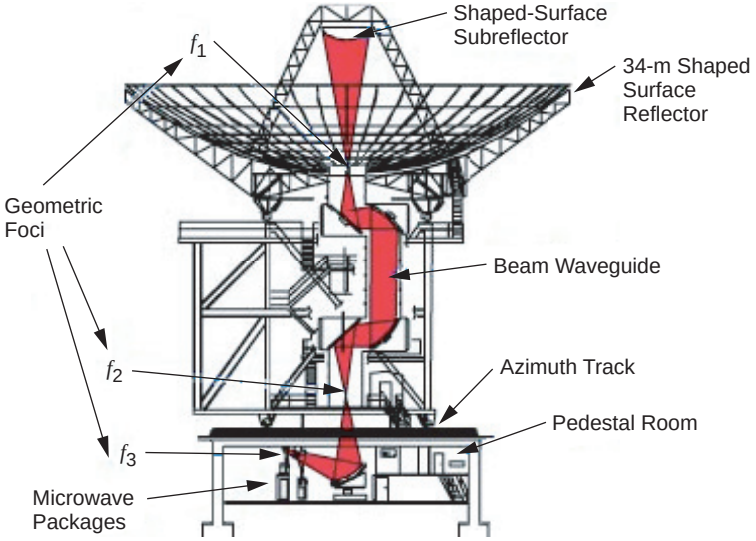


Fig. 7-2. Schematic of a 34-m BWG antenna indicating TPR placement of the microwave packages at *f*3 focus.

Fig. 7-3. 13.8-GHz microwave package TPR during testing prior to installation at DSS-13 at F_3.

In the following section we describe the general requirements for the DSN antenna-calibration effort. This is followed by a discussion of current methods, and their shortcomings, and a final section describes the new approach being now taken in delivering operational Antenna Calibration & Measurement Equipment (ACME) to the DSN.

7.2 Calibration System Requirements

The performance of a DSN antenna must be accurately characterized at the time it comes on line as a new instrument and when new upgrade capabilities are being implemented. Also, certain characteristics must be checked periodically to maintain performance as well as assess the cause of, and correct for, any observed anomaly during normal tracking of a spacecraft.

Calibration procedures require accurate measurement of the various parameters of interest. In the case of antenna calibration, these fall naturally into two categories, those derivable from the measurement of amplitude and phase of a received coherent microwave signal, and those derivable from the measurement of received power from a noncoherent source (such as a radio star). The former case involves the interference of received signals from the antenna under test (AUT) and a small, reference antenna mounted nearby (using the microwave holography described in Chapter 8), while the latter utilizes a total power radiometer (TPR) measuring system to determine the antenna temperature of the source.

The initial calibration consists of precision setting of the individual main reflector panels, precision alignment of the subreflector, determination of aperture efficiency versus antenna elevation angle, and development of a suitable pointing model to permit accurate “blind” antenna pointing.

All of the measurements needed to perform these calibrations involve far-field observations of monochromatic signals transmitted by satellite beacons or broadband radiation from various celestial sources. Holographic measurements are typically performed at X- or Ku-band, based on the availability and elevation angle of suitable geostationary satellite signals, while the remaining measurements utilize in-band S-, X-, and Ka-band frequencies, depending on the equipment planned for the particular antenna under test.

In all cases, some form of sampling of the source radiation is performed as a function of antenna offset from the source. The exact nature of the sampling, as well as the subsequent data processing, determine the precision and accuracy achieved in the overall calibration effort.

In the following section, we briefly enumerate and describe the deficiencies of the conventional approach that has been used in aperture efficiency and pointing measurements. The remainder of the chapter is devoted to a discussion of the approach now being pursued to significantly improve these measurements.

7.3 Conventional Approach to Aperture Efficiency and Pointing Measurements

To measure the gain of large antennas, one can measure the received power from a radio source that has been previously calibrated by independent means [2–5]. A practical method to accomplish this is to measure the received power from the calibration radio source and compare the result with the theoretical result one would measure with a "perfect" antenna. Expressed as a ratio, the result is the aperture efficiency $\eta(\psi)$, where ψ represents the orientation of the antenna, e.g., azimuth and elevation. The formal expression for $\eta(\psi)$ is:

$$\eta(\psi) = \frac{2kT_S(\psi)C_r(\psi)}{AS} \tag{7.3-1}$$

where

A = antenna aperture physical area, m^2

C_r = source size correction factor, unitless measure

k = Boltzmann's constant (1.38065×10^{-23} W/K-Hz)

S = flux density of radio source, W/(m^2-Hz)

T = temperature, kelvin (K)

T_S = antenna noise temperature increase due to the source, kelvin (K)

The antenna aperture, A, for a circular dish of diameter d is the geometrical area $A = \pi d^2/4$. In Eq. (7.3-1), it is assumed that the antenna points perfectly. In practice we attribute the losses due to mispointing of the antenna separately.

While radio astronomy telescopes are designed for maximum beam efficiency, the DSN antennas are designed for maximum aperture efficiency. Beam and aperture efficiencies are functions of the aperture illumination function. The aperture efficiency is at maximum with no taper, while the beam efficiency is at maximum with full taper. The aperture efficiency $\eta(\psi)$ is affected by the areas of the noise shield, subreflector, and struts blockages; the strut shadow; the amplitude taper illumination; the reflector surfaces root mean square (rms) errors; and the ohmic losses.

In radio astronomy, the process of measuring the antenna aperture efficiency is further complicated by the fact that sources for which radio-frequency (RF) flux densities have been accurately measured tend to be rather large in angular size relative to the antenna pattern of a large-aperture radio telescope such as a DSN antenna. In practice, the small angular width of the antenna beam partially resolves the angular structure of the radio source with the result that some of the radio flux density is not collected by the antenna when it is pointed "on source." A "correction for source size" $C_r(\psi)$, is

typically used to compensate for this effect. Note that the value of C_r is expected to vary with antenna orientation ψ because both the beam shape and the angular orientation of the radio source change as the source rises and sets across the sky.

Each of the quantities, C_r and S, contains a source of error, and the measurement method used to determine each must be addressed in any search for improvement. Note that in the DSN we define T at the input to the feedhorn aperture; and therefore, the antenna aperture efficiency, η, is also defined at the same reference point.

7.3.1 Source Size Correction Factor

The source size-correction factor is designed to account for the flux density of an extended source not collected by the antenna, and it is best understood with reference to the fundamental radiometric equation from which Eq. (7.3-1) is derived,

$$kT_S(\psi;\theta,\phi) = \tfrac{1}{2}\eta_\upsilon(\psi) A \iint_{\text{source}} B_\upsilon(\theta',\phi') P_{n,\upsilon}(\psi;\theta'-\theta,\phi'-\phi)\, d\Omega' \qquad (7.3\text{-}2)$$

where

B_υ is the source brightness function,

$P_{n,\upsilon}$ is the normalized antenna power pattern, and

(θ,ϕ) are rectangular, angular coordinates relative to the source center [6,7].

Here, we have been specific regarding the dependence of various quantities on the antenna pointing direction, ψ, and the operating frequency, υ, as well as the fact that the measured system noise temperature increase due to the source, T_S, depends on the antenna pointing.

The integral appearing in Eq. (7.3-2) is the source flux density collected by the antenna, and is smaller than the total source flux density emitted by the source

$$S_\upsilon = \iint_{\text{source}} B_\upsilon(\theta,\phi)\, d\Omega \qquad (7.3\text{-}3)$$

unless the source is much smaller in extent than the antenna main beam and the antenna is accurately pointed at the source. Equation (7.3-2) may be cast into the form of Eq. (7.3-1) by defining the source size correction factor,

$$C_{r,\upsilon}(\psi,\theta_m,\phi_m) = \frac{S_\upsilon}{S_{\text{coll},\upsilon}(\psi,\theta_m,\phi_m)} \geq 1 \tag{7.3-4}$$

where

$$\begin{aligned} S_{\text{coll},\upsilon}(\psi)\big|_{\max} &= S_{\text{coll},\upsilon}(\psi,\theta_m,\phi_m) \\ &= \iint_{\text{source}} B_\upsilon(\theta',\phi') P_{n,\upsilon}(\psi;\theta'-\theta_m,\phi'-\phi_m)\, d\Omega' \end{aligned} \tag{7.3-5}$$

is the *maximum* flux density collected by the antenna, that is, the antenna noise temperature field, $T_\upsilon(\psi;\theta,\phi)$, must be explored at a given elevation angle until the maximum value corresponding to the coordinates $(\theta_m,\phi_m)$is found. It should be noted that these coordinates will not be those for the source center unless the source happens to be symmetric.

Equations (7.3-4) and (7.3-5) imply that the determination of $C_{r,\upsilon}(\psi)$ requires a knowledge of the source brightness function and the normalized antenna power pattern. For those circumstances where $C_{r,\upsilon}(\psi)$ is within a few percent of 1, the usual approach to its evaluation has been to estimate both of these functions by symmetric Gaussians, in which case one obtains the oft-quoted formula

$$C_{r,\upsilon} = 1 + \left(\frac{\Theta_S}{\Theta_B}\right)^2 \tag{7.3-6}$$

where Θ_S and Θ_B are the source and antenna beam widths, respectively. For a disk-like distribution, the correction factor is

$$C_{r,\upsilon} = \left[\frac{1-e^{-x^2}}{x^2}\right]^{-1} \tag{7.3-7}$$

where,

$$x = (4\ln 2)^{\frac{1}{2}} * R/\Theta_B, \tag{7.3-8}$$

and R is the angular radius of the disk [8].

However, many commonly used sources have corrections approaching 100 percent for a large antenna operated at high frequency. The source size-correction value for Virgo A, at Ka-band, on the 70-m antenna, for example, is

calculated to be 1.90. At S-band, the value for source size correction for same source on the 70-m antenna is 1.205 [7]. Under these circumstances, the computation of the source-size correction must be carried out with more realistic functional representations of the source structure, and the source of these has been brightness maps measured with very long baseline interferometry (VLBI) arrays, or large antennas such as the 100-m antenna at Bonn, Germany.

It is possible, in principle, to perform a proper deconvolution of such maps to compute Eq. (7.3-5). Thus, an average brightness map obtained with an antenna having an equivalent normalized far-field power pattern, $P_{0,n}(\theta,\phi)$, given by

$$B_0(\theta,\phi) = \frac{1}{\Omega_0}\iint_{\text{source}} B(\theta',\phi') P_{0n}(\theta-\theta',\phi-\phi')\, d\Omega' \tag{7.3-9}$$

where Ω is the equivalent measuring beam solid angle, where we now drop the explicit frequency and elevation angle notation, and where for simplicity we assume beam symmetry so that the integral has the form of a convolution. Then, taking the Fourier transform of Eqs. (7.3-5) and (7.3-9), we have

$$\begin{aligned} \tilde{S}(\mu,\nu) &= \tilde{B}(\mu,\nu)\tilde{P}_n(\mu,\nu) \\ \tilde{B}_0(\mu,\nu) &= \frac{1}{\Omega_0}\tilde{B}(\mu,\nu)\tilde{P}_{0n}(\mu,\nu) \end{aligned} \tag{7.3-10}$$

from which we obtain

$$\tilde{S}(\mu,\nu) = \Omega_0 \frac{\tilde{P}_n(\mu,\nu)}{\tilde{P}_{0n}(\mu,\nu)} \tilde{B}_0(\mu,\nu) \tag{7.3-11}$$

so that performing the inverse Fourier transform yields $S_{\text{coll}}(\theta,\phi)$ from which $S_{\text{coll}}(\theta_m,\phi_m)$ may be found. In the above, (μ,ν) are the spatial frequency coordinates. This procedure has in fact been used to generate the C_r values currently used in the DSN for calibration purposes [7,8], but the approach has a number of limitations that become serious at high frequencies:

1) Maps are usually not available at the frequency of interest so that an interpolation procedure must be used to estimate a map at the required frequency.
2) Large antennas have significant flexure as a function of elevation angle due to gravitational loading resulting in aberrations that affect the beam pattern,

so that the values of C_r ought to be calculated as a function of elevation angle (see Eq. 7.3-4)).

3) Information on the mapping beam solid angle and shape is often approximate or unavailable in the literature. This can be obtained from holographic measurements [13].

An alternative approach is thus called for that eliminates the need for source size corrections, and this approach is described in a following section.

7.3.2 Flux Density

Since source flux densities are determined from the same equation used to determine aperture efficiency, Eq. (7.3-1), all of the sources of error attendant the latter must apply to the former as well. Thus, while the very brightest sources can be measured with a low-gain system such as a horn, for which calibration is relatively straightforward, the transfer of information from strong to weak sources, which are compact enough to serve as reasonable calibrators for large antennas, must be carried out with larger antennas. Then, Eq. (7.3-1) leads to the result

$$\frac{S_1}{S_2} = \frac{T_1}{T_2}\frac{C_{r_1}}{C_{r_2}} \tag{7.3-12}$$

where the subscripts refer to measurements of two different sources with the same antenna, and we see that not only antenna temperatures, but also source-size corrections, enter into the calculation of flux density ratios for different sources.

A survey of the literature on flux density measurements shows that the use of inaccurate C_r values contributes significantly to the error budget for such measurements [3,5,7,8], so that eliminating the need for such a correction would result in a significant increase in the accuracy of flux density determinations.

7.3.3 Source Temperature

The basic method for measuring the system noise-temperature increase due to a source involves some form of on-source, off-source subtraction. In the conventional approach (sometimes termed "autobore") this is accomplished by a boresight technique in which the antenna is successively offset in a given direction, say θ, relative to the source, by ±5, ±1/2, and zero antenna half-power beamwidths (HPBW). The resulting five data points are then fitted to a Gaussian function plus a linear background to account for the decrease in system noise temperature with elevation angle; and from this fit, the maximum,

or peak source noise temperature, and θ pointing error and beamwidth are determined. This pointing error is then used to execute an orthogonal boresight in the ϕ direction, and the process is repeated as the source is tracked.

While this works well at S-band, it is less satisfactory at X-band, and unsatisfactory at Ka-band, especially with regard to the pointing determination, where (for example) it has been unable to provide the requisite precision to meet the radio science requirements for the Cassini mission to Saturn [9]. Additionally, the method is inherently slow since each of the five measurements in a given direction requires that the antenna servos and mechanical structure settle at the offset specified before a noise-temperature measurement is made. A further problem is that the Gaussian fitting function only approximates the actual profile of the noise temperature measurement (which follows the antenna far-field pattern function), and for an extended source this approximation may not be very good.

In view of these limitations, one would like to have a source noise-temperature measurement of inherently greater accuracy. This would not only improve our knowledge of antenna gain and pointing, but it would also improve the calibration of weak sources by the comparison method described above.

In the following section we describe a new approach to the calibration of large, ground-based antennas that significantly improves the precision achieved by reducing or eliminating the above-noted sources of error inherent with present methods.

7.4 The Raster-Scan Method

The key to reducing the error sources discussed in the previous section lies in making system noise temperature measurements over a finite area of sky including the source, rather than along orthogonal cuts through the temperature profile. Thus, integration of Eq. (7.3-2) over the two dimensional angular field (θ, ϕ) gives

$$k \iint_{\substack{\text{source} \\ +\text{beam}}} T_S(\theta, \phi)\, d\Omega = \tfrac{1}{2} \eta A \Omega S \qquad (7.4\text{-}1)$$

where Ω is the antenna beam solid angle, and we have dropped the explicit frequency and elevation angle notation for simplicity.

If we now consider the application of the above equation to two sources, the equivalent of Eq. (7.3-12) becomes

$$\frac{\iint_{\substack{\text{source}\\+\text{beam}}} T_1(\theta,\phi)\,d\Omega}{\iint_{\substack{\text{source}\\+\text{beam}}} T_2(\theta,\phi)\,d\Omega} = \frac{S_1}{S_2} \tag{7.4-2}$$

as a result independent of source size corrections.

If the source considered in Eq. (7.4-1) is small enough relative to the main beam to be considered a point, then its brightness may be represented by

$$B(\theta,\phi) = S_p\,\delta(\theta)\,\delta(\phi) \tag{7.4-3}$$

where $\delta(x)$ is the Dirac delta function and S_p is the flux density of a point source. Then, Eqs. (7.3-4) and (7.3-5) show that $C_r = 1$ so that Eq. (7.3-1) becomes

$$\eta = \frac{2kT_p}{AS_p} \tag{7.4-4}$$

If the flux density, S_p, of this point source is known, then Eq. (7.4-4) immediately yields the aperture efficiency in terms of the peak source temperature. However, it is frequently the case that point sources bright enough for calibration purposes are also variable, so that one may not have a-priori knowledge of S_p. In this case, Eq. (7.4-2) may be used to determine S_p by comparison with an extended, calibrated source whose flux density, S_c, is known. Then, by combining Eqs. (7.4-2) and (7.4-4), we have

$$\eta = \frac{2kT_p}{AS_c}\,\frac{\iint T_c(\theta,\phi)\,d\Omega}{\iint T_p(\theta,\phi)\,d\Omega} \tag{7.4-5}$$

which now becomes the fundamental equation for determining aperture efficiency. These equations can now be arranged to solve for the source size corrections as follows:

$$C_r(\psi,\theta_m,\phi_m) = \frac{S_c}{S_p} = \frac{T_p(\theta_m,\phi_m)\iint T_c(\psi,\theta,\phi)d\Omega}{T_c(\theta_m,\phi_m)\iint T_p(\psi,\theta,\phi)d\Omega} \tag{7.4-6}$$

The data for the computation implied by Eqs. (7.4-5) and (7.4-6) are the temperature fields $T_p(\theta,\phi)$ and $T_c(\theta,\phi)$ for the point and extended calibration source, respectively; and these are obtained by sc anning the antenna beam across the source in a raster pattern, similar to a (non-interlaced) television (TV) scan (Fig. 7-4). In Fig. 7-4, the raster-scan is designed to image the planet at the center of the image. The deviation of the image position from the center is the result of pointing errors introduced by the antenna combined with difficult refraction correction computations at the low elevation angle of 7.6 deg. The color dots a bove and below the scans are the computed observation coordinate of the source at the mid point of each scan. The combined data from the 33 scans are then displayed at the top left of the display.

All the terms on the right side of Eq. (7.4-5) are either known values, or they are measured by the OTF-mapping system.

It should be noted here that the ratio of the integrals appearing in Eqs. (7.4-5) and (7.4-6) is, by Eq. (7.4-2), just the ratio of the flux densities for the two sources, that is, a constant. Thus, the measurement strategy should involve the alternate scanning of the two sources over a small but f inite elevation change so tha t the data points c orresponding to each integral, as a function of elevation, can be fitted to a linear function, or perhaps quadratic function. Then, it should be found that the ratio of these two fitting functions is constant and equal to S_c / S_p. These results illustrate some of the valuable features of the OTF-mapping system:

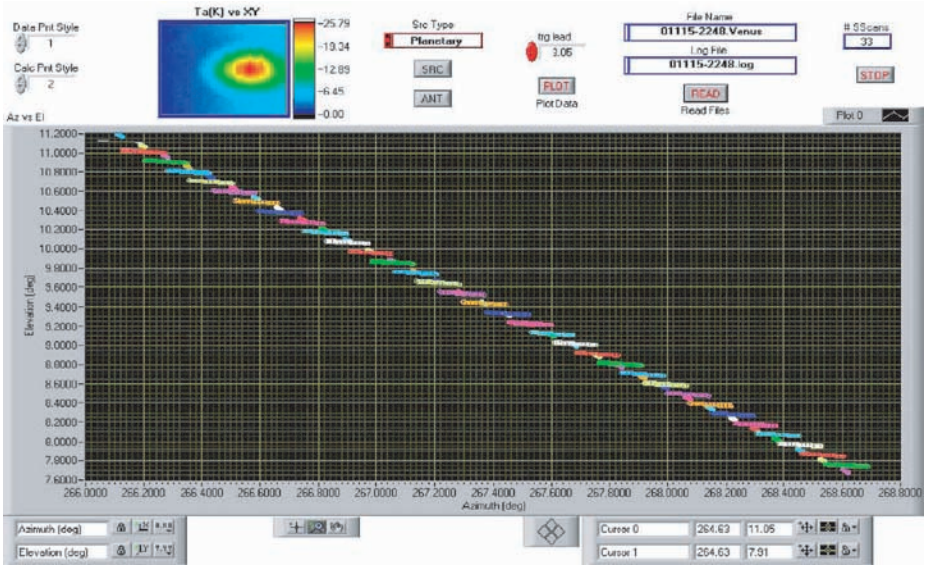


Fig. 7-4. Display showing a raster-scan performed at 13.8-GHz by TPR shown above, creating a 33 x 33 image of Venus as it sets.

1) The need to de rive source size correction factor $C_r(\psi)$ is eliminated. However, one has the option to c alculate $C_r(\psi)$ for various calibration radio sources using Eq. (7.4-6).
2) The need to derive accurate values of T_C for extended sources is eliminated. This is desirable because accurate derivations of T_C require mapping and deconvolving source structure from the antenna patterns, which change with antenna orientation (ψ). Using the OTF-mapping technique the accuracy of the antenna efficiency measurement is limited by the knowledge by which the calibration source flux density S_C is known.
3) The OTF-mapping system enables the experimenter to use compact point-like sources as secondary calibration sources for precision antenna calibrations. The vast majority of these radio sources are quasars, which are so distant that their angular sizes are very small. The problem is their radio brightness (flux density) values are highly va riable so one must calibrate them against the handful of absolutely calibrated radio sources that are available. With few exceptions, the time-scales of the quasar variations are typically a few days, so flux density calibration measurements can be done rather infrequently, and relative measurements of antenna performance with azimuth and elevation can be done almost any time.

As a practical matter, the extended calibration source 3C274 (Virgo A) and the variable point sour ce 3C273 serve admirably for such a strategy as they have nearly the same right ascension.

Substitution of Eq. (7.4-3) into Eq. (7.3-2) gives the result

$$T_p(\theta,\phi) = \frac{\eta A S_p}{2k} P_n(\theta,\phi) = T_p P_n(\theta,\phi) \tag{7.4-7}$$

so that the raster scan data set for the point source has a functional dependence determined by the beam pattern plus a background term due to the sky (in Eq. (7.4-9)), which may be linearly approximated over the small field scanned. Thus, if we assume that the antenna is in good a lignment, there are small system aberrations, and the main reflector is ne arly uniformly illuminated (which is a good approximation for the shaped reflector designs of the DSN), then $P_n(\theta,\phi)$ can be well approximated by an asymmetric Airy pattern

$$A\left(\kappa_\theta\theta,\kappa_\phi\phi\right) = \left[\frac{2J_1\sqrt{\kappa_\theta^2\theta^2+\kappa_\phi^2\phi^2}}{\sqrt{\kappa_\theta^2\theta^2+\kappa_\phi^2\phi^2}}\right]^2 \tag{7.4-8}$$

so that the system temperature data set for the point source raster scan has the form

$$T_{\text{op}}(\theta,\phi) = T_P A\left[\kappa_\theta(\theta-\theta_0),\kappa_\phi(\phi-\phi_0)\right] + T_{\text{op}} + a_\theta\theta + a_\phi\phi, \qquad (7.4\text{-}9)$$

where κ_θ and κ_ϕ are beamwidth parameters, θ_0 and ϕ_0 are the pointing errors, and a_θ and a_ϕ are the sky background coefficients for the θ and ϕ directions, and T_{op} is the system operating noise temperature.

The eight parameters appearing in Eq. (7.4-9) may be found from a nonlinear, least-squares fit [10] to the point-source raster-scan data, thus giving complete information on the peak temperature, and pointing errors and beamwidths for the two orthogonal directions corresponding to the scan axes. The precision of the resulting fit will depend on the noise fluctuations present in the noise temperature data, the scan parameters, and the data processing used, and these are dealt with in the following sections.

7.4.1 Fluctuations in System Noise Temperature

Three main sources of fluctuation of system noise temperature can be identified:

1) Thermal noise generated in the radiometer and atmosphere
2) Gain-bandwidth variations in the radiometer caused by ambient temperature fluctuations of electronic components, especially in the first stages
3) Fluctuations caused by variations in tropospheric density, especially of water vapor content. This is most significant at Ka-band.

In order to characterize and model the performance of the TPR, the two-sided power spectral density (PSD) was measured. The output fluctuations of a typical DSN Ka-band radiometer have been measured as a function of fluctuation frequency over the range 6.5×10^{-5} to 0.5 hertz (Hz), and the results compared with a model based on the above mechanisms. The results are shown in Fig. 7-5, where curve *a* corresponds to the radiometer looking at an ambient load and curve *b* was obtained with it looking at the zenith sky.

Curve *c* is a fit to curve *a* that is decreased by the square of the ratio of the system operating noise temperatures, $T_{\text{op}}\big|_{\text{amb}} \big/ T_{\text{op}}\big|_{\text{sky}} \cong 14.6$, and curve *d* corresponds to a statistical model for tropospheric fluctuations for average conditions at the DSN complex at Goldstone, California [11].

Curve *a*, which is constant at high frequencies and follows a $1/f^2$ dependence at low frequencies, corresponds to thermal noise and gain-bandwidth variations, and if these were the only terms present with the

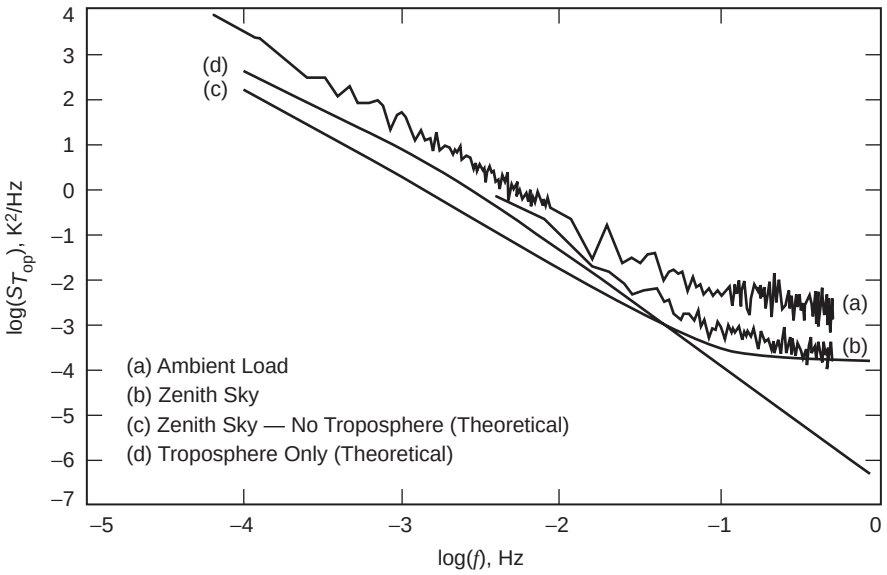


Fig. 7-5. Two-sided PSD of T_{op} fluctuations for a Ka-band radiometer.

radiometer looking at the zenith sky, the data of curve *b* would follow curve *c*. There is a significant departure from this; however, when curves *c* and *d* are added together, the result follows curve *b* closely. From this, we conclude that tropospheric fluctuations play an important role in the total radiometer fluctuations at frequencies below about 0.1 Hz.

Since we are interested in frequencies greater than 10^{-3} Hz when making gain calibrations, that is, times of interest are considerably shorter than 1000 s, we may consider only the high-frequency behavior of the Treuhaft-Lanyi model, which has a $1/f^{8/3}$ dependence so that curve *b* may be represented by the equation:

$$S_{T_{\text{op}}}(f) = S_0 + \frac{K_1}{f^2} + \frac{K_2}{f^{8/3}} \tag{7.4-10}$$

where the coefficients for the Ka-band radiometer tested have the values

$$S_0 = 1.50 \times 10^{-4}\,\text{K}^2/\text{Hz}$$
$$K_1 = 1.64 \times 10^{-6}\,\text{K}^2/\text{s}$$
$$K_2 = 2.36 \times 10^{-7}\,\text{K}^2/\text{s}^{5/3}$$

corresponding to average weather with the radiometer looking at the zenith sky and a T_{op} of approximately 100 K.

With the above form for the PSD of the fluctuations, one may determine the corresponding standard deviation of the fluctuations. This depends on the system operating temperature, T_{op}, the RF system bandwidth, B, the integration time, τ, used during the measurements, and, in view of the nonstationary behavior indicated by Eq. (7.4-10), the total duration of the measurement, T.

It can be shown that the variance of a random process, $X(t)$ of duration, T, having a high frequency cutoff, is given by

$$\sigma_X^2(\mathrm{T}) = 2\int_0^\infty \left[1 - \mathrm{sinc}^2(\pi f \mathrm{T})\right] S_X(f)\, df \tag{7.4-11}$$

where $\mathrm{sinc}(x) = \sin(x)/x$, and $S_X(f)$ is the two-sided PSD of the process. If the $X(t)$ signal is continuously averaged over a time interval τ, the resulting PSD is

$$S_{\bar{X}}(f) = \mathrm{sinc}^2(\pi f \tau) S_X(f) \tag{7.4-12}$$

so that the variance of the averaged process $\bar{T}_{\text{op}}(t)$, of duration T, is

$$\sigma^2_{\bar{T}_{\text{op}}}(\tau, \mathrm{T}) = 2\int_0^\infty \left[1 - \mathrm{sinc}^2(\pi f \mathrm{T})\right] \mathrm{sinc}^2(\pi f \tau) S_{T_{\text{op}}}(f)\, df \tag{7.4-13}$$

The evaluation of this integral for the spectrum given by Eq. (7.4-10) is accomplished by contour integration, with the result

$$\sigma_{\bar{T}_{\text{op}}}(\tau, \mathrm{T}) = \sqrt{\frac{S_0}{\tau} + \frac{2\pi^2 K_1}{3}\mathrm{T} + 18.3 K_2 \mathrm{T}^{5/3}} \tag{7.4-14}$$

where it has been assumed that the measurement duration is considerably longer than the integration time, that is, T >> τ.

The duration of the measurement of interest in the raster scan method depends on the rate at which the data are taken, and the details of the analysis. For example, if one were to operate at a lower frequency than Ka-band, the $T^{5/3}$ term in the above equation, corresponding to tropospheric fluctuations, would be absent, and if a radiometer gain calibration were carried out at the conclusion of each line of the scan, then the appropriate time would be the time required for the execution of a single scan line. Generally speaking, however, T

will be the time required for one complete raster, and an important conclusion to be drawn from Eq. (7.4-14) is the need for *short* measurement times. This, perhaps counterintuitive conclusion, has been born out in actual tests, as will be shown below.

7.4.2 OTF-Mapping Research and Development System Design

The analysis carried out in Sections 7.3, 7.4, and 7.4.1, culminating in Eq. (7.4-5) expressing the aperture efficiency, η, as a function of the source temperature of a point source, demonstrates that the raster scan geometry and timing should be determined primarily by the need to accurately derive this quantity, T_p, and this question is discussed in the following.

Equation (7.4-14), together with the need to avoid settling problems with the antenna mechanical system, suggest that the raster scan should be performed with a continuous motion at a constant, high angular velocity in a given direction, say θ, while discontinuously stepping in the orthogonal direction, again, mimicking a TV scan (Fig. 7-4). This means that the data are taken "on the fly," hence the term, OTF-mapping. In so doing, the averaging process referred to above will contribute to a distortion of the signal that must be taken into account.

A second (and related) consideration is selection of the sampling interval t_s. In view of the Fourier transform relationship between the complex, far-field amplitude, $U_n(\theta,\phi)$, and the complex aperture field, $G(x,y)$, the scan signal for a single line of a point source is absolutely bandlimited.

Thus, for a coherent detection scheme such as that used in the microwave holography system [17,24], which also uses a raster-scan format, the signal is of the form $VU_n(\dot{\theta}t,\phi)$ where V is an arbitrary amplitude factor related to the antenna gain, and $\dot{\theta} = d\theta/dt$ is the constant scan angular velocity. The spectrum of this signal has, by virtue of the clearly defined antenna aperture, a sharp cutoff at $f_0 = \dot{\theta}/2\Theta_B$, where $\Theta_B = \lambda/2a \approx$ antenna main beam-width. This cutoff, moreover, is independent of the main reflector shape, illumination, and system aberrations; and it depends only on the maximum dimension of the aperture, d in the scanned direction.

Similarly, for the noncoherent (TPR) detection used in gain measurements, the signal is of the form

$$VP_n\left(\dot{\theta}t,\phi\right) = V\left|U_n\left(\dot{\theta}t,\phi\right)\right|^2 \tag{7.4-15}$$

so that its spectrum is given by the autocorrelation of the coherent spectrum, and consequently has a cutoff frequency twice as high.

From the above, we infer that Nyquist sampling for a coherent system requires a minimum of one sample per beamwidth; while for a noncoherent system, a minimum of two samples per beamwidth is required. Also, since the signal spectrum is bandlimited in both cases, a sharp cutoff digital filter can be used to remove noise above the cutoff, and this can then be followed by a suitable Wiener filter to compensate for the distortion introduced by the integration, with minimal loss of high-frequency information.

Figure 7-6 shows a schematic block diagram of the antenna calibration OTF-mapping research and development (R&D) system. It can calibrate any of the DSN antennas via interfaces to their antenna controller, encoders read out, and diodes and microwave controls. All the interfaces must be done locally at the antenna under test (AUT). The system achieves high accuracy of raster alignments by the virtue of interfaces to the antenna angle encoders, which are being read at a high speed of 1000 readings/s. An internal computation engine in the data acquisition converts the sidereal motion of radio sources from right-ascension declination (RA-DEC) to antenna coordinates in azimuth-elevation (AZ-EL) in real-time, allowing for a tight feedback loop control of synchronizing receiver triggering (TPR) to the antenna position.

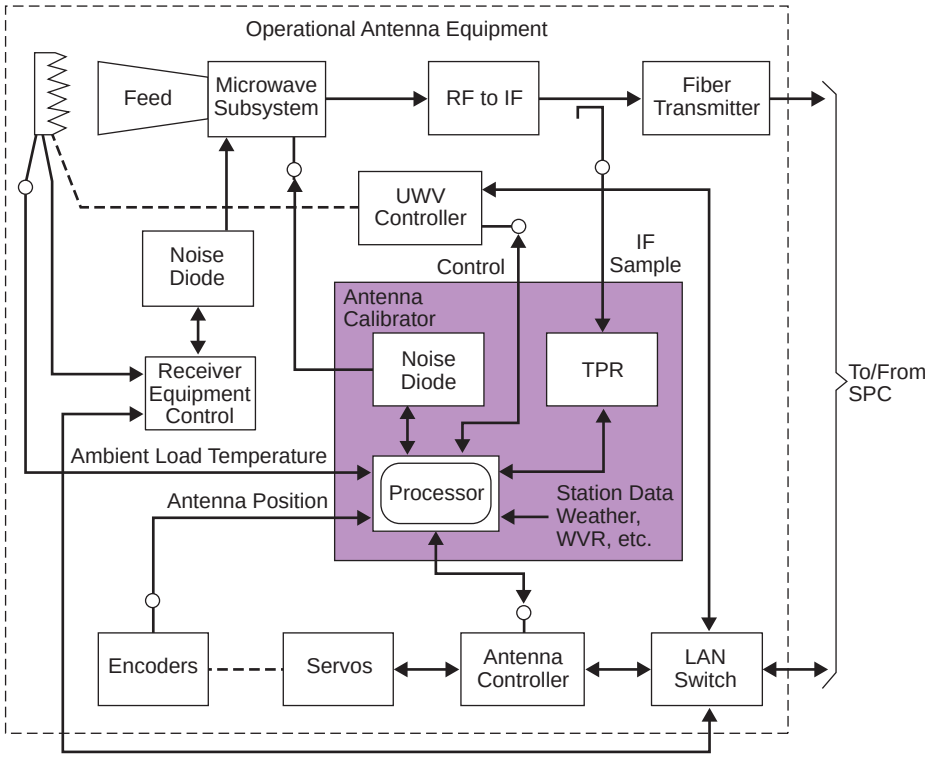


Fig. 7-6. OTF-mapping (raster-scan) R&D system, in shaded area showing interfaces to any DSN antenna (UWV = microwave).

The first trigger position in each of the raster subscans is determined by position synchronization, while the remaining data are triggered via time synchronization locked to the internal system clock. As a result, one of the critical elements of this design is the need to maintain a constant, known angular velocity of antenna motion during the taking of data along a given direction. The TPR is sampled at a constant known rate so that the relative position at which the data are taken is known with high accuracy; and thus, antenna-settling time is no longer an issue. This design ensures the alignment of the individual subscans within the full raster. The OTF–mapping R&D data acquisition algorithms include the computation of the radio source positions such that at any given time the position of the antenna relative to the source is known. Since the data are taken "on the fly," the integration occurring during the sampling interval results in an attenuation of high-frequency information ("smearing"), but this can be recovered by an inverse filtering process (Wiener filter). Since the source is scanned in a raster-type pattern by stepping from line to line, a complete data set corresponding to a complete raster contains all of the relevant data and not just a sampling of it along two orthogonal directions. This means that one is effectively including all of the source radiation so that no source-size correction is necessary (Eqs. (7.4-5) and (7.4-6)). The resulting three-dimensional (3-D) data set is then used to determine, by means of least-squares fitting, the main beam pattern, from which, the relevant calibration parameters are directly determined.

Figures 7-7 and 7-8 show the real-time display of the OTF-mapping R&D data acquisition instrumentation display for the two-dimensional (2-D) and 3-D cases, respectively. In both cases Eq. (7.4-9) is solved using the Levenberg-Marquardt method to determine the non-linear set of coefficients of these equations, which minimizes a chi-square quantity. In Fig. 7-7, the white dots

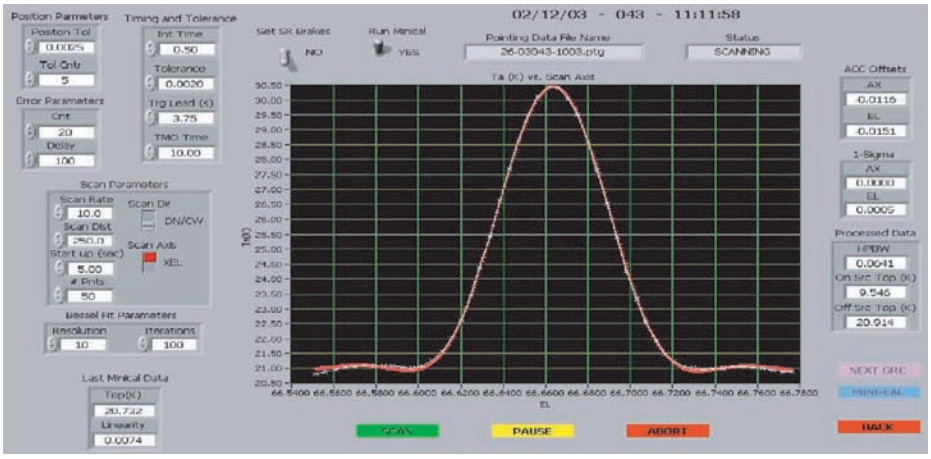


Fig. 7-7. OTF-mapping R&D data acquisition real-time display.

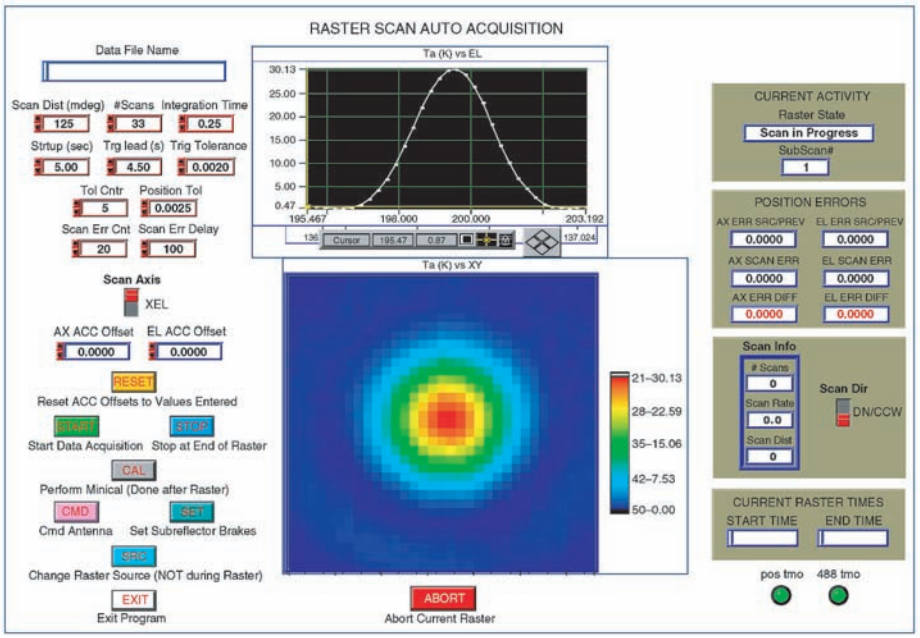


Fig. 7-8. Real-time display of the OTF-mapping R&D instrumentation during the acquisition of a 3-D raster.

are the raw data, and the red line is the fitting Airy function. Figure 7-8 shows the real-time display during the acquisition of 3-D raster (middle color plot), while individual subscans are shown above in white over black plot. In both cases, the precision in the estimation of the equation parameters are determined from the diagonal elements of the covariance matrix and also displayed in real-time for parameters of interest. Figure 7-9 shows the main panel program of the R&D OTF-mapping R&D system.

In order to study the interaction between scan velocity, $\dot{\theta}$, array size, N, and integration time, τ, computations have been made of the errors expected in the fitted parameter T_p for a range of values for each of these parameters for a one-dimensional fit corresponding to a single scan line, and the results are shown in Table 7-1. In Table 7-1, rms background noise from a single scan line, σ_l, and complete raster, σ_r, are computed for given scan line duration, t_l, and raster duration, t_r, from a model based on the measured power spectral density for the radiometer system. In all cases, the sampling interval, $t_s = \tau/2$. The computed errors in T_p are based on a general, nonlinear least-squares fitting analysis, using Eq. (7.4-14) to estimate the noise standard deviation, and a Gaussian beam pattern rather than an Airy pattern, for simplicity.

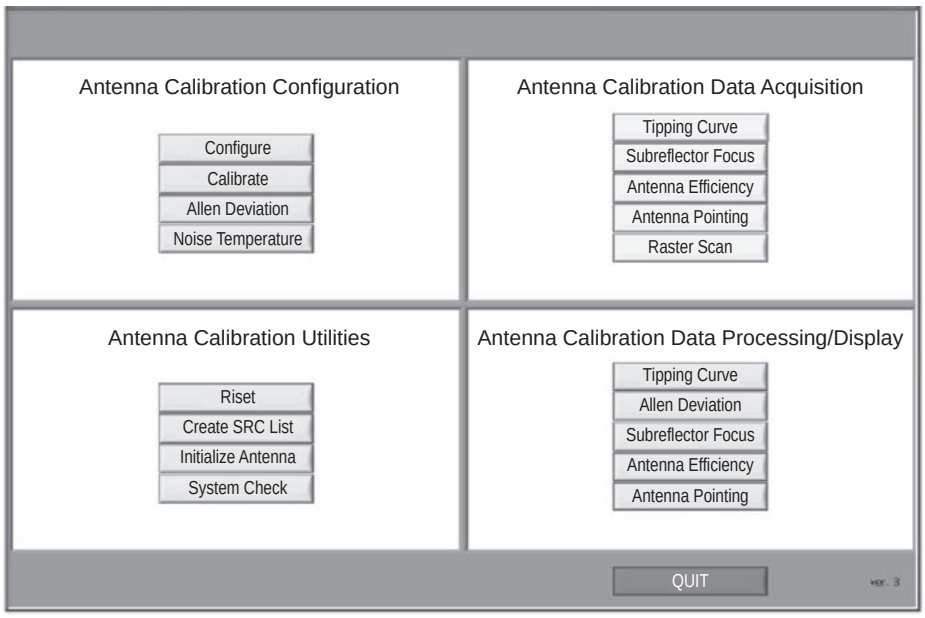


Fig. 7-9. OTF-mapping R&D main program panel.

Also shown in Table 7-1 are the rms fitting errors of aperture efficiency versus elevation curves based on quadratic fits to the data for a complete 6-hour pass of the source. These errors are inversely proportional to $\sqrt{N_r}$, where N_r is the number of complete rasters executed during the pass, each of which yields an estimate of all of the fitting parameters.

$$T_P(\theta) = T_P \exp\left[\frac{-\kappa_\theta^2(\theta-\theta_0)^2}{2}\right] \tag{7.4-16}$$

7.4.3 Test Results

Field test results agree very closely with the computed performance predictions presented in Table 7-1. As a typical example, an observation of Venus with the following parameters (see Fig. 7-10):

Source: Venus

Raster size: 33 × 33

Total measurement distances: 125 × 125 millidegrees (mdeg)

$\tau = 0.25$-s

Scan velocity = 15.6-mdeg/s

Total measurement duration = 430-s

Table 7-1. Theoretical simulation errors for the source temperature, and aperture efficiency η(ψ) versus elevation curve, as a function of the scan velocity, array size, N, based on nonlinear least-squares fitting analysis for raster scan data acquired with 34-m antenna at Ka-Band with an elevation of 90 deg and with a troposphere retrace time of 2 s.

τ(s)	t_l(s)	t_r(s)	$\dot{\theta}$(m deg/s)	$\sigma_l(K)$	$\sigma_r(K)$	$\sigma_T(K)$	$\bar{\sigma}_{FIT}$(%)	N_r
				$N = 17$				
0.1	0.8	47.6	100	0.036	0.068	0.037	0.042	453
0.2	1.6	61.2	50	0.026	0.074	0.041	0.052	352
0.4	3.2	88.4	25	0.019	0.094	0.052	0.080	244
0.8	6.4	142.8	12.5	0.017	0.136	0.075	0.147	151
1.6	12.8	251.6	6.24	0.022	0.214	0.118	0.309	85
				$N = 33$				
0.1	1.6	118.8	50	0.037	0.123	0.034	0.060	181
0.2	3.2	171.6	25	0.027	0.159	0.044	0.095	125
0.4	6.4	277.2	12.5	0.026	0.232	0.064	0.176	77
0.8	12.8	488.4	6.24	0.024	0.369	0.101	0.370	44
				$N = 65$				
0.1	3.2	338.0	25	0.038	0.275	0.038	0.115	63
0.2	6.4	546.0	12.5	0.029	0.405	0.055	0.215	39

Encoder sampling rate = 1000/s

TPR sampling rate = 150/s

The peak temperature was obtained by fitting the data to a 3-D Airy function, which resulted in 1-sigma error of 0.085 K. This result is very much in agreement with the predictions computed in Table 7-1.

The system also computes the reduced chi-square such that if high values are computed, it can be concluded that a main source of the error is due to the fact that the fitting function cannot follow the data to within the limit imposed by the random data errors. In field measurements, we found out that indeed this is the case when the antenna sidelobes are asymmetric as would be due to poor subreflector alignment.

7.5 Blind-Pointing Calibration

The OTF-mapping R&D system also proved itself capable of providing a new record of best blind-pointing performance, which was achieved on the DSN 34-m BWG antennas. In doing so, two new technologies were used:

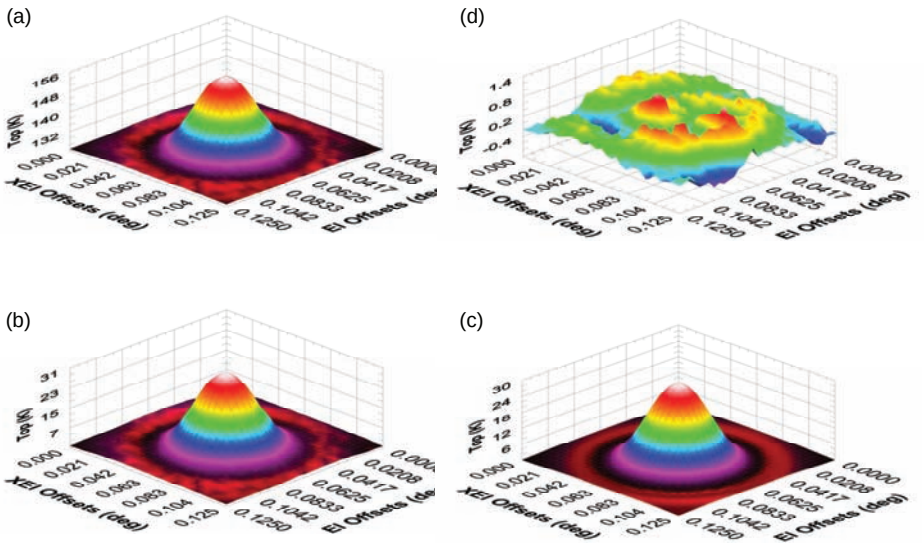


Fig. 7-10. Screen image displaying error analysis for a full 3-D raster scan indicates that an accuracy of 0.1 K was achieved in determining the source peak temperature: Top left: Raw data. Bottom left: Raw data after removal of background and slope. Bottom right: Fitting function to raw data on bottom left. Top right: Map-differencing between raw data and fitting function.

1) The OTF-Mapping R&D system [12] and
2) New 4th order pointing model software [14].

The 4th order pointing model was devised as a result of noticing systematic error residuals remaining in the data after applying the conventional 1st order model. The 1st order model, which typically has six to eight mathematical terms (Fig. 7-11), is a physical model originally developed by Peter Stumpff and published in "Astronomical Pointing Theory for Radio Telescopes" in 1972 [23]. The 4th order model (Fig. 7-12) was derived by expanding the spherical harmonics that are related to the associated Legendre polynomials by Eqs. (7.5-1) and (7.5-2) below, to the 4th order, resulting in 50 mathematical terms:

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{2l+1(l-m)!}{4\pi(l+m)!}} P_l^m \cos(\theta) e^{im\phi} \qquad (7.5\text{-}1)$$

where,

$$P_l^m(x) = (-1)^m \left(1-x^2\right)^{m/2} \frac{d^m}{dx^m} P_l(x) \qquad (7.5\text{-}2)$$

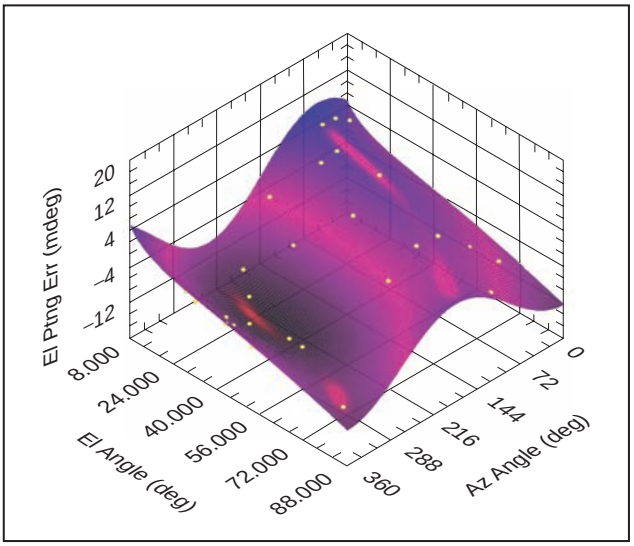


FIg. 7-11. Traditional 1st order pointing model resulting in a predicted performance of 2.74-mdeg MRE.

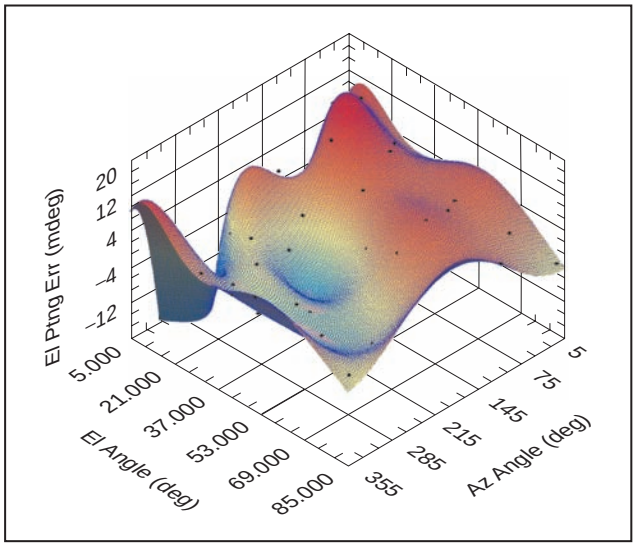


FIg. 7-12. New 4th order pointing model resulting in a predicted performance of 1.49-mdeg MRE.

All the physical terms from the 1st order model that did not appear in the expansion were retained in the new model, resulting in a total of 59 mathematical terms. As shown in Figs. 7-11 and 7-12, the application of the new 4th order model reduces the predicted mean radial error (MRE) by a factor

of approximately 2. Additional field tests confirmed that the blind pointing performance improved by approximately a factor 2–3 relative to the 1st order model.

To facilitate an efficient all-sky survey for the observation of radio sources, a scheduling program was written and integrated within the OTF-mapping R&D system. Fig. 7-13 is the output produced by the scheduling program for DSS-13, where each (yellow) dot represents a radio source to be observed and data recorded utilizing a 2-D cross-scan as illustrated in Fig. 7-7. The gathered data are then processed by the 4th order pointing model software that computes a new pointing model for the antenna. For now, we only want to present the final proven results. When the derived model was applied to the DSS-26 BWG antenna at Goldstone, California, and used in operational activity to track Voyager I, a new record of performance level of 3.5-mdeg mean radial error (MRE) was achieved, that was previously never attained on any of the 34-m BWG antennas (typical performance level of these antennas was 7–10 mdeg MRE at best). The result of this track is presented in Fig. 7-14.

What else is special about the data distribution of Fig. 7-14 is that in addition to the low MRE value, the data have a zero mean. This is particularly interesting, because in prior residual pointing error data plots, a sharp transition was observed as spacecraft moved across the meridian. This result was helpful in helping diagnosing the cause of that hysterisis and attributing it to an elevation encoder coupler [15].

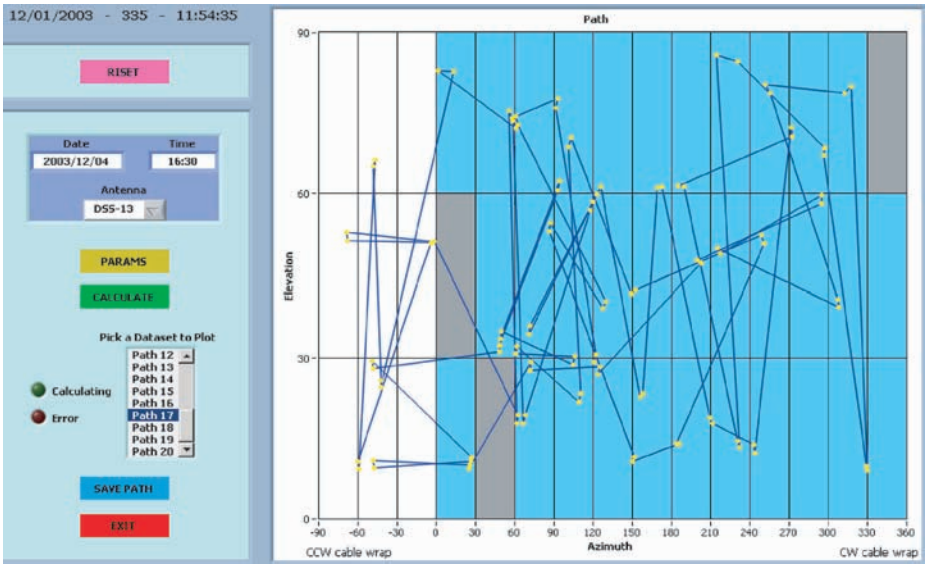


Fig. 7-13. OTF-mapping R&D scheduling display for efficient all-sky source selection and observation (CW is clockwise and CCW is counterclockwise).

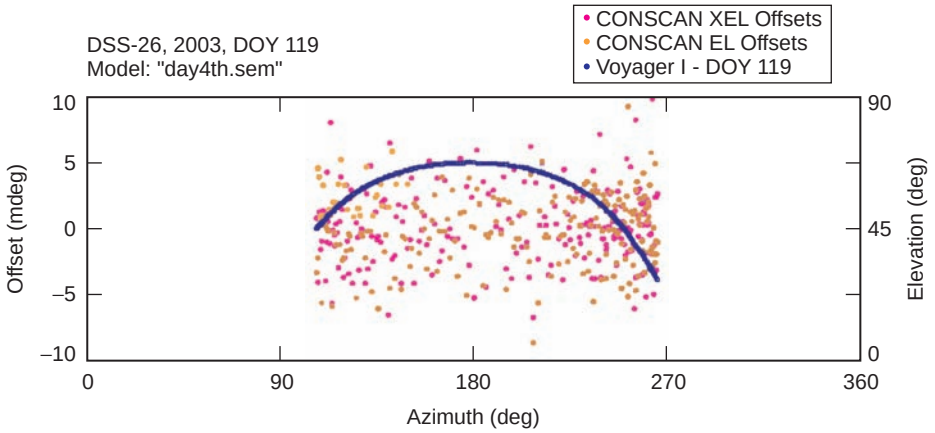


Fig. 7-14. DSS-26 tracking Voyager I (2003, DOY 119) with 3.5-mdeg MRE using 4th order model ("day4th.sem").

7.6 Cassini-Jupiter Microwave Observation Campaign (Cassini JMOC)

7.6.1 Introduction

The objectives of the Radar Instrument on board the Cassini-Huygens spacecraft are to map the surface of Titan and to measure properties of Saturn's rings and atmosphere. However, utilizing the Cassini radar as a radiometer can provide invaluable information regarding the atmosphere and surface compositions of Saturn and its moons, as well as Jupiter. However, the fact that the Cassini-Huygens onboard radar was not calibrated as a radiometer prior to launch was a deterrent for making such high-accuracy measurements. The flyby of the Cassini-Huygens spacecraft past Jupiter in December 2000 provided an opportunity to calibrate the onboard radar as a radiometer utilizing Jupiter as its known temperature load. The accuracy with which Jupiter's disc temperature could be determined from ground observations would translate directly to the accuracy of the calibration of the onboard radiometer; and consequently, it would determine the accuracy with which atmosphere and surface measurements of Saturn and Titan can be made. The fact that the Cassini onboard radar operates at 13.78 GHz was an additional challenge since none of the JPL-DSN ground antennas was equipped with a feed at this exact frequency.

In a presentation made to the principal Investigator (D. Rochblatt, presentation to Mike Klein and Mike Janssen, dated May 7, 1999), the strategy for the measurements and calibrations was laid out. The goal was to measure Jupiter's disk temperature with a 1-sigma accuracy of 2-percent, which if it

could be achieved, would enable new science. The technique is based on performing high-accuracy ground-based calibration measurements simultaneously with the spacecraft observations and at the exact same frequency of 13.78 GHz. This allows us to transfer to the Cassini radar receiver the ground-based radio astronomy flux calibration with high accuracy, using Jupiter as a common reference source. What made this calibration challenging is the fact that absolute calibration measurements of radio sources near 13 GHz did not exist. Current estimates of the absolute uncertainty of the radio astronomy flux calibration scale tend to increase with frequency in the centimeter-to-millimeter radio astronomy bands. Typical estimates of systematic errors in radio source flux measurements near 5 GHz are ~2 percent (1-sigma), whereas the estimates near 22 GHz are ~10 percent (one-sigma).

To achieve the maximum accuracy, a ground based TPR (Fig. 7-3) was designed, built, and installed at DSS-13, 34-m BWG R&D antenna (Fig. 7-2), while incorporating the OTF-mapping R&D system technique described above. After the installation of the TPR, and to support these in-flight calibrations, a coordinated series of ground-based observations named the Cassini-Jupiter Microwave Observing Campaign (Cassini-JMOC) was carried out from November 2000 through April 2001.

The second objective of the Cassini-JMOC project included an educational component that allowed middle-school and high-school students to participate directly in the ground-based observations and data analysis. The students made their observations as part of the Goldstone Apple Valley Radio Telescope (GAVRT) project.

7.6.2 Observations

The 34-m GAVRT antenna was used to participate in a multi-frequency campaign to study Jupiter's synchrotron radiation [16]. GAVRT students and teachers teamed with professional scientists and engineers to measure the ratio of Jupiter's flux density relative to those of six calibration sources that were selected to mitigate different sources of random and systematic errors. The calibration source selection criteria included the following:

- Flux density greater than 2 Jy to ensure high signal-to-noise ($5 < SNR < 10$) for individual measurements.
- Spectral Index is *known* with sufficient accuracy to interpolate the flux density at 13.8 GHz.
- Angular size should be small compared to 0.041 deg (the 3-dB width of the 34-m antenna beam at 13.8 GHz).

The source 3C405 (Cygnus A) was *exempted* from these selection criteria because it is one of the sources that was also being measured directly from Cassini during special calibration sequences in the fall of 2000 and other times during the mission. There is evidence that the source does not vary with time

and that its circular polarization is small (4 percent). Its proximity to Jupiter in the sky (right ascension and declination) was an advantage. All measurements of Jupiter and the calibration sources were processed to remove sources of error caused by changes in system performance with antenna tracking in azimuth and elevation. System "mini-cal" sequences were performed about three times per hour to monitor subtle changes in receiving system gain, stability and linearity.

7.6.3 Results

Tests were conducted in April 2001 at DSS-13 using the OTF-mapping R&D instrumentation to observe Venus, Jupiter, 3c405, 3c273, 3c274, 3c286, NGC7027, 3c123, and 3c84. Given that the antenna half power beamwidth (HPBW) for the 34-m antenna operating at 13.8 GHz is approximately 0.041 degrees, the raster dimensions were scaled for approximately three times the HPBW to produce maps with spatial dimensions of 0.125 × 0.125 deg on the sky. The temporal resolution along the scan corresponded to about 1/10 HPBW (approximately 0.004 deg), which resulted in data arrays of 33 × 33 points (Fig. 7-8). Typical raster-scans required 15–20 minutes to complete. When the weather was calm, excellent raster alignment was achieved. However, when the wind speed was above 16 km /hr (10 mph) apparent misalignments in the raster were observed (see Fig. 7-15, the raster data at 34.3-deg elevation). The source of this problem is believed to be sub-reflector oscillations because no apparent misalignment in the antenna angle encoders registering was noticed.

In deriving the DSS-13 antenna efficiency, the OTF-mapping R&D system was used tracking 3c273 (point source) and 3c274 (calibrated source) near the rigging angle of 49 deg, where an efficiency of 61 $\pm$2 percent was computed. Then the OTF-mapping data of Venus obtained from 8.7- to 58-deg (Fig. 7-15) elevation was also calibrated at 49-deg elevation to that value. In the process, a source size correction of 1.070 was computed for 3c274 near the same rigging angle, which compares well with the independently derived value of 1.075, obtained using the other scanning techniques that required many more observations.

The data and the plots demonstrate the capability of the OTF-mapping R&D system to reveal distortions in the antenna beam pattern at low elevation angles as is clearly shown at 8.7- and 12.7-deg elevations. These distortions are

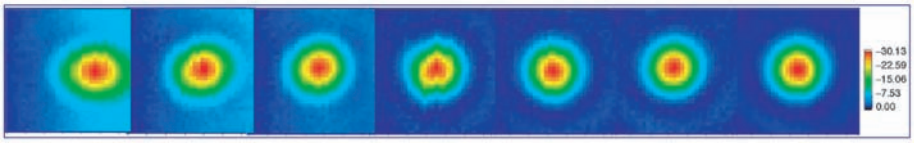


Fig. 7-15. A series of patterns taken of Venus at source elevations from 58 deg to 8.7 deg.

due to gravity-induced deformation of the main reflector surface [17]. At 8.7 deg, the antenna pointing error causes the image to be off center. The mispointing of the antenna is due in large part to errors in estimating the refraction correction at such a low elevation angle.

Figures 7-16(a) and (b) illustrate the stages of data processing for the OTF-mapping R&D system. Figure 7-16(a) shows the 3-D response of the raster

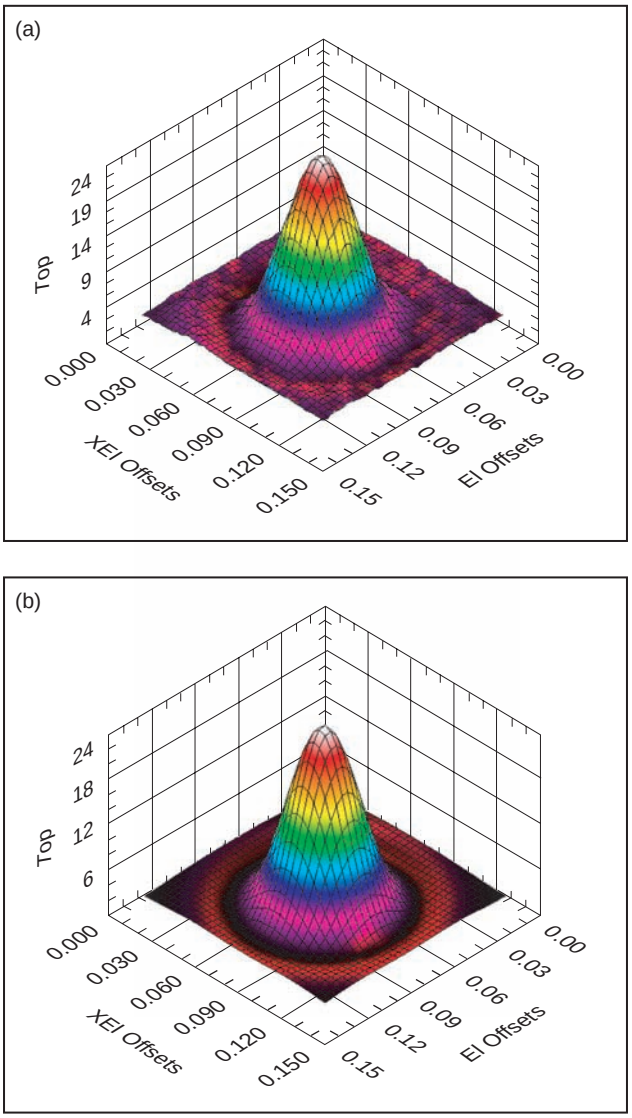


Fig. 7-16. Stages of data processing for the OTF-mapping R&D system, including (a) 3-D response from Venus and (b) fitting of raw data with an Airy function. (XEL is cross elevation).

scan data taken across Venus when it was near 58-deg elevation. The plot was constructed after removing the background noise, which is fitted to a two-dimensional baseline with arbitrary slope (Eq. (7.4-9)). Venus was close to Earth; and therefore, it was a very strong radio source when the measurements were made. Consequently, the noise level in the map is very low, and the smooth surface of the 3-D plot indicates the excellent alignment of the individual raster sub-scans. The x and y-axis coordinates correspond to elevation (El) and cross elevation (XEl) of the maps. The samples along sub-scan direction are approximately 0.0039 deg. The z-axis shows the measured system noise temperature in kelvins.

Figure 7-16(b) is the result of fitting the raw data with an Airy function (Eq. (7.4-8)). The mathematical expression that describes the spatial smoothing is caused by diffraction when radio waves (or light waves) are reflected off a circular aperture, which in this case is the 34-m effective-parabolic dish. The formal equation for diffraction of radio antennas is the Jacobi-Bessel series expansion of the far-field pattern of the antenna.

Figure 7-17 shows the data processing of OTF-Mapping data of 3c405, Cygnus A, using DSS-13. Since 3c405 is an extended source for this antenna at 13.8 GHz, the convolution of the source with the antenna main-beam resulted in the image on the top and lower left of Fig. 7-17. (The lower left of Fig. 7-17 is derived after the removal of the atmospheric component contribution to the noise temperature). The lower right corner of this figure displays the Airy pattern model of the antenna main-beam. After the subtraction of the main-beam from the data, the double-lobed shape of Cygnus A is revealed. A VLBI image of Cygnus A taken by the VLA is shown in Fig. 7-18 for reference. The criticality of a 3-D raster-scan for accurate determination of source size correction is clearly demonstrated by this process.

The observed ratios of Jupiter to the six calibration sources were used to calculate the effective disk temperature of Jupiter from each calibrator. The spectral indices of Venus and the sources 3C286, 3C123, and NGC 7027 were updated with new results from the National Radio Astronomy Observatory (NRAO). The result is shown in Fig. 7-19. The average disk temperature was computed to be 165 K ±2 K. This signifies an accuracy of 1.2 percent, which exceeds the project goal. This accuracy translates directly to the accuracy with which the Cassini Radar can be used as a radiometer to study the atmosphere and surfaces of Saturn and its moons. Most notable, these results were possible due to the new set of observations carried out at Goldstone to map the brightness distribution of 3C405 and 3C274 using the OTF-mapping technique described above. If these observations had been made by conventional auto-bore measurement techniques, an accuracy of only 4.3 percent would have been achieved [18].

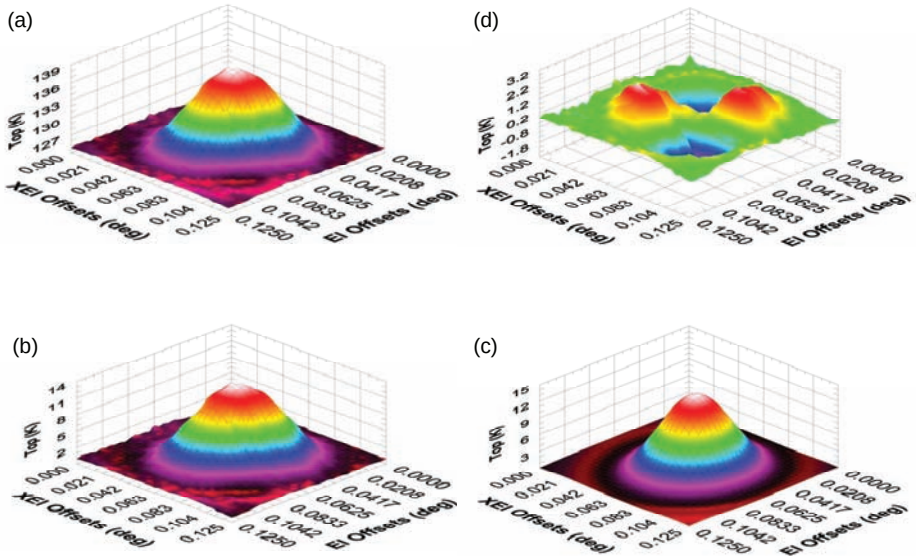


Fig. 7-17. 3c405, Cygnus A mapping by a single 34-m antenna at 13.78 GHz.

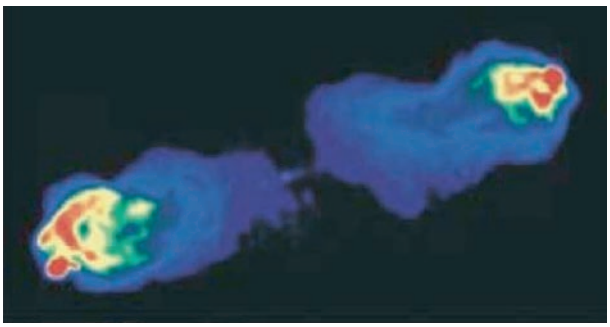


Fig. 7-18. Cygnus A image courtesy NRAO/AUI/NSF
Investigator: R.A. Perly.

The raster scan technique reduces the uncertainty in the total flux density measurement that arises when the antenna beam partially resolves the spatial dimensions of an extended radio source.

7.7 Operational Antenna Calibration & Measurement Equipment (ACME) for the DSN

The OTF-mapping R&D system provided a complete functionality in a portable package; however, it did not provide the best architecture suitable for an operational DSN environment. It was desired to have an antenna calibration system that could provide all these functionalities from the centralized DSN

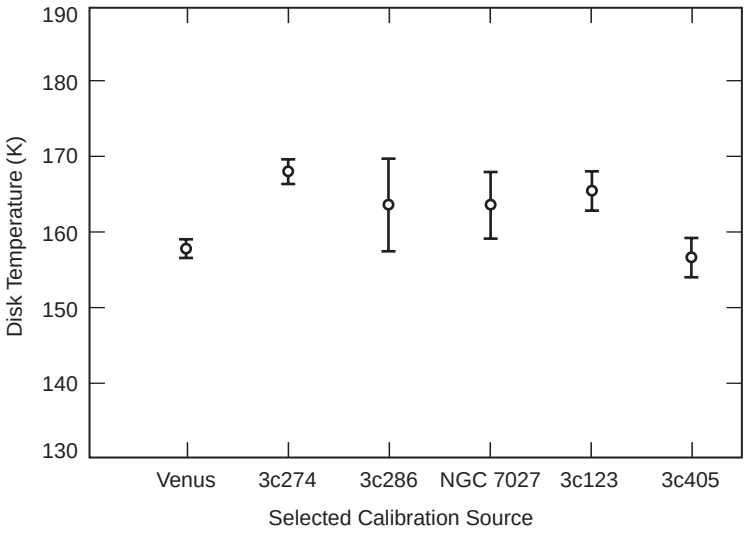


Fig 7-19. Jupiter disk temperature determined by OTF-mapping technique. (Data in the figure were taken with an accuracy of 1.2 percent, a Jupiter disk temperature of 165 K, and an observation frequency of 13.8 GHz.)

Signal Processing Center (SPC) and interfacing it to the Network Monitor and Control (NMC) of the DSN subsystem. ACME [25] was designed to run over the SPC-LAN automating these procedures using standard monitor and control without modifying operational environments.

The other key design feature of the operational system that is different from its R&D predecessor is in its synchronization implementation. While in the R&D system synchronization is based on position and timing (Section 7.4.2 above), ACME synchronization is based on time alone. The time synchronization is provided by computing a predict file for the antenna controller, which describes the exact antenna positions relative to the radio sources during a complete raster (for either the 2-D or 3-D scans) as a function of absolute time. These time stamps within the predict file provide the synchronization with the radiometer recorded values. A block diagram of ACME interface in the DSN environment is shown in Fig. 7-20.

7.7.1 ACME Major Capabilities

ACME uses noise-adding radiometer techniques to compute system noise temperature (SNT) values that can be used to compute pointing offsets and antenna efficiency and subreflector optimization for different feeds.

The two channels provided with the system, enable simultaneous measurements of two frequency bands, or two polarizations in the same band. This feature is especially useful in determining antenna beam coincidence at

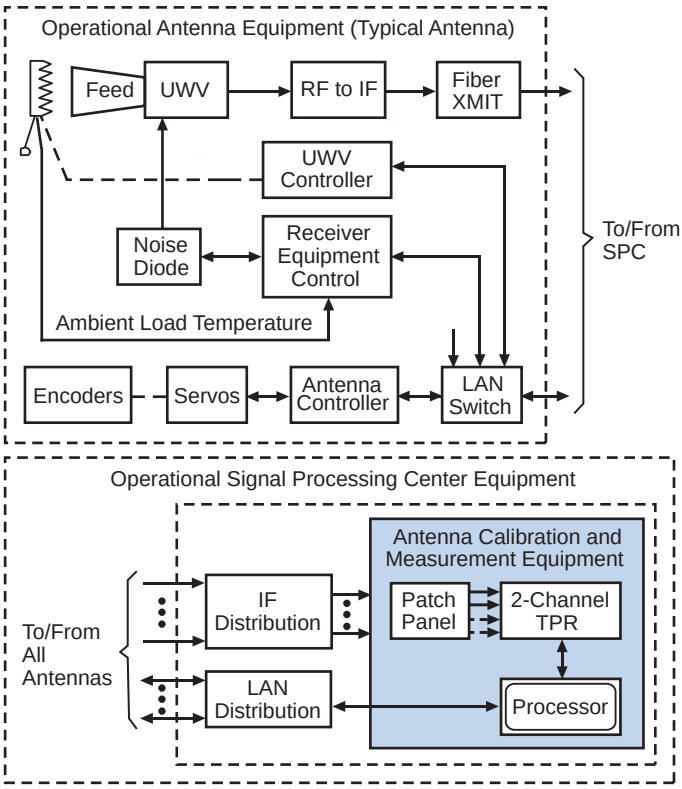


Fig. 7-20. ACME system block diagram.

different frequency bands and greatly improves productivity of time devoted to calibration. The new system can be used to evaluate non-modeled phenomena such as coupler hysteresis (el or az different readings at same position coming from different directions), and weather affects on a ntenna pointing performance. Surface deformations caused by temperature gradients and strong wind and bad refraction correction can be examples of weather interference in antenna pointing performance. In addition the system can be used to measure the antenna track level unevenness, and detection of servo anomalies. The system provides for the maintenance of pointing models from previous observations, for refining accuracy and provides general archive of observation data, for trend and historical data analysis.

7.7.2 Subsystem Design and Description

ACME is designed to perform the calibration activities by interfacing with the existing resources at the station and measuring the noise power with a square-law power meter. A single equipment rack is installed at each SPC that interfaces with the antenna servo, the microwave switches, the noise diodes and

the antenna controller through "predict" distribution. The software uses current communication protocols used on the DSN SPC LAN.

The calibration activities are performed without changing the station operational configuration. This dramatically reduces the risk over the subsequent spacecraft tracking passes. The preparation time is small enough to allow making use of virtually all antenna free times to obtain usable data. More calibration data will be available in the next years to increase the knowledge over the station calibration status, the degradation rates, and a number of hitherto unknown factors that impact the pointing accuracy and efficiency of the antenna.

The power measurements are being carried out by a set of band-limited filters of 5-percent bandwidth (BW) and using broad-bandwidth square-law detectors. Signals from the complex interface (IF) switch distribution allow for selecting any front end, and can be applied to either a 250-MHz (BW = 12.5-MHz) center frequency, 321-MHz (BW = 16-MHz) center frequency, or a tunable filter from 200 to 400 MHz (BW = 5 percent).

ACME uses noise adding radiometry (NAR) [19] techniques with the 50-K diode to calibrate operational parameters. The process is highly automated and does not require any manual intervention for configuration. The radio source catalog from year 2000 is maintained within the system; and it computes nutation, precession, and diurnal and annual aberration to determine current position and build cosine director type of "predicts" for later antenna controller distribution and synchronization. When executing continuous scans, the width of each scan is typically set to as much as five times the HPBW over the source. The system radiometer measures and integrates noise power to derive a far-field antenna pattern over a calibrated rectangular coordinate system, normally, elevation versus cross-elevation.

7.7.3 Radiometer Calibration

Power measurements are derived by switching a 50-K diode as reference over ambient load and sky in a NAR [20,21] mode. Microwave switch configurations, as well as diode modulating control signal, are fully automated under ACME control. The precision achieved during measurement of total power is 1 percent, and while operating in a NAR mode it is 1.2 percent.

7.7.4 Pointing Measurements

As of today, pointing is the main application of ACME. As DSN moves up its operating frequency, pointing precision becomes more and more relevant.

The calibration system must maintain systematic error models and is able to collect data in a variety of conditions without interfering with the DSN operations schedule. ACME performs these functions, and it is able to give an

overall picture of pointing quality in less than 5 hours under normal weather conditions.

ACME is operated with a user friendly graphic user interface (GUI). Clicking over a source on the general source display map, causes the given source to be included in the source list for a given session observation. The "predicts" are built for the sources on the list for later distribution to antenna controllers which, will direct antenna movements to scan the sources.

The main computations engine of the system is based on a nonlinear Levenberg-Marquardt regression using Eqs. (7.4-8) and (7.4-9) from which SNT values are derived from measured data.

The composition of the two-axis (El and Xel) scan gives the basic data to the system for calibration. The center of the scan is the theoretical position of the radio source. From the distance of the maximum noise power relative to the center of the scan, the position error of the antenna in the measured axis is computed. The base line of the noise is the background noise, so the curve T_{op} is the source temperature measured with the antenna. If the source is an accurately modeled radio source, the antenna efficiency can be derived from this measurement. The width of the curve at the –3-dB level is the main beam HPBW.

Data derived from computed offsets are used to derive either first- or fourth-order systematic error models. ACME provides a model calculator that can read data from the system or other sources. In addition, it provides an input filter to apply to input data. Another feature is the ability to "fill" empty areas of the sky with data from previously built models.

7.7.5 Subreflector Optimization

Subreflector misalignment translates into antenna efficiency loss. The Ruze equation gives an expression for this loss:

$$\eta = e^{-\left(\frac{4\pi\varepsilon}{\lambda}\right)^2} \tag{7.7-1}$$

where ε is proportional to the subreflector displacement [22].

To determine the optimal subreflector position, ACME determines the maxima of the curve derived from the subreflector movement over the selected axis. Two orthogonal scans are performed for every position.

The equation used to calculate loss of efficiency derived for small pointing errors is:

$$T = T_0 \cdot e^{-\left(\frac{2.273\cdot\theta^2}{HPBW^2}\right)} \tag{7.7-2}$$

7.8 Conclusions

A theoretical analysis of gain and pointing calibration methods, together with a realistic assessment of system noise characteristics, has led to the conclusion that significant improvement in performance can be realized by performing rapid, continuous raster scans of point and extended radio sources, and by determining temperature and pointing information from two-dimensional, nonlinear, least-squares fits of the data to realistic beam patterns.

The method has the further advantage that source-size corrections, which presently represent a significant source of error in both gain measurements and source flux-density determinations, are not needed since essentially all of the source flux density is collected during the raster scan.

The use of rapid scanning also results in the collection of vastly more data than with conventional techniques, so that errors in gain, or aperture efficiency versus elevation curves can be greatly reduced.

Measurements at Ku-band and Ka-band, based on 2-D and 3-D fitting are in good agreement with theoretical calculations using measured power spectral density data to predict the background noise during a scan, and using this background noise as input to a nonlinear, least-squares model to predict fitting parameter errors.

This application of the OTF-mapping R&D system for the Cassini-JMOC calibration work demonstrates some of the valuable attributes of the system for the calibration and performance analysis of the DSN antennas for telemetry and for radio science.

The OTF-mapping R&D system has been developed into an operational antenna calibration and measurement equipment (ACME) system. ACME initial delivery is the first step in an ambitious project to provide the DSN with a standard calibration tool. This is the first attempt to use a unified evaluation criterion, allowing larger quantities of data to be collected and improving its quality. It will allow the technical community to know the state and evolution of all antennas via a common database. Set up of the tool is quick and operationally safe. Calibration times depend on the type of measurements, but experience gathered so far indicates that it will easily be made compatible with DSN routine maintenance and operations.

As the system is used by the calibration engineers at the stations, more feedback is expected and more and better improvements will be added to the future work.

References

[1] J. Bautista, R. Clauss, S. Petty, and J. Shell, "DSN Low Noise Amplifiers In the New Millennium," *JPL-TMOD Technology and Science Program News*, vol. 82, no. 5, January 2001.

[2] L. W. M. Baars, R. Genzel, I. I. K. Pauliny-Toth, and A. Witzel, "The Absolute Spectrum of CasA: An Accurate Flux Density Scale and a Set of Secondary Calibrators," *Astronomy and Astrophysics*, vol. 61, pp. 99–106, 1977.

[3] M. J. Klein and C. T. Stelzried, "Calibration Radio Sources for Radio Astronomy: Precision Flux Density Measurements at 2295 M Hz," *the Astronomical Journal*, vol. 81, no. 12, pp. 1078–1083, 1976.

[4] P. H. Richter and S. D. Slobin, "DSN 70-Meter Antenna X- and S-Band Calibration Part I: Gain Measurements," *The Telecommunications and Data Acquisition Progress Report 42-97*, *January–March 1997*, pp. 315–351, Jet Propulsion Laboratory, Pasadena, California, May 15, 1989. http://ipnpr.jpl.nasa.gov/progress_report/

[5] A. J. Freiley, P. D. Batelaan, and D. A. Bathker, *Absolute Flux Density Calibrations of Radio Sources at 2.3 GHz*, JPL Technical Memorandum 33-806, Jet Propulsion Laboratory, Pasadena, California, December 1, 1977.

[6] J. D. Kraus, *Radio Astronomy*, McGraw-Hill, New York, New York, pp. 97–101, 1966.

[7] P. Richter, *Radio Source List for Antenna Calibration*, JPL D-3801, DSN No. 890-269 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 15, 1994.

[8] B. L. Ulich and R. W. Haas, "Absolute Calibration of Millimeter-Wavelength Spectral Lines," *Astrophysical Journal Supplement Series*, vol. 30, pp. 247–258, 1976.

[9] M. Adler, *Cassini Project Policies and Documents*, JPL D-9945, Rev. D (internal document), Jet Propulsion Laboratory, Pasadena, California, 1995.

[10] P. Richter, "Estimating Errors in Least-Squares Fitting," *The Telecommunications and Data Acquisition Progress Report 42-122, April–June 1995*, Jet Propulsion Laboratory, Pasadena, California, pp. 107–137, August 15, 1995. http://ipnpr.jpl.nasa.gov/progress_report/

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

[12] D. Rochblatt, P. Richter, and P. Withington, "On-the-Fly Mapping for Calibrating Directional Antennas," *NASA Tech Briefs* (NPO-30648), vol. 28, no. 8, pp. 53–55, August 2004.

[13] P. W. Gorham and D. J. Rochblatt, "Effect of Antenna-Pointing Errors on Phase Stability and Interferometric Delay," *The Telecommunications and Data Acquisition Progress Report 42-132, October–December 1981*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–19, February 15, 1998. http://ipnpr.jpl.nasa.gov/progress_report/

[14] D. Rochblatt and P. Withington, "Precision Blind Pointing Calibration of the NASA-JPL-DSN Large Reflector Antennas at Ka-Band (32-GHz)," *EuCAP 2006 – European Conference on Antennas & Propagation*, Session 4A10A – Antenna Measurements (12j), November 9, 2006.

[15] *1st DSN Antenna Calibration Workshop*, JPL D-29992 (internal document), Jet Propulsion Laboratory, Pasadena, California, July 6–9, 2004.

[16] S. J. Bolton, M. Janssen, R. Thorne, S. Levin, M. Klein, S. Gulkis, T. Bastian, R. Sault, C. Elachi, M. Hofstadter, A. Bunker, G. Dulik, E. Gudim, G. Hamilton, W. T. K. Johnson, Y. Leblanc, O. Liepack, R. McLeod, J. Roller, L. Roth, and R. West, "Ultra-Relativistic Electrons in Jupiter's Radiation Belts," *Nature*, vol. 415, pp. 987–991 February 28, 2002.

[17] D. J. Rochblatt and B. L. Seidel, "Performance Improvement of DSS-13 34-Meter Beam-Waveguide Antenna Using the JPL Microwave Holography Methodology," *The Telecommunications and Data Acquisition Progress Report 42-108, October–December 1991*, Jet Propulsion Laboratory, Pasadena, California, pp. 253–270, February 15, 1992. http://ipnpr.jpl.nasa.gov/progress_report/

[18] M. J. Klein, S. J. Bolton, A. J. Freiley, S. Gulkis, M. A. Janssen, S. J. Levin, D. J. Rochblatt, J. P. Roller, and R. K. McLeod, "DSN and GAVRT Observations of Jupiter at 13 GHz and the Calibration of the Cassini Radar Instrument for Passive Radiometry," *URSI General Assembly*, Naastrichthe, Netherlands; August 17–24, 2002.

[19] C. Stelzried, "Noise Adding Radiometer Performance Analysis," *The Telecommunications and Data Acquisition Progress Report 42-59, July and August 1980*, Jet Propulsion Laboratory, Pasadena, California, pp. 98–106, October 15, 1980. http://ipnpr.jpl.nasa.gov/progress_report/

[20] C. Stelzried, "Correction of High-Frequency Noise-Temperature Inaccuracies," *The Telecommunications and Data Acquisition Progress Report 42-111, July–September 1992*, Jet Propulsion Laboratory, Pasadena, California, pp. 269–277, November 15, 1992. http://ipnpr.jpl.nasa.gov/progress_report/

[21] C. T. Stelzried, "Noise Temperature and Noise Figure Concepts: DC to Light," *The Telecommunications and Data Acquisition Progress Report 42-67, November and December 1981*, Jet Propulsion Laboratory, Pasadena, California, pp. 100–111, February 15, 1982. http://ipnpr.jpl.nasa.gov/progress_report/

[22] R. Levy, *Structural Engineering of Microwave Antennas*, IEEE Press: Hackensack, New Jersey, 1996.

[23] P. Stumpff, translation of "Astronomische Pointing Theorid Fuer Radioteleskope," *Klein Heubacher Berichte*, vol. 15, Fornmolde Technischon Zentralamt, Darmstadt, Germany, pp. 432–437, 1972.

[24] D. J. Rochblatt, "Holographic Measurements of the NASA-JPL Deep Space Network Antennas," *Proceedings of the 1998 IEEE Aerospace Conference*, Vol. 3, Snowmass at Aspen, Colorado, pp. 441–452, March 21–28, 1988.

[25] D. Rochblatt, P. Richter, P. Withington, M. Vasquez, and J. Calvo, "New Antenna Calibration Techniques in the Deep Space Network," *The Interplanetary Network Progress Report* vol. 42-169, pp. 1–34, May 15, 2007. http://ipnpr.jpl.nasa.gov/progress_report/

Chapter 8
Microwave Antenna Holography

David J. Rochblatt

8.1 Introduction

The National Aeronautics and Space Administration (NASA)–Jet Propulsion Laboratory (JPL) Deep Space Network (DSN) of large reflector antennas is subject to continuous demands for improved signal reception sensitivity, as well as increased transmitting power, dynamic range, navigational accuracy, and frequency stability. In addition, once-in-a-lifetime science opportunities have increased requirements on the DSN performance reliability, while needs for reduction of operational costs and increased automation have created more demands for the development of user friendly instruments. The increase in the antenna operational frequencies to X-band (8.45 gigahertz (GHz)) and Ka-band (32 GHz), for both telemetry and radio science, proportionately increased the requirements of the antenna calibration accuracy and precision. These include the root-mean-square (rms) of the main reflector surface, subreflector alignment, pointing, and amplitude and phase stability. As an example, for an adequate performance of an antenna at a given frequency, it is required that the reflector surface rms accuracy be approximately $\lambda/20$ (0.46 millimeter (mm) at Ka-band) and that the mean radial error (MRE) pointing accuracy be approximately $\lambda/(10*D)$, or a tenth of the beamwidth (1.6 millidegrees (mdeg) for a 34-meter (m) antenna at Ka-band).

Antenna microwave holography has been used to improve DSN performance. Microwave holography, as applied to reflector antennas, is a technique that utilizes the Fourier transform relation between the complex far-field radiation pattern of an antenna and the complex aperture distribution. Resulting aperture phase and amplitude-distribution data are used to precisely

characterize various crucial performance parameters, including panel alignment, subreflector position, antenna aperture illumination, directivity at various frequencies, and gravity deformation effects. The holography technique provides a methodology for analysis, evaluation, and radio-frequency (RF) performance improvement of large reflector and beam waveguide antennas. Strong continuous-wave (CW) signals can be obtained from geostationary satellites and used as far-field sources. Microwave holography has been one of the most economical techniques for increasing the performance of the large DSN antennas in terms of cost-to-performance ratio. This chapter describes the instrument design and the mathematical algorithms and software for the development of the holographic measurement system. In addition, it describes its application in the DSN to improve, optimize, and maintain its performance to prescribed specifications.

The word "holography" is derived from the Greek "holos," which means "whole." Therefore, a hologram is created when the whole information can be recorded and presented graphically. In the antenna engineering case, the whole information is the amplitude, frequency, and phase of the signal or the transfer function of the antenna. We know that in optics, when a hologram is created we can see the depth of the image, which is absent in regular camera images. The reason is that a camera records only the intensity (square amplitude) of light at a given frequency (for which the film is sensitive). In holographic recording, the phase is recorded in addition to the amplitude at a given frequency range. It is the recording of phase that contains the depth, or the third dimension, that gives it a life-like perception. Good painters know how to play with light and shadow to create the "feel" of depth in their painting. It is the recording of the phase in the antenna holography that enables us to derive the misalignment of the antenna panels in the direction perpendicular to the x-y plane or the z-axis.

To obtain a hologram of the entire antenna dish surface, a two-dimensional sampling of the antenna far-field pattern must be recorded. Holographic recording acquires the phase and amplitude information utilizing raster-scan patterns of the antenna angular response. (Note: raster scans are the most popular, although other scan geometries are possible and could be advantageous under certain conditions.) Holographic metrology is based on interferometrically connecting a reference antenna to the large test antenna and digitally recording the test antenna amplitude and phase response. This is done by continuously scanning the test antenna against a signal source from a geosynchronous satellite, following a two-dimensional grid (Fig. 8-1).

Celestial radio sources can also be used but require a different receiver architecture. Their usual lower signal strength imposes limitations and introduces additional complexity to the measurement and the data processing. Their sidereal motion across the sky would require a faster data acquisition to avoid smearing of the resulting surface map.

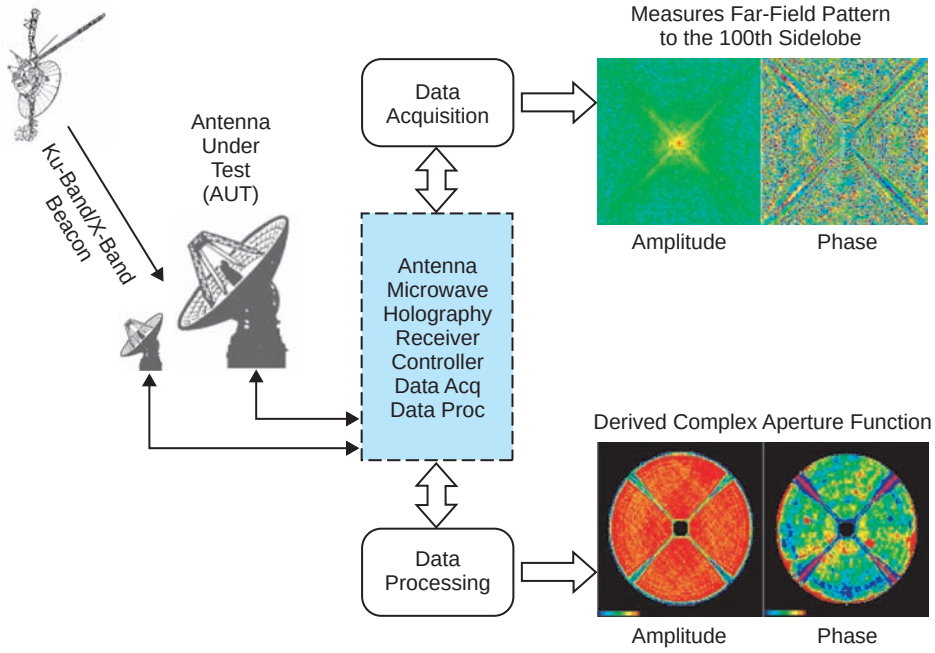


FIg. 8-1. Antenna microwave holography activities diagram.

When a regularized far-field grid can be measured, an inverse fast Fourier transform (FFT) algorithm can then be used to obtain the desired information, consisting of the test antenna aperture amplitude and phase response [1–6]. Other irregularized grids [7] can also be used for the measurements followed by other inversion techniques than the FFT (as an example, a singular value decomposition). The angular extent of the response that must be acquired is inversely proportional to the size of the desired resolution cell in the processed holographic maps. From the aperture phase response, the surface error map is calculated, and the amplitude response is directly displayed. The information in the surface error map is used to calculate the adjustments of the individual panels in an overall main reflector best-fit reference frame. The amplitude map provides valuable information about the energy distribution in the antenna aperture (Fig. 8-1).

The ultimate performance of a large, steerable, reflector antenna is limited by imperfections of the reflecting surface. The size of the panels that form the surface of the antenna and the allowable level of losses due to surface inaccuracies dictate the required resolution of the measurements. For a maximum of 0.1-decibel (dB) degradation in antenna efficiency due to surface imperfections, the rms surface error (ε) must be no greater than $0.012\,\lambda$ where λ is the wavelength of the operating antenna frequency. When the surface error

is $0.024\,\lambda$, degradation in antenna efficiency is 0.4 dB, which demonstrates the exponential relationship between surface error and gain loss.

For $\varepsilon/\lambda < 1/4\pi$, this relationship (also known as the Ruze formula) can be expressed [8]:

$$\frac{\eta_A}{\eta_0} = \exp-\left(\frac{4\pi\varepsilon}{\lambda}\right)^2 + \left(\frac{2r_o}{D}\right)\left[1-\exp-\left(\frac{4\pi\varepsilon}{\lambda}\right)^2\right] \qquad (8.1\text{-}1)$$

where,

η_A = efficiency of the physical antenna

η_0 = efficiency of a hypothetical antenna with absence of surface errors

ε = rms of surface deviation in the axial direction

λ = wavelength

r_o = correlation radius

D = antenna diameter

For totally random surface phase errors, $r_o = 0$, and the Ruze formula, Eq. (8.1-1) reduces to its first term. When r_o is significant relative to D, the second term in Eq. (8.1-1) adds to the first term to yield a higher efficiency value than in the case where the errors are totally random. Therefore, using the first term in Eq. (8.1-1) yields the worst-case value for a given surface rms value.

Figure 8-2 shows the gain loss of a reflector antenna (of any size) as a function of its rms surface error using Eq. (8.1-1) and assuming $r_o = 0$. As can be seen from this plot, reducing the reflector effective rms error from 0.67-mm to 0.25-mm will result in an antenna gain increase of 3.0 dB at Ka-Band (32 GHz). In the DSN, this indeed has been the case as most of the 34-m Beam

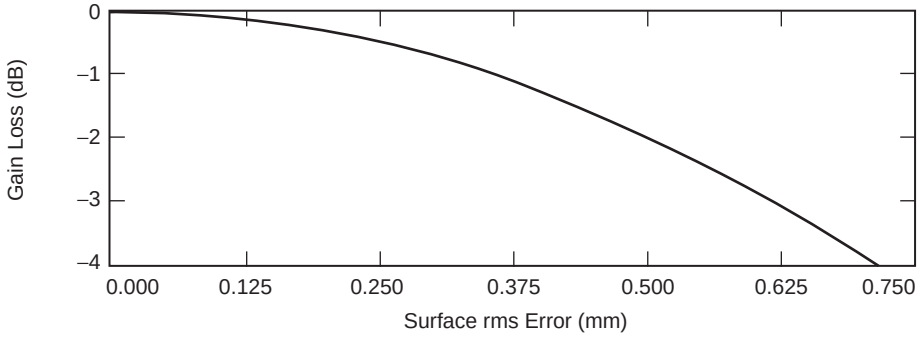


Fig. 8-2. Antenna gain loss versus rms surface error at 32 GHz.

Waveguide (BWG) antenna's panels were initially set by a theodolite technique and later refined by the holographic technique that improved their performances by 3 dB (on average).

Antenna microwave holography implementation typically has to meet several requirements. It must satisfy the requirements for a fast (45-minute) "health check" measurement, utilizing low-resolution medium-precision images to determine the antenna status. It also must provide high-resolution, high-precision images measured over a period of no longer than 12 hours to avoid thermal diurnal effects.

8.2 Holography System Simulation

Two approaches for the system architecture are typically used. These are based on a wide bandwidth or a narrow bandwidth receiver architecture. To facilitate the examination of either of the two approaches, we developed simulation algorithms.

The parameters critical for the quality of the images derived from holographic measurements are signal-to-noise ratio, maximum scan angle, instrumentation dynamic range, related approximations (may be included due to different sampling techniques) and overall system accuracy. A detailed mathematical derivation of the related equations can be found in [9,10]. In general, to derive the standard deviation in the final holographic map from simulation, we first compute the far-field pattern of the perfect reflector antenna. A simulation tool is developed by superimposing the contribution of the measurement system noise on the far-field patterns [9]. By processing the new far-field data and displaying the images, one can derive the standard deviation error in processed holographic maps. The NASA-DSN-JPL 64-m antennas prior to their upgrade to 70-m diameter [11] were used for the simulations case study. These three Cassegrain antennas (located at Goldstone, California; Robledo de Chavela, Spain; and Tidbinbilla, Australia) were designed with –13 dB amplitude aperture taper illumination. Prior to the upgrade of the three antennas to 70-m, these antennas had on average an rms surface error of 1.34-mm (details for each antenna are provided in Table 8-1). In the post 70-m upgrade, these antennas were all set holographically [12] to an average value of 0.65 mm, which improved their performance at X-band by approximately 0.75 dB.

Table 8-1. Holography historical data.

Antenna Diameter (m)	DSS*	Date	Meas. Freq. (GHz)	Resolution (m)	Elevation (deg)	Initial rms (mm)	Final rms (mm)	Gain Improvement (dB) Frequency Band S	X	Ka
70-m Antennas										
70	14	4/88	12.198	0.42	47.0	**1.26**	**0.64**	0.05	0.64	9.3
70	43	10/87	12.750	0.44	47.0	**1.18**	**0.65**	0.04	0.59	8.5
70	63	7/87	11.451	0.42	42.0	**1.58**	**0.65**	0.09	1.17	16.9
DSS-13 34-m R&D Antenna										
34	13	9/90	12.198	0.32	46.0	**0.88**	**0.43**	0.02	0.32	4.6
34	13	1/92	12.198	0.32	46.0	**0.68**	**0.37**	0.01	0.18	2.5
34	13	2/94	12.198	0.32	46.0	**0.38**	**0.31**	0.002	0.03	0.32
34-m Operational Antennas										
34	24	5/94	11.922	0.33	46.3	**0.50**	**0.25**	0.007	0.1	1.27
34	25	6/96	11.913	0.33	47.0	**0.50**	**0.25**	0.007	0.1	1.27
34	26	10/96	11.913	0.33	47.0	**0.42**	**0.25**	0.004	0.05	0.76
34	54	5/98	12.502	0.32	43.2	**0.79**	**0.32**	0.02	0.25	4.0
34	34	6/98	12.748	0.315	48.3	**0.47**	**0.26**	0.006	0.08	1.2
34	55	7/03	11.450	0.33	43.1	**0.90**	**0.25**	0.03	0.41	5.8

* DSS = Deep Space Station (antenna's designation in the DSN):
DSS-14, 43, and 63 are the 70-m antennas
DSS-13 is the 34-m research and development (R&D) BWG antenna
DSS-24, 25, 26, 34, 54, and 55 are the 34-m BWG antennas

The accuracy in the final holographic maps and the resolution in of the images are interrelated. We formulated the accuracy from the simulation results to be

$$\sigma \cong 0.082 \frac{\lambda D}{\delta \mathrm{SNR}} \tag{8.2-1}$$

where

σ = standard deviation (accuracy) in r ecovering the mean position of a resolution cell

λ = wavelength

D = reflector diameter

δ = spatial resolution in the aperture plane (defined below)

SNR = beam peak voltage signal-to-noise ratio (SNR) in the test (antenna) channel.

Equation (8.2-1) agrees well with the analytical expressions derived in [2]. Here, the constant 0.082 was empirically determined based on the simulation results, which agree well with the analytically derived constant of $1/4\pi$. As will be shown in the simulation results, the accuracy across holographic maps varies with the aperture amplitude taper illumination. Results are better at the center of the dish and gradually become worse toward the edge of the dish. For a uniformly illuminated dish, accuracy stays relatively constant through most of the dish and quickly becomes worse just at the edge where the illumination falls off rapidly.

To define the angular resolution δ in the processed holographic maps, consider a square grid containing N^2 sampled data points separated by less than λ / D or one antenna beamwidth. Let

$$\Delta u = \Delta v = \frac{\mathrm{k}\lambda}{D} \tag{8.2-2}$$

where

$\Delta u, \Delta v$ = separation between two adjacent points in two orthogonal axes

k = a constant $0.5 < k < 1.0$

The length of this grid (L) is then:

$$L = N\mathrm{k}\left(\frac{\lambda}{D}\right) \tag{8.2-3}$$

Consider the Fourier transform of a rectangular pulse extended from $+(N/2)(\mathrm{k}\lambda D)$ to $-(N/2)(\mathrm{k}\lambda D)$. The function transforms from a pulse to:

$$\frac{\sin\left(\dfrac{\pi N\,\mathrm{k}\,\lambda x}{D}\right)}{\dfrac{\pi N\,\mathrm{k}\,\lambda x}{D}} \tag{8.2-4}$$

The two nulls of this function occur at $+D/\mathrm{k}\lambda N$ for a full null width of $2D/\mathrm{k}\lambda N$. We now define the spatial resolution to be at the 50-percent width and obtain:

$$\delta = \frac{D}{\mathrm{k}\,N} \tag{8.2-5}$$

where

δ = spatial resolution in the aperture plane

Aperture simulation models are used because of their simplicity and usefulness. They allow an examination of the interrelations between the standard deviation in the holographic maps and a known feature on the reflector surface. In these models, aperture and phase distribution are typically defined, and then far-field data are constructed. In general, an integration or FFT scheme may be used to obtain the far-field data. However, for certain special aperture distributions, such as those that are circularly symmetric, *closed-form* expressions can be used [13]. This allows an accurate and efficient far-field pattern generation. The steps of this aperture model follow.

The geometry of a circular aperture, with different annular regions designated by red and blue colors, is shown in Fig. 8-3. The green color represents a perfect dish surface relative to the best-fit paraboloid, while it is assumed that the red and blue colors represent regions that are deformed by a constant value of ±0.2 mm, respectively, causing constant phase irregularities and resulting in a main reflector surface error of 0.11-mm root-mean-square (rms). (Please note that the scale in Fig. 8-3 is ±0.35 mm). We further assume that the amplitude and phase distributions across the aperture are circularly

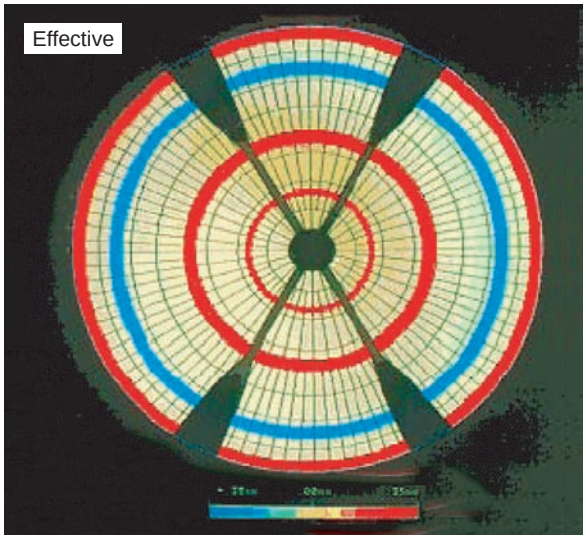


Fig. 8-3. Simulation I: geometry of simulated reflector distortions; no noise was injected.

symmetric, closed-form functions. These assumptions allow us to express the far-field integral in terms of a one-dimensional integral.

Furthermore, for an appropriately chosen amplitude distribution, this integral can be integrated in a closed form. This closed-form expression can be used to construct the far field pattern.

Once the far-field amplitude and phase data are generated, they can then be used in the error simulation algorithm by appropriately injecting noise in a manner that simulates the exact architecture of the holographic measurement system and its front-end thermal noise [9].

A narrow-bandwidth system can be designed with a wide dynamic range and linear response. Such a system will make use of geostationary satellite beacon signals (nearly CW) available on nearly all satellites at Ku-band (10.8 to 12.9 GHz), X-band (7.7 GHz), S-band (2.2 GHz), and on other bands as well. The antenna microwave holography receiver block diagram is shown in Fig. 8-4.

The back-end receiver uses synchronous detectors for the in-phase (I) and quadrature-phase (Q) components of the test and reference channels. The analog signal is digitized utilizing a 19-bit resolution analog-to-digital (A/D) converter, to form the ratio (rather than multiplication) of the test-to-reference channel signals. This provides the real and imaginary components of the complex far-field function. Amplitude variations in the satellite signal cancel out in the division operation. This feature is especially critical since no control over the satellite signal power level is available. Also, since the reference channel SNR in this scheme can easily be 40 dB or better, it can be safely used in the denominator. (This would not be desirable for weak reference signals.)

The antenna microwave holography provides a linear dynamic range of better than 96 dB down to integration periods of 0.2 millisecond (ms). When the satellite beacon effective isotropic radiated power (EIRP) is about 11 dB referenced to watts (dBW), a beam peak SNR of 73 dB is achieved on the 70-m antenna at Ku-band (12 GHz) with a 0.1-s integration period using a simple room-temperature field effect transistor (FET) (100 K) amplifier, while a 2.8-m reference dish provides 40–45 dB in SNR, using a room-temperature (100 K) FET.

For a multiplier integrator as well as a divider integrator receiver architecture, the effective signal SNR can be expressed as

$$\mathrm{SNR}_E = \left[\sqrt{\frac{1}{\mathrm{SNR}_T^2} + \frac{1}{\mathrm{SNR}_R^2} + \frac{1}{\mathrm{SNR}_T^2\,\mathrm{SNR}_R^2}} \right]^{-1} \tag{8.2-6}$$

where SNR_T and SNR_R are the test channel and reference channel SNR, respectively. The generality of this formulation makes it useful for many different receiver architectures.

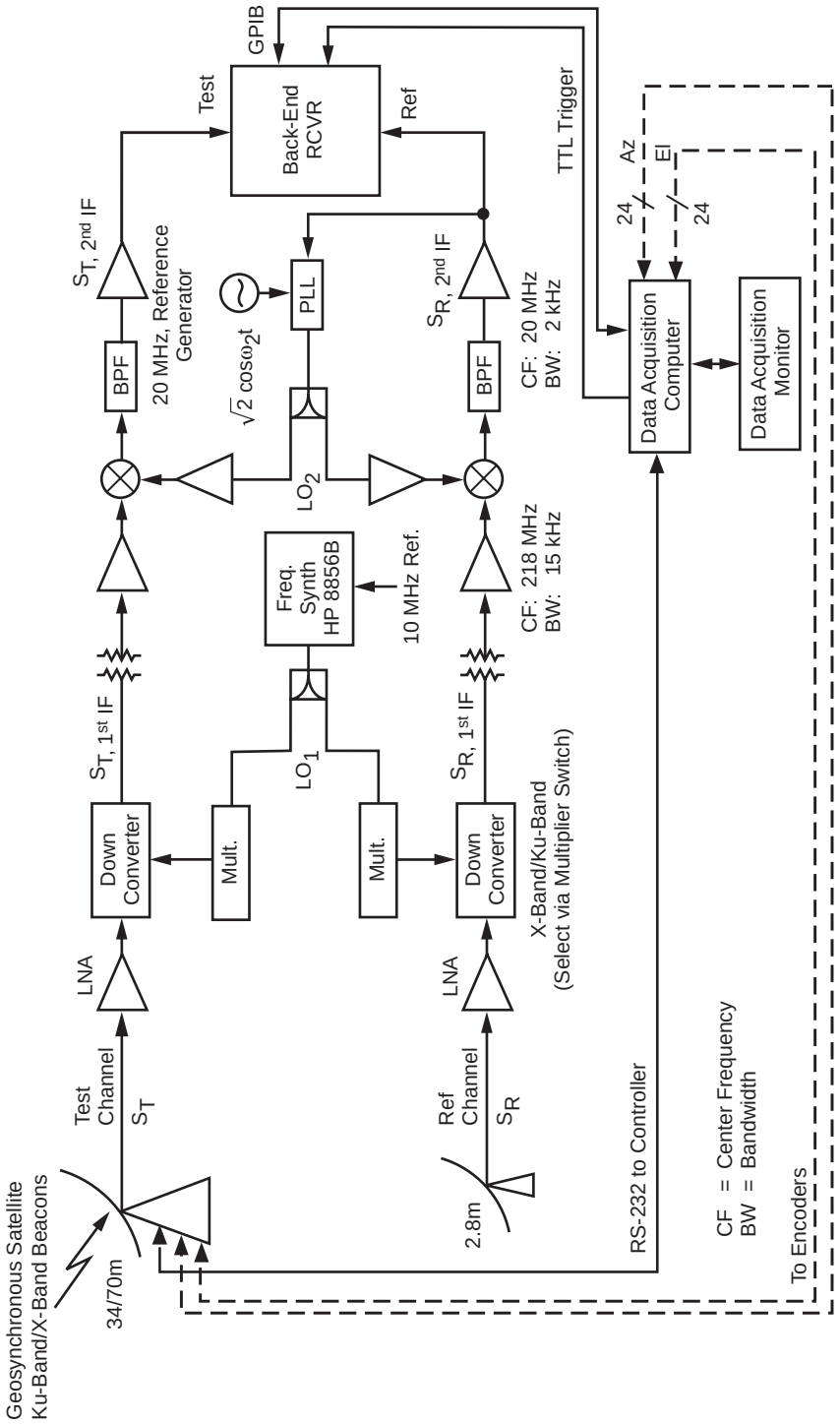


Fig. 8-4. Antenna microwave holography block diagram (notations defined in Section 8.3).

From Eq. (8.2-6), it is a pparent that the effective SNR_E is dominated by the weaker of the two channels. What this means is that the beam peak SNR_T (of 73 dB) is not realized, and the first few data points on beam peak and a few sidelobes have an effective SNR_E (of approximately 45 dB) of the reference SNR_R. Once the test channel SNR_T drops below the reference antenna SNR_R (45 dB), it doe s degrade the effective SNR_E, which, from that point on, follows the same function as the test antenna beam patterns (SNR_T). This is acceptable since very few data points are affected (approximately 0.5 percent), and since by the nature of the data processing through the Fourier transform operation, all the data points in the far field contribute to each and every point in the aperture, as is shown in the simulations below.

A simulation algorithm was developed to derive the relationships between the standard deviation in the final holographic maps and the measurement SNR. The receiver back-end architecture, which comprises I and Q separation of both the test and reference channels, has been modeled in the simulation [9]. This was done by a dding independent noise-processing components n, for which the 1σ in the random Gaussian function, for the test and reference signals, respectively, is

$$1\sigma_T = \frac{\mathrm{amp}_T(\max)}{\mathrm{SNR}_T} \tag{8.2-7}$$

where $\mathrm{amp}_T(\max)$ = beam peak amplitude in the test channel

$$1\sigma_R = \frac{\mathrm{amp}_R(\max)}{\mathrm{SNR}_R} \tag{8.2-8}$$

where $\mathrm{amp}_R(\max)$ = beam peak amplitude in the reference channel.

The function of the receiver described in Fig. 8-4 was used in the simulation to provide the resultant measured complex quantity, including noise:

$$\text{complex field} = \frac{\mathrm{amp}_T(\theta_i)e^{j\phi_{Ti}} + n_{R_i}^T + jn_{I_i}^T}{\mathrm{amp}_R\, e^{j0} + n_{R_i}^R + jn_{I_i}^R} \tag{8.2-9}$$

where

$\mathrm{amp}_T(\theta_i)$ = test antenna far-field amplitude voltage at the sampled data position θ_i.

ϕ_{T_i} = test antenna far-field phase at the sampled data position *i*.

$n_{R_i}^{T}$ = noise component in the complex real part of the digitized data sample *i* in the test channel.

$n_{I_i}^{T}$ = noise component in the complex imaginary part of the digitized data sample *i* in the test channel.

$n_{R_i}^{R}$ = noise component in the complex real part of the digitized data sample *i* in the reference channel.

$n_{I_i}^{R}$ = noise component in the complex imaginary part of the digitized data sample *i* in the reference channel.

$\text{amp}_R\, e^{j0}$ = reference channel far-field constant amplitude and phase value.

This simulation (Figs. 8-3, 8-5, 8-6, and 8-7) examined the effect of the SNR in the reference and test antennas on measurement accuracy.

In the simulation, four rings of panels were intentionally displaced by 0.2 mm ($\lambda/130$ at 11.45 GHz, Fig. 8-3). Three rings were displaced positively, and one was displaced negatively. The width of the three outmost rings was 2.0 m (76 λ), and the innermost rings was 1.0 m wide. The rms surface error of this model (Fig. 8-3) is 0.11 mm. The far-field for the above reflector geometry was generated and then contaminated with noise due to the front end, according to the model represented by Eq. (8.2-4). The far-field data were then processed to display the recovered surface error maps and to compute the surface rms errors.

In Fig. 8-3, Simulation I, the far-field was processed with no noise added to it. This simulated an SNR of more than 90 dB. The computer computational errors are at a level of about $\lambda/5000$ (11.45 GHz). By subtracting (map differencing) this model from subsequent simulations, we obtained a measure comparable to the measurement system standard deviation. Simulation II, (Fig. 8-5) models the conditions where the test antenna SNR on beam peak during the 0.1-s integration period is 73 dB, and the reference antenna constant SNR is 40 dB. The recovered rms of the test antenna surface is 0.12 mm, and the measurement system standard deviation is 0.07 mm ($\lambda/370$ at 11.45 GHz).

Simulation III, (Fig. 8-6) simulates conditions in which the test antenna beam peak SNR in the 0.1-s integration period is 68 dB and the reference antenna SNR is 40 dB. The recovered surface rms is 0.16 mm with a standard deviation of 0.13 mm.

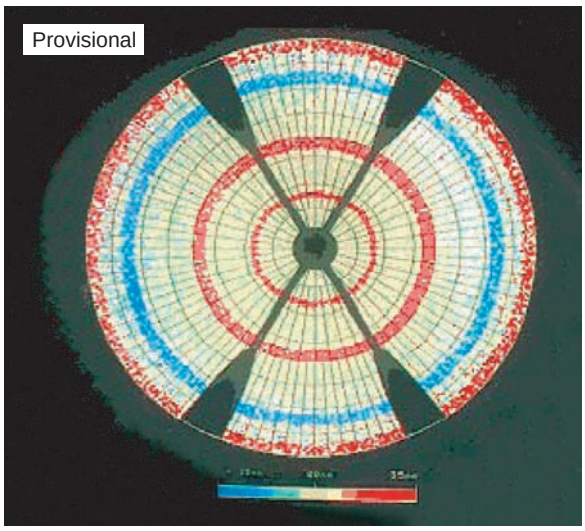


Fig. 8-5. Simulation II: test antenna SNR on beam peak in the 0.1-s integration period is 73 dB, and the reference antenna constant SNR is 40 dB.

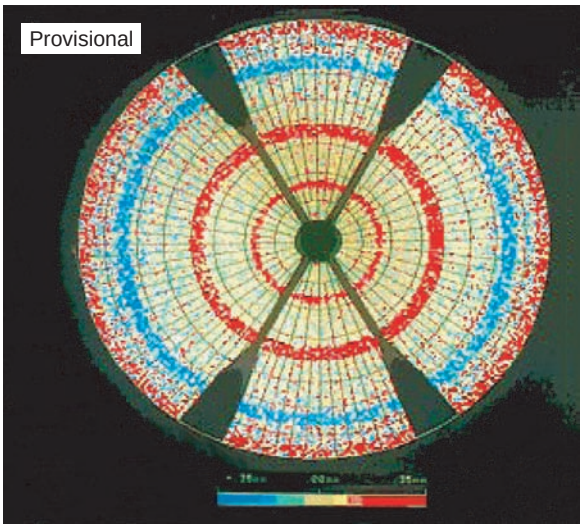


Fig. 8-6. Simulation III: test antenna beam peak SNR in the 0.1-s integration period is 68 dB, and the reference antenna constant SNR is 40 dB.

From simulation IV (Fig. 8-7), it is c lear that the recovery of the dish surface error is very poor when the SNR drops to 58 dB. The recovered surface

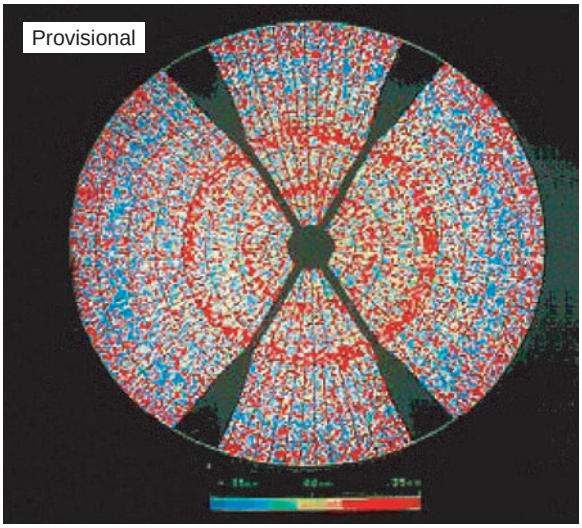


Fig. 8-7. Simulation IV: SNR dropped to 58 dB.

rms is 0. 43 mm, and the standard deviation of this map is 0. 41 mm. The necessity of a high-beam peak SNR for high-resolution, high-precision holographic measurement is clearly demonstrated [9].

8.3 Holography Receiver Signal Analysis

The MAHST design features a dual channel coherent CW receiver where the reference signal is pr ovided by a small (2.8-m diameter dish) reference antenna, which is bor e-sighted at a geostationary satellite while the antenna under test (AUT) is acquiring the signal while performing a continuous raster-scan relative to the moving spacecraft (see Figs. 7-4 and 8-1).

Referring to F ig. 8-4, the microwave signal at the test antenna (upper channel) may be modeled as

$$S_T = \sqrt{2P_T(t)}\sin(\omega_o t + \theta + \varphi) \tag{8.3-1}$$

where

S_T = test signal

$P_T(t)$ = nominal power of the received signal at the test channel

ω_o = nominal (angular) frequency ($2\pi f_o$) of the received microwave test signal

θ = the function account for Doppler effect

φ = phase

The departure of the instantaneous frequency from its nominal value of ω_o is accounted for by the time derivative of θ. The purpose of the receiver is the measurement of the test antenna amplitude and phase represented by square root of $2P_T(t)$ and φ.

The microwave signal at the reference antenna is modeled as

$$S_R = \sqrt{2P_R(t)}\sin(\omega_o t + \theta) \tag{8.3-2}$$

where:

S_R = reference signal

$P_R(t)$ = nominal power of the received signal in t he reference channel

The reference channel serves as a phase reference for the phase measurements, as well as for tracking out the Doppler effects introduced by the movement's drifts of the spacecraft. The one-sided noise spectral density N_{OT} of the receiving channels equals the equivalent noise temperature multiplied by Boltzmann's constant. The equivalent noise temperature in the two channels must be referenced to the same point in the receiving chain (for example, the input of the low-noise amplifier).

The frequency of the local oscillator (LO) in the first down-converter mixer in the receiver front end is selectable via three multipliers to cover the entire Ku-band frequency range of 10.8 to 12.8 GHz, as well as X-Band (7.7 GHz). The first LO (LO_1) is common to both the test and the reference channels for phase coherent detection, and it can be modeled as

$$\mathrm{LO}_1 = 2\cos\left[\left(\omega_o - \omega_1\right)t\right] \tag{8.3-3}$$

where ω_1 is the nominal (angular) frequency of the signal in the first intermediate-frequency (IF) stage. The resultant signals output of the first down conversion stages become:

$$S_{T,1\mathrm{IF}} = \sqrt{2P_T(t)G_T}\,\sin\left(\omega_1 t + \theta + \varphi + \xi\right) \tag{8.3-4}$$

and

$$S_{R,1\mathrm{IF}} = \sqrt{2P_R(t)G_R}\,\sin\left(\omega_1 t + \theta\right) \tag{8.3-5}$$

where ξ is the differential phase delay between the reference and test channels. G_T and G_R are the power gains in the test and reference channels,

respectively. The one-sided noise spectral densities within the first IFs are $N_{OT}G_T$ and $N_{OR}G_R$.

The second LO (LO_2) is derived from the output of a phase-locked loop (PLL) that tracks the reference channel signal Doppler effects. The PLL itself is tied into the Frequency and Timing Subsystem (FTS) station standard stable oscillator.

$$LO_2 = 2\cos\left[\left(\omega_1 - \omega_2\right)t + \theta - \phi\right] \tag{8.3-6}$$

The PLL output is used to further down convert the first IF signals in the test and reference channels. When the PLL tracks perfectly, $\phi = 0$. The test and reference signals in the second IF stages are then given by:

$$S_{T,\,2\,\mathrm{IF}} = \sqrt{2P_T(t)G_T}\,\sin\left(\omega_2 t + \varphi + \xi + \phi\right) \tag{8.3-7}$$

and

$$S_{R,\,2\,\mathrm{IF}} = \sqrt{2P_R(t)G_R}\,\sin\left(\omega_2 t + \phi\right) \tag{8.3-8}$$

where ω_2 is the nominal (angular) frequency of the second IF and ϕ is the phase-tracking error in the PLL.

The gains of the channels between the first and second IFs are incorporated into G_T and G_R. The differential phase delay between the reference and test channels that occurs between the first and second IFs is incorporated into ξ. The one-sided noise spectral densities within the second IFs are $N_{OT}G_T$ and $N_{OR}G_R$.

The phase transfer function of the PLL is given by:

$$H(s) = \frac{KF(s)}{s + KF(s)} \tag{8.3-9}$$

where

$H(s)$ = Laplace transform
K = cumulative loop gain
$F(s)$ = transfer function of the loop filter

The noise-equivalent bandwidth B of this phase transfer functions is:

$$B = \int_0^\infty |H(j2\pi f)|^2\, df \tag{8.3-10}$$

and the loop phase error variance is:

$$\sigma_\phi^2 = \frac{N_{OR}B}{P_R} \tag{8.3-11}$$

The back-end portion of the receiver measures the amplitude and phase of the signals in the test channel second IF relative to the amplitude and phase of the signal in the reference channel second IF. It is this relative amplitude and phase that is required for holography measurements.

The test channel signal is given by Eq. (8.3-7) and that of the reference channel by Eq. (8.3-8). The frequency of these two signals is stable because the PLL has removed the time-varying Doppler effect. It might seem that the gains G_T and G_R and the differential phase delay ξ obscure the parameters of interest. For the purpose of holography, however, it is only necessary to measure how the relative amplitude and phase change with time. As long as G_T, G_R, and ξ remain approximately constant during the course of the observation. The receiver back-end works as follows. The test and reference channel signals (at approximately 20 MHz) are further downconverted to 100 kHz. The signals are then subjected to automatic gain control (AGC). In each channel, there are in-phase and quadrature detectors followed by analog-to-digital (A/D) converters. The amplitude and phase of the test signal relative to the reference signal are computed as described by Eq. (8.2-9). The AGC removes much of the amplitude variation from the signals. This is not a problem because the variations applied to each channel are recorded. Recorded AGC gain represents a coarse measure of the signal amplitude. These recorded AGC values are then entered into the final calculation of the relative amplitude.

The relative phase $\varphi + \xi$ is measured. As mentioned above, as long as the instrumental delay ξ is approximately constant during the observation; the variation of φ is reflected in the measured result. The loop phase error ϕ is not present because it is a common-mode error in the test and reference signals. (However, it is still important to keep the loop phase error variance Eq. (8.3-11) small in order to minimize cycle slips in the loop.) The variance in the relative phase measurement due to receiver noise is given by

$$\frac{N_{OT}}{2P_T\,\mathrm{T}} + \frac{N_{OR}}{2P_R\,\mathrm{T}}\ \mathrm{rad}^2 \tag{8.3-12}$$

where T is the integration time for each measured phase.

By the virtue of the reference antenna continuous boresight on the spacecraft signal source, under ideal conditions its signal power would be a constant during the observation period. In addition, if the gains of the test and

reference channels would also stay constant during the measurements, under these conditions the relative amplitude measured would be:

$$\sqrt{\frac{P_T(t)G_T}{P_R G_R}} \tag{8.3-13}$$

This is proportional to the test antenna far-field pattern amplitude.

The variance in the relative amplitude measurement due to receiver noise is given by

$$\frac{P_T G_T}{P_R G_R}\left[\frac{N_{OT}}{2P_T\,\mathrm{T}} + \frac{N_{OR}}{2P_R\,\mathrm{T}}\right] \tag{8.3-14}$$

where T is the integration time for each measured amplitude.

In practice, the antenna microwave holography receiver was designed with a second-order PLL exhibiting a lag-lead loop filter and a selectable (variable) phase-locked loop (PLL) bandwidth designed to ope rate with phase noise values of 1–3 deg at 50 dB-Hz. This enables the receiver to track over a wide range of the commercially available geosynchronous satellites.

Figure 8-8 shows typical antenna far-field amplitude and phase pattern measured by the antenna microwave holography receiver described above. Figure 8-8 is the result of sampling a 127 × 127 data array from a 34-m diameter antenna (DSS-13) scanning at ±2.65 deg relative to the satellite nominal position.

8.4 Mathematical Formulation Data Processing

The mathematical relationship between an antenna far-field radiation pattern (T) and the antenna surface-induced current distribution (J) is given by

$$\begin{aligned}\vec{T}(u,v) = \iint_S \tilde{J}(x',y')\exp(jkz')\bullet\left[\exp\left[-jkz'(1-\cos\theta\right]\right] \\ \bullet\exp\left[jk(ux'+vy')\right]dx'dy'\end{aligned} \tag{8.4-1}$$

where

$z'(x',y')$ = defines the surface S

u, v = direction cosine space

θ = observation angle

For a small angular extent of the far-field pattern, this expression reduces to

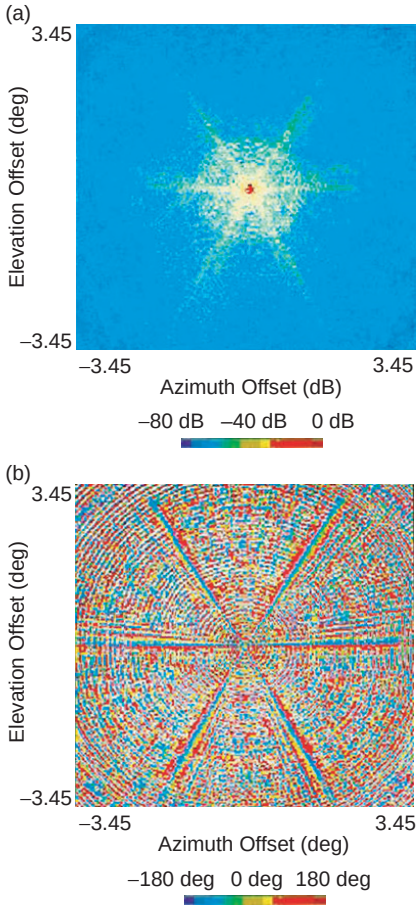


Fig. 8-8. Antenna far-field pattern measured by antenna microwave holography at Ku-band (12-GHz) showing (a) amplitude and (b) phase.

$$\vec{T}(u,v) = \iint_S \tilde{J}(x',y')\exp(jkz')\bullet\exp\left[jk(ux'+vy')\right]dx'dy' \qquad (8.4\text{-}2)$$

Equation (8.4-2) is a n exact Fourier transform of the induced surface current. To derive the residual surface error, geometrical optics ray tracing is used to relate the normal error, ε, to the axial error and phase in a main-reflector paraboloid geometry (Fig. 8-9):

$$\frac{1}{2}\Delta PL = \frac{1}{2}\left[P^1P + PQ\right] = \frac{1}{2}\left[\frac{\varepsilon}{\cos\varphi} + \frac{\varepsilon\cos 2\varphi}{\cos\varphi}\right] = \varepsilon\cos\varphi \qquad (8.4\text{-}3)$$

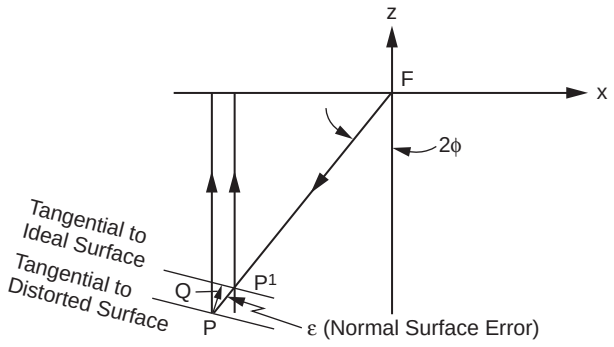


Fig. 8-9. Surface distortion geometry.

$$\text{Phase}\,(\Delta PL) = \frac{4\pi}{\lambda}\varepsilon\cos\varphi \tag{8.4-4}$$

and

$$\cos\varphi = \frac{1}{\sqrt{1+\dfrac{X^2+Y^2}{4F^2}}} \tag{8.4-5}$$

where F is focal length.

Allowing for the removal of a constant phase term and substituting Eq. (8.4-4) into Eq. (8.4-2) yields

$$\begin{aligned}\vec{T}(u,v) = \exp(-j2kF)\iint_s \tilde{J}(x',y')\bullet\exp\left(j4\pi\frac{\varepsilon}{\lambda}\cos\varphi\right)\\ \bullet\exp\left[jk(ux'+vy')\right]dx'dy'\end{aligned} \tag{8.4-6}$$

For the processing of sampled data, the associated discrete Fourier transform (DFT) is utilized:

$$\begin{aligned}T(p\Delta u, q\Delta v) = sx\,sy\sum_{n=-N1/2}^{N1/2-1}\sum_{m=-N2/2}^{N2/2-1} J(nsx, msy)\\ \bullet\exp\left[j2\pi\left(\frac{np}{N1}+\frac{mq}{N2}\right)\right]\end{aligned} \tag{8.4-7}$$

where

$N1\times N2$ = measured data array size

sx, sy = sampling intervals on the aperture, coordinates

n, m, p, q = integers indexing the discrete samples

$\Delta u, \Delta v$ = sampling interval in the far-field space

Since the magnitude of the far-field pattern is essentially bounded, the fast Fourier transform (FFT) is usually used for computation, and it is symbolized here by (F). Solving for the residual normal surface error and substituting Eq. (8.4-5), we obtain

$$\varepsilon(x,y) = \frac{\lambda}{4\pi}\sqrt{1+\frac{x^2+y^2}{4F^2}}\,\text{Phase}\left[\exp(j2kF)F^{-1}\left[T(u,v)\right]\right] \qquad (8.4\text{-}8)$$

The spatial resolution in the final holographic maps was defined in Eq. (8.2-5).

The resulting aperture function needs to be corrected for modulo-2π phase errors and a global least-squares fit performed on the data to the "best-fit" paraboloid. This process also allows for the correction of antenna-pointing errors introduced during the measurement. The "best-fit" paraboloid is found by minimizing S, the sum squares of the residual path length changes:

$$S = \sum_{i=1}^{N^2} \Gamma(\Delta PL_i)^2 A_i \qquad (8.4\text{-}9)$$

where

Γ = support domain constraints masking operator

ΔPL_i = path length change

A_i = amplitude weighting factor

with respect to 6 degrees-of-freedom of the reflector motion; three vertex translations, two rotations, and a focal length change. The six partial differential equations, which are solved simultaneously, are of the form [14]:

$$\frac{\partial S}{\partial \text{Par}} = 2\sum_{i=1}^{N^2} \Gamma \frac{\partial \Delta PL_i}{\partial \text{Par}} \Delta PL_i A_i = 0 \qquad (8.4\text{-}10)$$

where Par is one parameter of the 6 degrees-of-freedom.

It is correct to apply the best-fit paraboloid algorithm to either the conventional Cassegrain paraboloid-hyperboloid or dual-shaped reflector systems even though the latter do not use a paraboloid as the main reflector. Either design is a planewave-to-point source transformer, differing only in the field intensity distribution. The resultant aperture function at the end of this process is referred to as an "Effective Map" since it includes all phase effects that are contributing to the antenna performance. These effects include the subreflector scattered (frequency-dependent) feed phase function. Removal of the feed-phase function and subreflector support structure diffraction effects

results in a frequency independent map, which is referred to below as the "Mechanical Map."

Panel setting information is derived by sorting together all the data points within each panel and performing a least-squares fit. The algorithms allow for one translation and two rotations, S^k, α^κ, β^κ; hence, it can be referred to as a rigid body motion. For each panel and its associated n data points, we solve for the motion parameters via Eq. (8.4-11) (Fig. 8-10). This mathematical process also increases the accuracy in determining the screw adjustment correction [14] by a factor of $\sqrt{n}$.

$$\begin{bmatrix} \sum_{i=1}^{n}\cos^2(\gamma_i) & \sum_{i=1}^{n} d_i * \cos^2(\gamma_i) & -\sum_{i=1}^{n} e_i * \cos(\gamma_i) \\ \sum_{i=1}^{n} d_i * \cos^2(\gamma_i) & \sum_{i=1}^{n} d_i^2 * \cos^2(\gamma_i) & -\sum_{i=1}^{n} e_i * d_i * \cos(\gamma_i) \\ -\sum_{i=1}^{n} e_i * \cos(\gamma_i) & -\sum_{i=1}^{n} d_i * e_i * \cos(\gamma_i) & \sum_{i=1}^{n} e_i^2 \end{bmatrix} \begin{bmatrix} S^k \\ \alpha^\kappa \\ \beta^\kappa \end{bmatrix} = \begin{bmatrix} -\sum_{i=1}^{n} \varepsilon_i * \cos(\gamma_i) \\ \sum_{i=1}^{n} \varepsilon_i * d_i * \cos^2(\gamma_i) \\ \sum_{i=1}^{n} \varepsilon_i * e_i * \cos(\gamma_i) \end{bmatrix} \tag{8.4-11}$$

For optimal telecommunication performance in terms of G/T, the Cassegrain dual shaped-reflector antenna system has the advantage of providing higher aperture gain and aperture efficiency relative to the traditional paraboloid–hyperboloid Cassegrain design. However, for radio astronomy, where high beam efficiency is needed, the latter provides the better choice. The

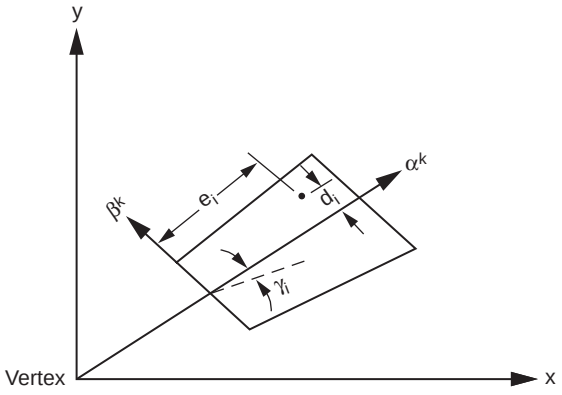
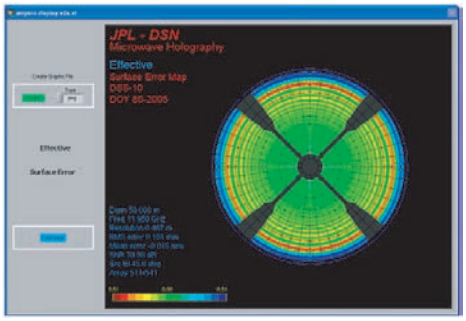


Fig. 8-10. Reflector panel geometry.

challenge of holographic applications for these antenna systems is the high level of diffraction effects, due to the subreflector edge, appearing towards the edge of the dish m ain reflector aperture (see simulation of feed diffraction effects on hologr aphic processing Fig. [8-11(a)]). The holographic map that includes these effects is termed "Effective Map" in contrast to the "Mechanical Map" where the diffraction effects have been removed. If the diffraction effect due to the feed phase function is kept during the derivation of the panel setting corrections, it w ill tend to mechanically "tune" and improve the dish performance at the measurement frequencies, but it w ill degrade the dish performance at other frequencies. If the antenna performance is to be optimized over a wide range of operating frequencies, as is the casee for the NASA-JPL-antennas, the diffraction effects must be removed prior to de riving the panel setting correction. Figure 8-11(b) shows the "Mechanical Map" where diffraction effects were reduced considerably resulted from applying the diffraction cancellation operation [14]. The error in panel setting correction resulted from the residual diffraction effect is on the order of 30 micrometers (μm).

(a)



(b)

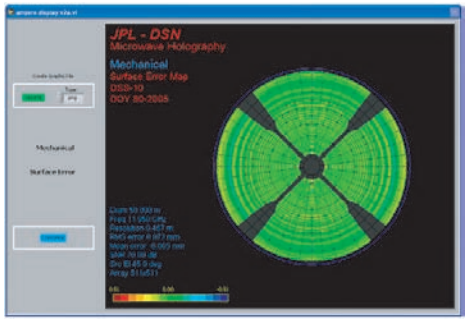


Fig. 8-11. Simulated maps, including (a) simulated effective map and (b) simulated mechanical map.

8.5 Applications

8.5.1 34-m BWG Research and Development Antenna

In August 1990, holographic measurements from the Cassegrain *f*1 focus of the new DSS-13 BWG antenna in Goldstone, California, were made (Fig. 8-12 and Fig. 7-2). Strong CW signals from geostationary satellite beacons were used as far-field sources. Three different geostationary satellites were scanned, producing high- and medium-resolution data sets at elevation angles of 46.5, 37, and 12.7 deg The measurements obtained provided the necessary subreflector position inf ormation, panel setting information, a look a t the adjusted surface of the antenna, and information about the gravity performance of the structure at a low elevation angles. The holographic antenna measurements used satellite signal and ephemeris information supplied by several commercial companies for GTE (GSTAR W103), GE (SatCom K1), and ComSat (Intelsat V).

Functionally, the outer 0.6 m of the DSS-13 antenna is designed as a noise shield. The rms error obtained from analysis of the central 32 m of the antenna is, therefore, more representative of the actual surface than the rms obtained from examination of the full 34-m dish. Therefore, we will present here only the rms values for the central 32 m of the antenna. The precision in the derived surface error maps is 0.05 mm (50 μm). In general, the indicated rms increases as the lateral resolution of the measurement increases (i.e., δ is a s maller numerical value) (Eq. (8.2-1). This is a n expected result as there is le ss area averaging occurring as the resolution increases. The asymptotic or infinite

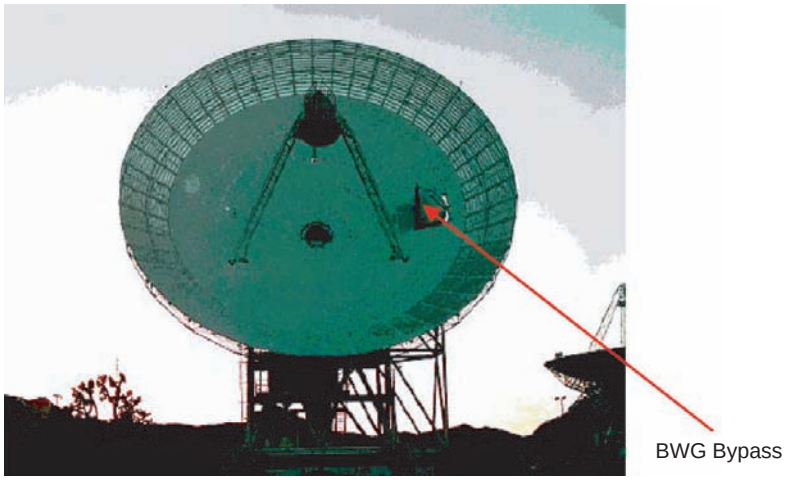


Fig. 8-12. The DSS-13 BWG antenna, which includes the BWG bypass structure in August 1990.

resolution rms can be estimated by analyzing the scan data at varying resolutions.

It is estimated that the rms error found by holography high-resolution (0.32-m) scans is 8 percent below the infinite resolution rms. Figure 8-13 shows the surface error map of the central 32 m of the DSS-13 antenna surface as found on August 28, 1990, at 46.5 deg elevation. The main reflector surface normal rms error was found to be 0.88 mm (0.77 mm axial) at a resolution of 0.32 m [15,16].

This measurement supplied the data required for verifying the subreflector position, analyzing the antenna surface, and providing the panel-setting information. The surface images derived from the aperture-plane phase represent the antenna surface deviations from ideal in the surface normal direction. In the images, the subreflector, the tripod and its shadows, and the bypass beam waveguide are intentionally masked out. The remaining surface is overlaid with an outline of each reflecting panel. The surface error information is shown in pseudo color, with red and blue indicating the high and low deviations of ±1.25 mm, respectively. The panel-setting information derived from this scan was applied to the 348 panels adjusted by the 1716 adjustment screws for this antenna. As a scheduling expedient, it was decided to adjust the surface panels by turning the adjusting screws to the nearest 1/8 of a turn (0.16 mm). Screws requiring adjustment of less than ±1/8 of a turn were not touched.

Figure 8-14 was derived on September 7, 1990 after the first application of holography panel resetting. The image reveals that the panels in the outer two rings are overbent. The rms surface error achieved by holography-based rigid

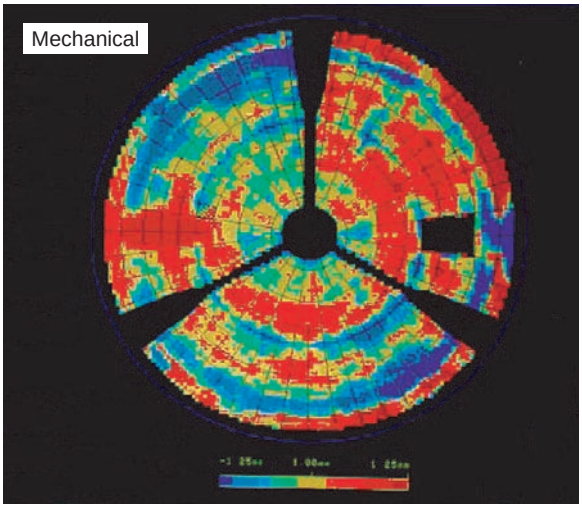


Fig. 8-13. DSS-13 after the initial theodolite setting.

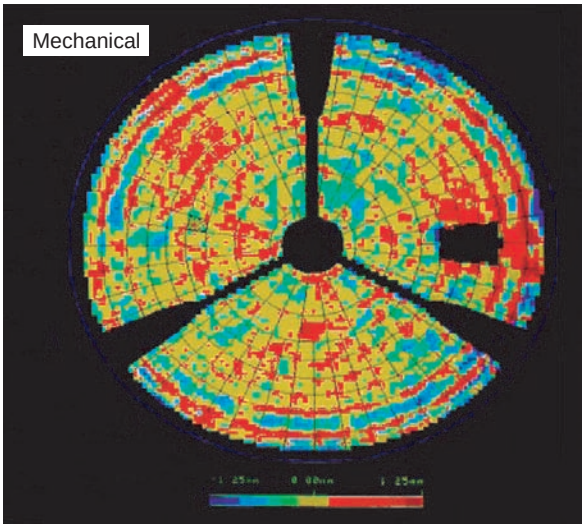


Fig. 8-14. DSS-13 after first application of holography.

body panel adjustment is 0.43 mm surface normal at a resolution of 0.32 m (Table 8-1). The post-holography surface provides a performance improvement (G/T) of about 0.32 dB at X-band (8.45 GHz), and 4.6 dB at Ka-band (32 GHz). The measured antenna efficiency at 46 deg elevation, after the holography panel setting, was 52.3 percent at Ka-band (32 GHz) and 75.4 percent at X-band (8.45 GHz) [15,16].

Please note that additional gain improvements due to subreflector position corrections are not accounted for in Table 8-1.

8.5.2 Gravity Performance of the BWG Antennas

The surface-error map shown in Fig. 8-15 was derived from medium-resolution, 0.80-m holography measurements made on September 18, 1990 at an elevation angle of 12.7 deg. The surface normal rms error at this low elevation angle and resolution is 0.50 mm. The measured antenna efficiency at 12.7 deg elevation was 39.4 percent at Ka-band and 74.0 percent at X-band.

The asymmetry revealed in this holographic low elevation map is attributed to the beam waveguide (BWG) bypass structure (shown in Fig. 8-12). Expectations were that removing the bypass would eliminate the asymmetrical gravity distortion and improve the 32-GHz gravity performance of the antenna. After the removal of the bypass BWG, and replacement with four panels, the measured gravity distortion function was indeed symmetrical (Fig. 8-16). However, the performance of the antenna as a function of elevation angle did not improve. Actually, the antenna gravity performance roll-off after the bypass

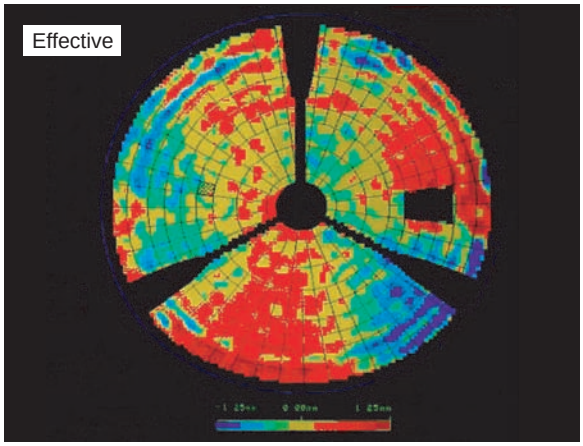


Fig. 8-15. Holographic surface-error map imaging at 12.7-deg elevation.

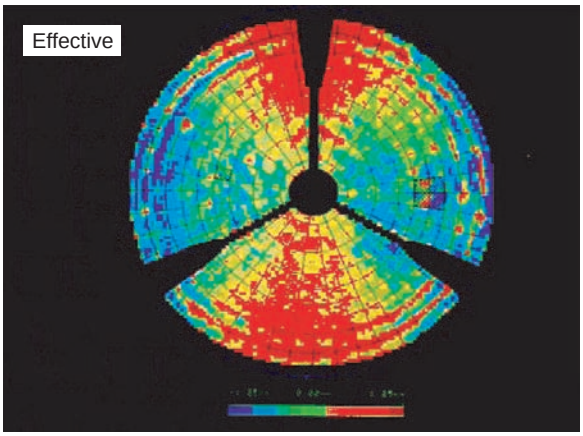


Fig. 8-16. DSS-13 imaging at 12.5-deg elevation after removing the bypass.

removal was 2.3 dB between 46- and 12.7-deg elevations at Ka-band, while with the bypass in place, it was only 1.2 dB. From this test (and combined with structural analysis by Roy Levy [17]) it was clear that the bypass structure, although causing asymmetrical gravity response, was adding significant stiffness to the antenna backup structure—a highly desirable feature for improved Ka-band performance. This lesson was applied to all future NASA–

JPL DSN 34-m BWG antennas, building them stiffer, and thereby achieving a gravity deformation loss of only 1.0 dB at Ka-band. (The new 34-m BWG antennas have a quadripod support for the subreflector instead of a tripod and this increases stiffness significantly.

Figure 8-17 shows the mechanical surface error map that was obtained (January 1992, Table 8-1) from the holographic measurements made on the DSS-13 after the removal of the BWG bypass and the application of a second holographic panel setting to the antenna. The normal rms surface error achieved was 0.38 mm, which agrees well with the 1990 predicted (best achievable by panel setting) surface of 0.36 mm. This reduction in rms (down from 0.43 mm) contributed an additional 0.26-dB performance increase (G/T) at Ka-band. The deformed panels in rings 8 and 9 are clearly noticeable in Fig. 8-17. The images in Fig. 8-17 were derived at Ku-Band (12.1795 GHz) by raster-scanning the antenna beam across a commercial geostationary satellite. The very high lateral

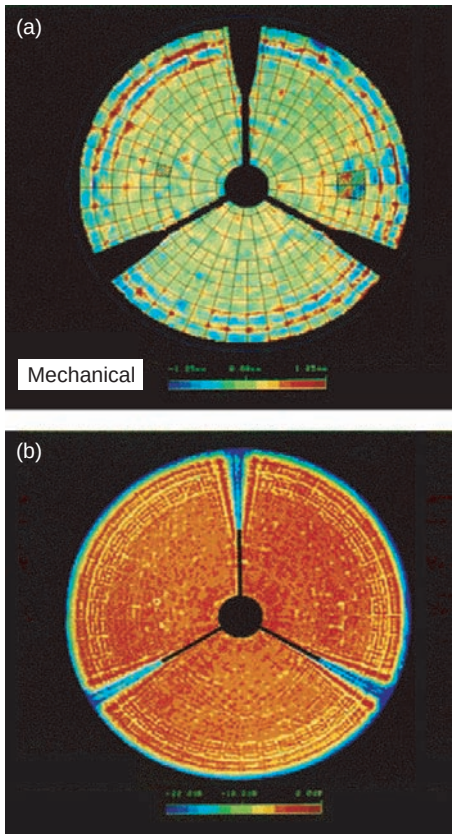


Fig. 8-17. Super-high resolution, 20-cm holographic imaging of DSS-13 showing (a) surface error and (b) current intensity map.

resolution of 20 cm is the result of a complex data array of 38,809 (197 × 197) samples of the far field of the antenna; the data array includes sampling to the antenna 150th side lobe. The surface current intensity map (Fig. 8-17(a)) and surface error (Fig. 8-17(b)) confirm features of the antenna's mechanical and electromagnetic designs. These images also provide information that can be used to physic ally correct a broad range of possible design deficiencies. Irregularities in the shape of the reflecting surface and in the intensity (power) distribution are revealed in the "light" of the microwave illumination. The surface-current map confirms the uniform illumination design of this dual-shaped reflector antenna.

Each of the 348 individual reflecting panels, as outlined in the surface-error map, is characterized by an average of 94 accurate data cells, from which information to mechanically adjust each panel is derived and applied by adjusting the 1716 antenna adjustment screws.

The rms surface achieved of $\lambda/25$ is capable of operating at Ka-band with only 0.4-dB efficiency loss due to sur face rms error. Estimates based on holographic measurements before and after adjustment of the surface indicate that more than 4 dB of performance was gained at 32 GHz. Radiometric measurements of antenna efficiency at 32 GHz confirm the excellent results obtained through the application of microwave holography.

During the planning stages of the DSS-13 BWG antenna project in 1988, it was decided to compromise and manufacture the main reflector panels by utilizing the existing DSS-15 34-m High Efficiency (HEF) antenna panel molds. The differences in the shape of the panels were thought to be minor, and it was believed that they would not signif icantly affect the required performance of the new research and development antenna. After the initial holographic imaging of DSS-13, it was clear from the images that the panels in rings 8 and 9 were systematically overbent (especially noticeable in the super-high resolution map of Fig. 8-17). Assuming that the panels on the DSS-13 antenna were made accurately from the DSS-15 manufacturing contours, the panels were mathematically best-fitted to the DSS-13 design contour. The axial errors between these two contours were calculated for each of the nine rings and then subtracted from the reference DSS-13 required shape.

The errors in the first seven panel rings are minor and cause no significant loss at 32 GHz. The errors in r ings 8 a nd 9 a re much more sizable (smooth curves in F igs. 8-18 and 8-19) and contribute noticeably to the antenna RF performance at 32 GHz. The step-wise linear plots in F igs. 8-18 and 8-19 are the holographically derived errors in the panels in r ings 8 a nd 9. The holography data are sampled over seven points a cross each panel, and these data agree very well with the mechanically derived smooth curves. Indeed, when we compute the antenna rms error for the inner seven rings only in Fig. 8-17, the result is 0.28 mm, while the rms error for the outer rings 8 and 9 (excluding the noise shield) is 0.60 mm [16].

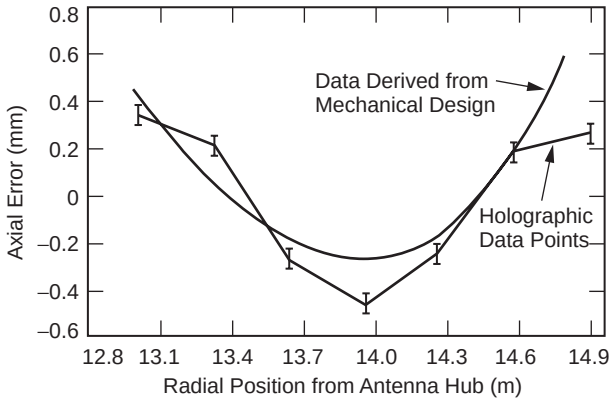


Fig. 8-18. Mechanical error in panel ring 8.

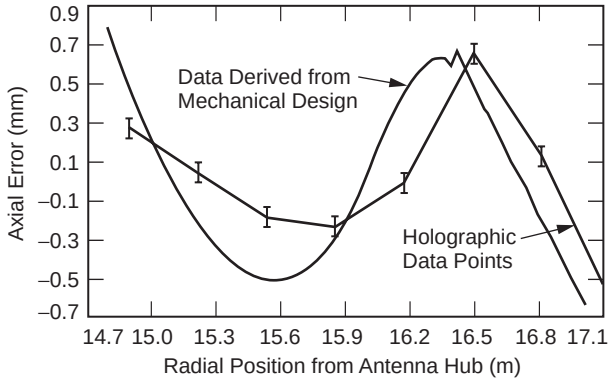


Fig. 8-19. Mechanical error in panel ring 9.

The potential increase of performance at Ka-band by achieving 0.28-mm rms for the entire dish is 0.6 dB. The DSS-13 antenna efficiency at the time of these measurements was 52 percent at f1 focus at Ka-Band (32 GHz), and it was predicted that the antenna efficiency would increase to 60 percent if the 0.6-dB opportunity were pursued. This proved to be the case; the current efficiency of the DSN 34-m BWG subnet antennas is 60 percent from f3 focus due to their effective rms of 0.25 mm, which was established by the microwave holography technique.

Replacing the damaged bent panels in rings 8 and 9 would have cost $300k; so instead a proposal to unbend them using the holographic technique was accepted. During the early part of February 1994, holographic measurements were made at DSS-13 to apply the panel unbending procedure to the panels in rings 8 and 9. The DSS-13 antenna surface error was further reduced from 0.38 mm to 0.31 mm corresponding to an additional 0.32 dB performance improvement at Ka-band (Table 8-1). Applying a total of four panel

setting/unbending sessions at DSS-13 between 1990 and 1994 resulted in reduction of its surface rms error from an initial 0.88 mm to 0.31 mm, which improved its RF performance at Ka-band (32 GHz) by approximately 5.3 dB. The efficiency of the DSS-13 measured from the BWG focus (F3) is 57 percent. This corresponds to an estimated efficiency from the Cassegrain F1 focus of approximately 65 percent.

8.5.3 Operational DSN 34-m BWG Antenna Network

Between May of 1994 and July of 2003, six newly constructed NASA–JPL DSN 34-m BWG antennas were measured holographically, and their panels and subreflectors were set and aligned. Three of the six antennas are located in the Goldstone Deep Space Communication Complex (GDSCC) in California and are designated DSS-24, DSS-25, and DSS-26. Two 34-m BWG antennas (designated DSS-54, and DSS-55) are located near Madrid, Spain (MDSCC), and one 34-m BWG antenna (designated DSS-34) is located near Canberra, Australia (CDSCC). The summary results of these measurements are presented in Table 8-1.

At GDSCC, the measurements were made from the Cassegrain F1 focus, utilizing a Ku-band (11.9-GHz) beacon signal from the GSTAR-4 satellite observed at the nominal elevation angle of 47 deg. Gravity and performance measurements at low elevation angles were taken using the beacon signal of the Intelsat-307 satellite observed at the nominal elevation angle of 12.7 deg. At 47-deg elevation, the normal rms surface errors of the DSS-24, DSS-25, and DSS-26 (as set by the theodolite) were 0.50 mm, 0.50 mm, and 0.42 mm, respectively (Table 8-1). Figure 8-20 shows the holographically derived surface error map of DSS-24 after the alignment of the panels using the theodolite technique, achieving a normal rms surface error of 0.50 mm [18,19].

Figure 8-21 shows the holographically derived surface error map of DSS-24 after applying only one session of holography derived panel setting, achieving a normal rms surface error of 0.25 mm (the color scales in the images of Figs. 8-20 and 8-21 is ±1.25 mm). The 34-m BWG network antennas have 348 panels and 1716 adjusting screws, with the rms surface of the individual panels specified at 0.127 mm, and the rms surface error of the subreflector is 0.125 mm. Since a precision panel adjusting tool was not used in order to reduce antenna down time, the panel listing data were rounded to the nearest ±1/8 of a screw turn. This enabled resetting the entire dish in an 8-hour period. The inferred root sum square (rss) panel setting accuracy is therefore 0.175 mm rms. The precision of the antenna surfaces in terms of diameter/rms is 1.36×10^5, and the gain limit of the antennas occurs at 95 GHz. The resulting measured efficiencies of the antenna from the Cassegrain f1 focus at 46.3-deg elevation were: 75.25 percent at X-band and 60.6 percent at Ka-band

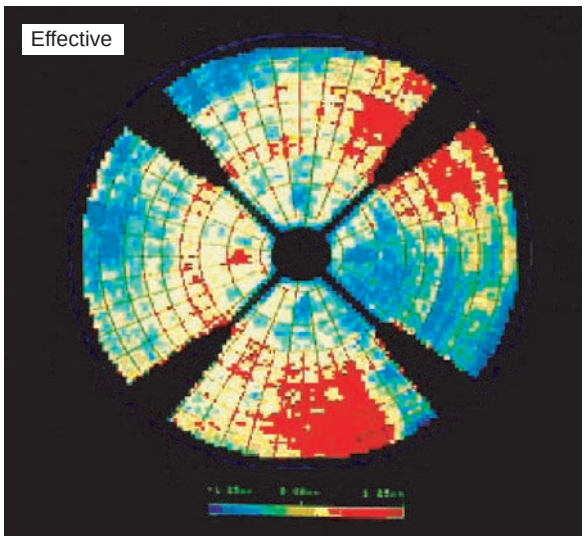


Fig. 8-20. Holographically derived DSS-24 surface error map before holographic alignment (theodolite alignment only), 0.50-mm rms.

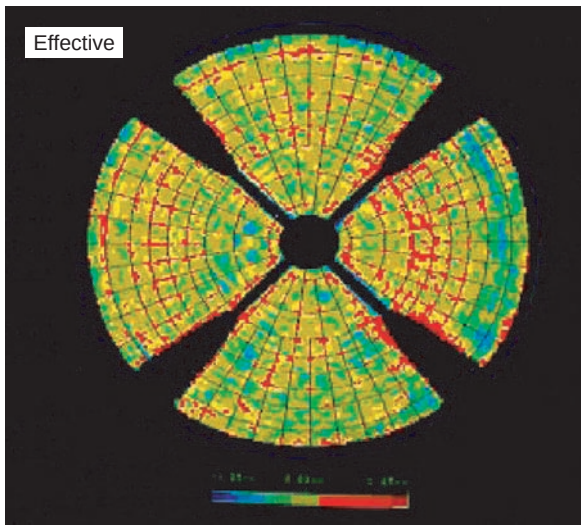


Fig. 8-21. DSS-24 surface-error map after holographic alignment, 0.25-mm rms.

(referenced to the input LNA). At 32 GHz (Ka-band), the averaged improved performance for each of the GDSCC antennas due to holography panel setting is estimated to be 1.1 dB. The antennas rms surface error at 12.6-deg elevation

averages approximately 0.50 mm, and it is mostly characterized by astigmatism due to gravity deformation, as expected.

The conventional process was used to set the panels of the first five 34-m BWG antennas (Table 8-1) that came on line. This process consisted of initial theodolite metrology panel setting to bring the antenna rms surface error down to 0.54 mm (0.02 in.), followed by holographic panel setting that further reduced the antennas rms error to 0.25 mm. The theodolite panel setting typically required 6 weeks of antenna down time. Given the experienced gain with these five antennas, and noticing the high efficiency of the holographic panel setting application, it was decided to relax the (total station) theodolite setting to a "rough" 1.0 mm (0.05 in.) (thereby reducing the antenna down time from 6 weeks to 1 week) and let holography bring the antenna rms down to 0.25 mm. This new process promised a saving a total of approximately 5 weeks of antenna down time, human resources, and cost to the project if it were successful.

This new process was applied successfully for the first time at MDSCC on DSS-55 in July of 2003 (Table 8-1). Figures 8-22 and 8-23 are the holographically derived surface error maps of DSS-55 before (Fig. 8-22) and after (Fig. 8-23) holographically application showing the reduction of the antenna rms surface error from 0.90 mm to 0.25 mm presented on a ±0.73-mm color scale.

8.5.4 Subreflector Position Correction

The theory for the subreflector position correction via holography can be found in [14]. The subreflector position correction is derived from the low-order distortions in the antenna aperture phase function, which is derived from low-resolution holographic imaging (25 × 25 array for a 34-m antenna (or 51 × 51 for a 70-m antenna). Two low-resolution measurements are usually required due to the interaction of cubic and linear terms; the latter is due to systematic pointing errors. The time required for a single low-resolution measurement is approximately 45 minutes, and the data processing time is approximately 16 minutes.

Figure 8-24 shows the far-field amplitude pattern of DSS-24 as found in the initial stage of the holographic measurements. Corrections to the subreflector controller X, Y, and Z axis were applied as follows: 1.31 cm in the –X direction (The derivation of the subreflector correction in the X-axis is especially critical since no servo drive controller is available for this axis), 0.952 cm in the +Y direction, and 0.343 cm in the +Z direction resulting in the antenna far-field pattern shown in Fig. 8-25. From observing the antenna far-field pattern in these images, it is clear that the antenna went through a transformation from

unfocused to focused. The performance improvement obtained by setting the subreflector alone is 0.25 dB at X-band (8.45 GHz) and 3.6 dB at Ka-band (32 GHz).

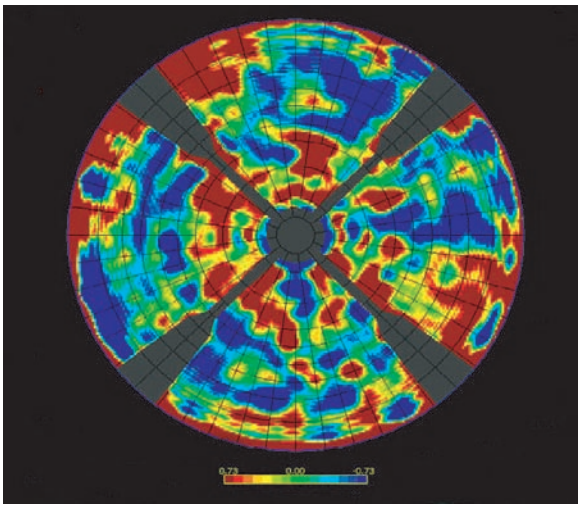


Fig. 8-22. DSS-55 surface error map before holography, 0.90-mm rms.

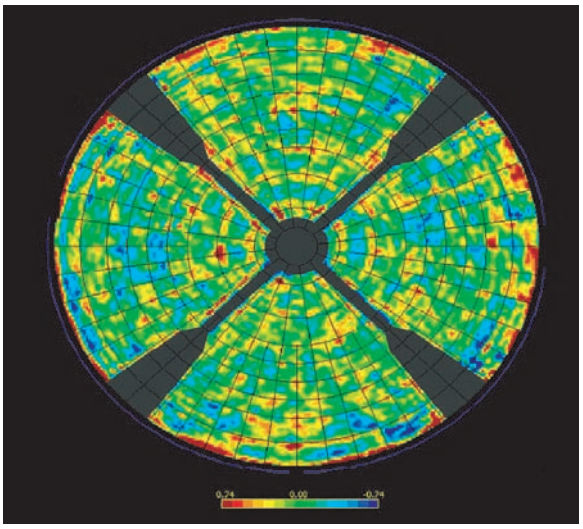


Fig. 8-23. DSS-55 surface error map after holography, 0.25-mm rms.

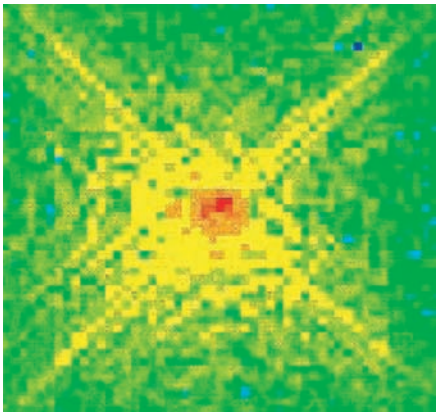


Fig. 8-24. DSS-24 far-field amplitude pattern before holographic corrections.

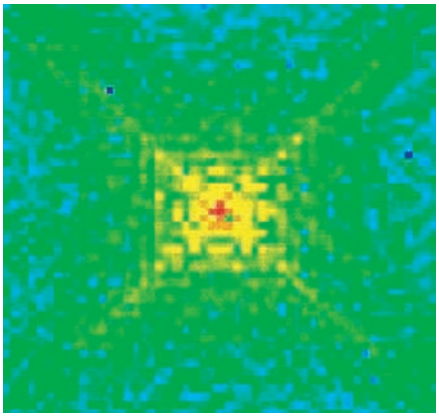


Fig. 8-25. DSS-24 far-field amplitude pattern after holographic corrections.

8.6 Conclusion

Microwave holography has proven to be an invaluable tool in the development and maintenance of large, high-performance ground antennas. The effective and highly successful application of microwave antenna holography to the large NASA–JPL DSN antennas has significantly improved their microwave and mechanical performance. For the 34-m BWG antenna subnet, the application of microwave antenna holography, combined with the implementation of low-noise system temperature of 22.3 K, resulted in a maximum *G/T* performance at Ka-band of 65.6 dB (on average and in vacuum) for each antenna at its rigging angle.

This improved performance has enabled new technologies and science advances. The added Ka-band observation frequency with the 34-m BWG subnet (which provided excellent amplitude and phase stability, high gain, and excellent blind pointing pe rformances) enabled the highly successful Cassini radio science data return from Saturn ring occultation and bistatic radar. Another example is the high data rate achieved of 6 megabits per second (Mb/s) communicating with the Mars Reconnaissance Orbiter (MRO) while at a distance of 0.225 astronomical units (AU) using the 34-m BWG at Ka-band.

References

[1] J. C. Bennet, A. P. Anderson, P. A. McInnes, and A. J. T. Whitaker, "Microwave Holographic Metrology of Large Reflector Antennas," *IEEE Transactions on Antennas and Propagation*, vol. AP-24, pp. 295–303, 1976.

[2] P. F. Scott and M. Ryle, "A Rapid Method for Measuring the Figure of a Radio Telescope Reflector," *Royal Astronomical Society, Monthly Notices*, vol. 178, pp. 539–545, March 1977.

[3] M. P. Godwin, A. J. T. Whitaker, and A. P. Anderson, "Microwave Diagnostics of the Chilbolton 25M Antenna Using the OTS Satellite," *International Conference on Antennas and Propagation, 2nd, York, England, Proceedings*, Part 1, pp. 232–236, April 13–16,1981.

[4] D. J. Rochblatt and B. L. Seidel, "DSN Microwave Antenna Holography," *The Telecommunications and Data Acquisition Progress Report 42-76*, October–December 1983, pp. 27–42, February 15, 1984. http://ipnpr.jpl.nasa.gov/progress_report

[5] D. J. Rochblatt, Y. Rahmat-Samii, and J. H. Mumford, "DSN Microwave Antenna Holography Part II: Data Processing and Display of High-Resolution Effective Maps," *The Telecommunications and Data Acquisition Progress Report 42-87*, July–September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 92–97, 1986.
http://ipnpr.jpl.nasa.gov/progress_report

[6] C. E. Mayer, J. H. Davis, W. L. Peters, III, and W. J. Vogel, "A Holographic Surface Measurement of the Texas 4.9-m Antenna at 86.1 GHz," *IEEE Transactions On Instrumentation and Measurement*, vol. IM-32, pp. 102–109, 1983.

[7] O. M. Bucci, G. D'Elia, G. Franceschetti, and R. Pierri, "Efficient Computation of the Far Field of Parabolic Reflectors by Pseudo-Sampling Algorithm," *IEEE Transactions on Antennas and Propagation*, vol. AP-31, no. 6, pp. 931–937, November 1983.

[8] J. Ruze, "Antenna Tolerance Theory—A Review," *Proceedings of the IEEE*, vol. 54, no. 4, pp. 633–640, April 1966.

[9] D. J. Rochblatt, "Systems Analysis for DSN Microwave Antenna Holography," *The Telecommunications and Data Acquisition Progress Report 42-96*, October–December 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 132–157, February 15, 1989. http://ipnpr.jpl.nasa.gov/progress_report

[10] D. J. Rochblatt and Y. Rahmat-Samii, "Effects of Measurement Errors on Microwave Antenna Holography," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 7, pp. 933–942, July 1991.

[11] D. Bathker, A. G. Cha, D. J. Rochblatt, B. L. Seidel, and S. D. Slobin, "70-meter Deep Space Network Antenna Upgrade Performance," *ISAP Japan 1989, Proceeding of the 1989 International Symposium on Antennas and Propagation*, Vol. 3, pp. 647–650, 1989.

[12] M. P. Godwin, M. F. Adams, P. J. Richards, E. P. Schoessow, and P. Wright, *Final Report on Ku-Band and X-Band Holographic Tests on the DSS-14 70-meter Antenna*, Eikontech Limited Report No A104 prepared for JPL, JPL D-40452, (internal document), Jet Propulsion Laboratory, Pasadena, California, October 1988.

[13] Y. Rahmat-Samii, "Microwave Holography of Large Reflector Antennas—Simulation Algorithms," *IEEE Transactions on Antennas and Propagation*, vol. AP-33, pp. 1194–1203, 1985 (see corrections, *IEEE Transactions on Antennas and Propagation*, vol. AP-34, p. 853, 1986.

[14] D. J. Rochblatt, "A Microwave Holography Methodology for Diagnostics and Performance Improvement for Large Reflector Antennas," *The Telecommunications and Data Acquisition Progress Report 42-108*, October–December 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 235–252, February 15, 1992. http://ipnpr.jpl.nasa.gov/progress_report

[15] D. J. Rochblatt and B. L. Seidel, "Microwave Antenna Holography," *IEEE Transactions on Microwave Theory and Techniques*, Special Issue on Microwaves in Space, vol. 40, no. 6, pp. 1294–1300, June 1992.

[16] D. J. Rochblatt and B. L. Seidel, "Performance Improvement of DSS-13 34-Meter Beam-Waveguide Antenna Using the JPL Microwave Holography Methodology," *The Telecommunications and Data Acquisition Progress Report 42-108*, October–December 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 253–270, February 15, 1992. http://ipnpr.jpl.nasa.gov/progress_report

[17] R. Levy, *Structural Engineering of Microwave Antennas: for Electrical, Mechanical, and Civil Engineering*, IEEE Antennas and Propagation Society, IEEE Press, New York, New York, 1996.

[18] D. J. Rochblatt, P. M. Withington, and H. J. Jackson, "DSS-24 Microwave Holography Measurements," *The Telecommunications and Data Acquisition Progress Report, No. 42-121*, January–March 1995, Jet Propulsion Laboratory, Pasadena, California, pp. 252–270, May 15, 1995. http://ipnpr.jpl.nasa.gov/progress_report

[19] D. J. Rochblatt, P. M. Withington, and H. J. Jackson, "DSS-24 Antenna RF Performance Measurements," *Microwave Holography Measurements*, 890-270 JPL D-12277, (internal document), Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 890-270: pp. 2-1 to 2-18, February 1, 1995.

Acronyms and Abbreviations

ACME	Antenna Calibration and Measurement Equipment
A/D	analog to digital
ADL	Arthur D. Little Corporation
AES	Auger electron spectroscopy
AGC	automatic gain control
AIAA	American Institute of Aeronautics and Astronautics
AIL	Airborne Instruments Laboratories
Al	aluminum
AlAs	aluminum arsenide
Al_2O_3	aluminum oxide
AlGaAs	aluminum gallium arsenide
AM	air mass
AMW	Antenna Microwave System
ANA	automatic network analyzer
As	arsenic
AsH_3	arsine
atm	atmosphere
Au	gold

A.U.	astronomical unit
AUT	antenna under test
AWVR	advanced water vapor radiometer
az-el	azimuth-elevation
Az-El	elevation-over-azimuth-drive
B	boron
B	radiometer factor, kelvins/watt (K/W)
Be	berylium
BP	bandpass
BPF	bandpass filter
bps	bits per second
BTL	Bell Telephone Laboratories
BW	bandwidth
BWG	beam waveguide (DSN 34-m antenna)
cal/g	calories per gram
cc	cubic centimeter
CCIR	Comité Consultatif International des Radio Communications (French: International Radio Consultative Committee)
CCR	closed-cycle refrigerator; closed cycle refrigeration
CCW	counterclockwise
CD	cumulative distribution of the atmospheric microwave loss, 0<CD<1
$(CH_3)_3Ga$	trimethylgallium (also abbreviated TMGa)
CH_4	methane
CMB	cosmic microwave background
CMF	cosmic microwave foreground
CO_2	carbon dioxide

Cr	chromium
Cr_2O_3	chromium oxide
CVD	chemical vapor deposition
CW	clockwise
CW	continuous wave
dB	decibel
dBi	decibels referenced to isotropic radiator
dBm	decibels referenced to milliwatts
dBW	decibels referenced to watts
dc	direct current
deg	degree
DFT	discrete Fourier transform
DGT	Deep Space Network Galileo Telemetry Receiver
DSCC	Deep Space Communication Complex
DSIF	Deep Space Instrumental Facility
DSIF	Deep Space Instrumentation Facility (combined with part of the Space Flight Operation Facility in 1963 to create the Deep Space Network)
DSMS	Deep Space Mission Systems
DSN	Deep Space Network
DSS	Deep Space Station
DSS-11	Pioneer Site at Goldstone, California
DSS-12	Echo Site at Goldstone, California
DSS-13	research station at Goldstone
DSS-14	70-m antenna at Goldstone
DSS-41	Woomera, Australia
DSS-42	near Canberra, Australia
DSS-43	70-m antenna in Australia

DSS-51	near Johannesburg, South Africa
DSS-54	antenna in Madrid
DSS-55	antenna in Madrid
DSS-63	70-m antenna antenna in Madrid
DSS-65	antenna in Madrid
DUT	device under test
DX (centers)	deep donor states (also referred to as deep-level trapping sites) related to the band structure of AlGaAs
EBL	electron beam lithography
EDM	electric discharge machining
EIRP	equivalent (or effective) isotropic radiated power
ESA	European Space Agency
eV	electron volt
f	frequency
feed assembly	feed horn + ambient waveguide components
feedcone	front-end assembly + supporting equipment + feedcone shell enclosure
FET	field effect transistor
FFT	fast-Fourier transform
FL	linearity factor
front-end assembly	feed assembly + LNA assembly
FTS	Frequency and Timing System
G	gain
G	gauss
Ga	gallium
GaAs	gallium arsenide

GAVRT	Goldstone-Apple Valley Radio Telescope program with 34-m antenna (DSS 12)
GCF	Ground Communications Facility
Ge	germanium
GE	General Electric Company
GHz	gigahertz
GIT	Georgia Institute of Technology (Atlanta, Georgia)
GM	Gifford-McMahon (cycle)
g/mm^3	grams per cubic millimeter
GPIB	general purpose interface bus
GSSR	Goldstone Solar System Radar
g/s	grams per second
G/T	gain-to-temperature ratio; antenna gain (proportional to antenna area) and the noise temperature performance; a system figure of merit
GUI	graphic user interface
G10	fiberglass and epoxy composite
H_2	hydrogen
H_2O	water
He	helium
HEF	High Efficiency (DSN 34-m antenna)
HEMT	high electron mobility transistor
HFET	heterostructure transistor (SDHT), and heterojunction FET (synonym for HEMT)
HP	horsepower
HPBW	half-power beamwidth (degrees)
HSB	High-Speed Beam Waveguide (antenna)
Hz	hertz

ICE	Inter-Cometary Explorer
IEEE	Institute of Electrical and Electronics Engineers
IF	interface
IF	intermediate-frequency
IMPATT	impact avalanche and transit time
In	indium
InAlAs	indium aluminum arsenide
InGaAs	indium gallium arsenide
InP	indium phosphide
ISO	International Organization for Standardization
ISS	impedance standard substrate
ITU	International Telecommunications Union
J	joule
J/K	joules/kelvin
JMOC	Jupiter Microwave Observation Campaign (Cassini)
JPL	Jet Propulsion Laboratory
J-s	joule-second
JT	Joule-Thomson
JWST	James Webb Space Telescope
K	kelvin
Ka-band	31.8–32.3 GHz (DSN frequency allocation within Ka-band)
KaBLE	Ka-Band Link Experiment
kbps	thousand bits per second
KHz	kilohertz
km	kilometer
Ku-band	15 GHz
kW	Kilowatt

l	liter
LAN	local-area network
L-band	960 MHz
LCER	Lewis Center for Educational Research
LCP	left-hand circularly polarized
LH_2	liquid hydrogen
LHe	liquid helium
LN_2	liquid nitrogen
LNA	low-noise amplifier
LNA assembly	low-noise amplifier + post amplifier + gain set attenuator
LO	local oscillator
lps	liters per second
LRM	line-reflect-match (calibration method)
LRO	Lunar Reconnaissance Orbiter
LWC	liquid water content
m	meter
maser	microwave amplification by stimulated emission of radiation
MBE	molecular beam epitaxy
Mb/s	megabit per second
mdeg	millidegree
MDSCC	Madrid Deep Space Communication Complex
MEC	Microwave Electronics Corporation
MESFET	metal semiconductor field effect transistor
MHEMT	metamorphic high electron mobility transistor
mho	(a unit of conductance equal to the reciprocal of an ohm)
MHz	megahertz
mm	millimeter

µm	micrometer
MMIC	microwave monolithic integrated circuit
MMICAD	(computer-aided engineering/computer-aided testing [CAE/CAT] software from Optotek, Ltd.)
MMS	(Agilent) modular measurement system
Mo	molybdenum
MOCVD	metal-organic chemical vapor deposition
MODFET	modulation-doped FET (synonym for HEMT)
MOVPE	metal-organic vapor phase epitaxy
MRE	mean radial error
MRO	Mars Reconnaissance Orbiter
ms	millisecond
µs	microsecond
m/s	meters per second
m^3/s	cubic meters per second
MSFN	Manned-Space-Flight-Network
mS/mm	millisiemens/millimeter
MTBF	mean time between failures
N_2	nitrogen
NAR	noise-adding radiometer
NASA	National Aeronautics and Space Administration
ND	noise diode
Ni	nickel
NL	nonlinearity
NMC	(Deep Space Network) Network Monitor and Control
NRAO	National Radio Astronomy Observatory
NRAO, AUI	National Radio Astronomy Observatory of the Associated Universities Inc.

OCR	Open-cycle refrigeration
OFHC	oxygen-free high conductivity
OTF	on-the-fly (mapping)
Ω	ohm
P	phosphorus
Pa	pascal
par-amp	parametric amplifier
PCG	phase calibration generator
Pd	palladium
Perspex	polymethyl methacrylate
PH_3	phosphine
PHEMT	pseudomorphic high electron mobility transistor
PLL	phase-locked loop
PM	power meter
PMAA	polymethyl methacrylate
PPC	persistent photoconductivity
PSD	power spectral density
psia	pounds per square inch absolute
Pt	platinum
RAC	radio astronomy controller
RA–DEC	right ascension–declination
RCP	right-hand circularly polarized
R&D	research and development
RF	radio frequency
RFI	radio-frequency interference
RHEED	reflection high energy electron diffraction

RIE	reactive ion etching
R-J	Rayleigh-Jeans
RMS, rms	root mean square
rpm	revolutions per minute
RSS	root sum square
RTA	rapid thermal annealing
RWM	reflected-wave maser
S	siemens
SAR	signal adding radiometer
Sb	antimony
S-band	2.20–2.40 GHz (DSN frequency allocation within S-band)
SCFM	standard cubic feet per minute
SD	standard deviation = sigma = σ
SDHT	selectively-doped heterostructure transistor (synonym for HEMT)
SEM	scanning electron microscope
SFOF	Space Flight Operations Facility (combined with part of the Deep Space Instrumentation Facility in 1963 to create the Deep Space Network)
Si	silicon
SiC	silicon carbide
SiH_4	Silane
Si_3N_4	silicon nitride
SIM	Space Interferometry Mission
SiO_2	silicon dioxide
SMART-2	Small Missions for Advanced Research in Technology–2
SNR	signal-to-noise ratio
SNT	system noise temperature
SOLT	short-open-load-thru (calibration method)

SPC	(Deep Space Network) Signal Processing Center
SPD	S-band polarization diversity (feedcone)
std	standard
STDN	S-band Tracking and Data Network
SWS	slow-wave structure
T	tesla
teflon	polytetrafluoroethylene
TEGFET	two-dimensional electron gas FET (synonym for HEMT)
Ti	titanium
TMGa	trimethylgallium, $(CH_3)_3Ga$
TPR	total power radiometer
TV	television
TWM	traveling wave maser
UCSD	University of California, San Diego
ULNA	ultra-low-noise (X-band TWM [amplifier])
URSI	International Union of Radio Science
UWV	microwave
VC	voltage-controlled crystal oscillator
VLA	National Radio Astronomy Observatory) Very Large Array (Socorro, New Mexico)
VLBI	very long baseline interferometry
VSWR	voltage standing wave ratio
W	watt
W/cm	watts per centimeter
WG	waveguide

W/g	watts per gram
W/s	watts per second
WR	rectangular waveguide
WVR	water vapor radiometer
X-band	8.40–8.45 GHz (DSN frequency allocation within X-band)
XEl	cross elevation
XMT	transmit
XRO	X-band receive operational (DSN feedcone assembly)
XTR	X-band transmit and receive feed assembly
YIG	yttrium iron garnet
2-D	two dimensional
3-D	three dimensional