AD NUMBER

AD466938

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution: Further dissemination only as directed by Air Force Electronic Systems Division, Attn: ESAT, Hanscom AFB, MA, OCT 1963, or higher DoD authority.

AUTHORITY

esd ltr, 1 may 1968

ESD-TDR-63-354, VOL 6

466938

AD No.

DDC FILE COPY

FINAL REPORT
ON
RANGE INSTRUMENTATION PLANNING STUDY
TECHNICAL DOCUMENTARY REPORT NO. ESD-TDR-63-354

VOLUME 6: TTC AND LONG-HAUL COMMUNICATIONS

OCTOBER 1963

DIRECTORATE OF AEROSPACE INSTRUMENTATION
DEPUTY FOR ENGINEERING AND TECHNOLOGY
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Mass.

Prepared under Contract No. AF 19(628)-2356

by

TRW Space Technology Laboratories
One Space Park
Redondo Beach, California

g 35

ESD-TDR-63-354, VOL. 6

FINAL REPORT,

[redacted]

(6) RANGE INSTRUMENTATION PLANNING STUDY.

~~TECHNICAL DOCUMENTARY REPORT NO. ESD-TDR-63-354~~

VOLUME 6. TTC AND LONG-HAUL COMMUNICATIONS.

(11) OCT[redacted]R 1963

DIRECTORATE OF AEROSPACE INSTRUMENTATION
DEPUTY FOR ENGINEERING AND TECHNOLOGY
ELECTRONIC SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
L. G. Hanscom Field, Bedford, Mass.

(15)

~~Prepared under~~ Contract No. AF 19(628)-2356

by

TRW Space Technology ~~Laboratories~~.

~~One Space Park~~

Redondo Beach, California

FOREWORD

This volume consists of Appendices XX through XXIV, dealing with certain facets of technology especially pertinent to the range instrumentation problem, and Appendix XXV, which surveys and summarizes the state of the long-haul communications art.

The numbering of the figures, tables, references, and equations is consecutive for the appendices in this volume. For the convenience of the reader, however, the pagination of each appendix reflects the appendix designator; for example, Appendix XXI has its pages numbered 21-1, 21-2, 21-3, etc.

This volume and the other seven volumes making up this report have a standard table of contents immediately following the abstract page. Additionally a major-heading table of contents for each of the other volumes is shown in reduced size immediately following the standard table of contents. The list of abbreviations and symbols used in the report has been included in each volume.

The STL Document Control Number for Volume 6 is 8691-6099-RU000.

This abstract is UNCLASSIFIED

RANGE INSTRUMENTATION PLANNING STUDY (U)

ABSTRACT

The emphasis in space technology over the past several years has shifted to very-long-range missiles, orbiting vehicles, and deep-space probes. Thus support from the test environment is now required on a truly global basis. For this reason, equipment compatibility and a means of integrating the operations of the existing ranges have become necessities.

TRW Space Technology Laboratories (formerly Space Technology Laboratories, Inc.) was awarded a contract by ESD to perform a study of the entire test environment problem for these classes of missions and tests. The primary objective was the definition of what the global test environment should be in the 1965 to 1970 period and the development of an implementation plan permitting the timely and efficient attainment of the recommended configuration.

Emphasis was placed on providing the capability to support the requirements imposed by the many programs involving missiles, large boosters, and spacecraft to be tested in the 1965 to 1970 period. Efficiency was emphasized because it was known that the costs associated with provision of a capability to support such a variety of missions would be very large under the best of circumstances.

STL's conclusion is that an integrated global test environment is not only feasible but highly desirable. The report presents specific recommendations for the choice of prime instrumentation in this integrated global test environment, and develops operational and management concepts appropriate to it. The report includes detailed recommendations, an implementation plan, and applied research and advanced development plans necessary or desirable to assure the timely implementation of the range and to provide a basis for a continuing upgrading of its capabilities.

Volume 1 consists of an overall introduction, a summary of the total report, and a presentation of the basic system concept.

Volume 2 contains a detailed description of the recommended instrumentation network and gives the implementation plan and cost estimates.

Volume 3 summarizes, as functions of time and location, the most stringent requirements imposed on the network; evaluates the network's capability of meeting the test requirements; and presents recommendations for applied research and advanced development programs required to implement the network or to advance the state of the art in pertinent instrumentation technology.

Volumes 4 through 8 contain supporting appendices for the findings, conclusions, and recommendations presented in Volumes 1 through 3.

Publication of this technical documentary report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

C. V. HORRIGAN
Acting Director
Aerospace Instrumentation
Deputy for Engineering and Technology
Electronic Systems Division

CONTENTS

APPENDIX XXV. LONG-HAUL COMMUNICATIONS (Cont'd)

APPENDIX XXV. LONG-HAUL COMMUNICATIONS (Cont'd)

ILLUSTRATIONS

APPENDIX XXII. ELECTRONICALLY SCANNED ANTENNAS

APPENDIX XXIV. NUCLEAR RADIATION HAZARD TO ONBOARD TELEMETRY, TRACKING AND COMMAND EQUIPMENT

APPENDIX XXV. LONG-HAUL COMMUNICATIONS

VOLUME 1
SYSTEM CONCEPT

VOLUME 2
RECOMMENDED INSTRUMENTATION NETWORK

VOLUME
EVALUATION, REQU AND TECHNOI

VOLUME 3
ION, REQUIREMENTS, D TECHNOLOGY

VOLUME 4
INSTRUMENTATION DATA REQUIREMENTS

2

VOLUME 5
TRACKING, TELEMETRY, AND COMMAND (TTC)

TTC
C

VOLUME 6
TTC AND LONG-HAUL COMMUNICATIONS

VOLUME 7
DATA PROCESSING AND DISPLAY

VOLUME 8
MISCELLANEOUS CLASSIFIED APPENDICES

3

LIST OF ABBREVIATIONS

ac	alternating current
A/C	aircraft
ACIC	Aeronautical Chart and Information Center
AFC	automatic frequency control
AFCRL	Air Force Cambridge Research Laboratory
AFMTC	Air Force Missile Test Center
AFSC	Air Force Systems Command
AGC	automatic gain control
AICBM	anti-intercontinental ballistic missile
AIL	Airborne Instruments Laboratory
alt	altitude
AM	amplitude modulation
AMR	Atlantic Missile Range
AMS	Army Map Service
approx	approximately
ARDC	Air Research and Development Command
ARPA	Advanced Research Projects Agency
ARS	Air Rescue Service
BECO	booster engine cutoff
BO	burn out
BOA	broad ocean area
bpS	bits per sample
BTL	Bell Telephone Laboratories
BWO	backward wave oscillator
CCMTA	Cape Canaveral Missile Test Annex
cm	centimeter
COM	Chief Operations Manager
CONUS	Continental United States
cps	cycles per second
CRT	cathode-ray tube
CW	continuous wave
CY	calendar year
db	decibel
dbm	decibels referred to one milliwatt
dc	direct current

DCA	Defense Communications Agency
DCS	Defense Communications System
DE	Delco
deg	degree
deg/sec	degrees per second
DOD WGS	Department of Defense World Geodetic System
DSIF	Deep Space Instrumentation Facility
EAFB	Edwards Air Force Base
EBPA	electron beam parametric amplifier
EEG	electroencephalogram
EGO	Eccentric Geophysical Observatory
EKG	electrocardiogram
ESD	Electronic Systems Division
FDM	frequency division multiplex
FM	frequency modulation
FOT	frequency of optimum transmission
fps	feet per second
FSK	frequency-shift keyed
ft	foot
ft/sec	feet per second
g	gravity, acceleration due to
gc	gigacycles per second
GD/A	General Dynamics/Astronautics
GDOP	geometric dilution of precision
GE	General Electric
GERTS	General Electric Range Tracking System
GRCC	Global Range Control Center
HF	high frequency
hr	hour
HU	Hughes Products
ICGM	Intercontinental Global Missile
IF	intermediate frequency
IFCS	information flow control station
IGC	inertial guidance and control
IMCC	Integrated Mission Control Center
in.	inch

IR	infrared
IRBM	intermediate range ballistic missile
IRIG	inter-range instrumentation group
ISB HF	independent sideband high frequency
JPL	Jet Propulsion Laboratory
^{o}K	degrees Kelvin
kbits/sec	kilobits per second
kc	kilocycles per second
Kev	thousands of electron volts
km	kilometer
kw	kilowatt
LASV	Low-Altitude Supersonic Vehicle
LF	low frequency
LOS	line of sight
LRCC	local range control center
LUF	lowest usable frequency
m	meter
MATS	Military Air Transport Service
max.	maximum
mc	megacycles per second
MCC	mission control center
MEC	Microwave Electronics Corporation
MECO	main engine cutoff
megw	megawatt
Mev	millions of electron volts
MH	Minneapolis Honeywell
mi	miles
MILS	missile impact location system
Mil Spec	military specification
MIT	Massachusetts Institute of Technology
mm Hg	millimeters of mercury
MMRBM	mobile, medium-range, ballistic missile
mr	milliradian
msec	millisecond
MO	Motorola
MTBF	mean time between failures

MTS	members of the technical staff
MUF	maximum usable frequency
mw	milliwatt
NASA	National Aeronautics and Space Administration
nmi	nautical mile
NMFPA	U. S. Naval Missile Facility, Point Arguello
OAO	Orbiting Astronomical Observatory
OD	Operations Directive
OGO	Orbiting Geophysical Observatory
ORV	ocean range vessel
OSO	Orbiting Solar Observatory
PACM	pulse-amplitude code modulation
PAFB	Patrick Air Force Base
PAM	pulse-amplitude modulation
PCA	polar cap absorption
PCM	pulse-code modulation
PD	pulsed doppler
PDM	pulse duration modulation
PELT	precision early launch tracker
PH	Philco
PIRD	Program Instrumentation Requirements Document
PM	phase modulation
PMR	Pacific Missile Range
PN	pseudonoise
POGO	Polar Orbiting Geophysical Observatory
PPE	preliminary planning estimate
ppi	plan position indicator
ppm	parts per million
pps	pulses per second
PRF	pulse repetition frequency
PSK	pulse-shift keyed
R and D	Research and Development
RARR	range and range rate
RAY	Raytheon
RCA	Radio Corporation of America
ref	reference

RF	radio frequency
RIPS	Range Instrumentation Planning Study
RISE	Research in Supersonic Environment
rms	root mean square
rss	root sum square
RTG	radio-isotope thermoelectric generators
SCF	Satellite Control Facility
sec	second
SGLS	Space-Ground Link Subsystem
SLAM	Supersonic Low-Altitude Missile
smi	statute mile
S/N	signal-to-noise
SOFAR	Sound Firing and Ranging
SPO	Systems Project Office
Sps	samples per second
SRM	Systems Requirements Model
SSB	single sideband
SSD	Space Systems Division
SSN	sunspot number
STL	TRW Space Technology Laboratories
STV	special test vehicle
SWF	short-wave fadeout
SY	Sylvania
TASI	time assignment speech interpolation
TDM	time division multiplex
TI	Texas Instruments
TLM	telemetry
TR	Transitron
TTC	tracking, telemetry, and command
TTY	teletype
TV	television
TWT	traveling-wave tube
UFS	Unified Frequency System
UHF	ultra-high frequency
μrad/sec	microradians per second
USA WGS	United States Army World Geodetic Systems

USAF WGS	United States Air Force World Geodetic System
μsec	microsecond
VAFB	Vandenberg Air Force Base
VECO	vernier engine cutoff
VHF	very-high frequency
VLF	very-low frequency
w	watt
WE	Western Electric
wpm	words per minute
ws	watt-second
WSMR	White Sands Missile Range
Xpdr	transponder

Appendix XX. LOW-NOISE RECEIVERS

A. INTRODUCTION

This appendix deals with low-noise system performance. A brief consideration is given to the noise aspects of the entire surveillance, communication, or telemetry system; system noise temperature is defined; and the manner in which various noise contributions enter into determining system sensitivity is presented.

The results of a state-of-the-art survey on low-noise antennas and a wide variety of low-noise preamplifiers comprise the main portion of the following discussion. This field has been a very active one during the past several years. The types of low-noise preamplifiers now include masers, parametric amplifiers (cooled and uncooled), traveling wave tubes, high frequency planar triodes, tunnel diode amplifiers, and transistor amplifiers. A brief history and description of the principle of operation of each of these devices are included, but the main emphasis is placed on providing up-to-date information on the low-noise capabilities of these devices, together with other capabilities of interest such as frequency range, bandwidth, stability, and dynamic range. The factors which determine the noisiness of the device are examined, and an attempt is made to project the degree of noise reduction which will take place by 1970 in both devices and systems.

Finally, a comparison is made between the performance capabilities of different types of low-noise preamplifiers, and a distinction is made as to their suitability for application in either ground-based or missile and space vehicle receivers.

B. THE RECEIVING SYSTEM

1. ANTENNA AND SKY NOISE

With the development of extremely low-noise receivers the sensitivity of a space communication or telemetry system has become limited to a great extent by natural sources of noise in the environment rather than by noise in the man-made devices. It has become customary to group these natural noise contributions and call them antenna noise temperature. The antenna noise temperature is made up of sky noise from the galaxy, the sun, and the atmosphere, plus earth noise.

Galactic noise depends upon which direction relative to the galaxy the antenna beam is pointing. The greatest noise occurs when the beam is pointed at the galaxy center; the next most unfavorable orientation occurs when the beam is pointing anywhere in the galactic plane; the lowest noise occurs when the antenna beam is pointed at the galactic poles. Since the probability that the beam will point in the galactic plane is low, it is reasonable to assume an average galactic noise contribution which is only slightly greater than when the beam is pointing at the galactic poles.

The effect of atmospheric absorption is to somewhat attenuate galactic noise while contributing its own atmospheric noise to the effective sky temperature. The effect of atmospheric absorption depends upon the angle the antenna beam makes with the zenith since the effective path length through the atmosphere will thus change. A plot of effective sky noise temperature for several antenna beam angles is shown in Figure 1. A "low-noise window" is seen to exist between 1 and 10 gc. This window is bounded on the low-frequency end by an increasing galactic noise contribution, and on the high-frequency end by absorption resonances for water vapor and oxygen. Figure 1 is not an exact representation of sky noise, since it does not take into account solar noise, variations in water vapor content of the atmosphere, or bad weather. Under unfavorable conditions the sky temperature can greatly exceed the values of Figure 1.

The ideal antenna noise temperature would be the sky noise temperature; for real antennas, however, there must be added a contribution from ground noise radiation via the side lobes, back lobe, and spillover of the antenna. In an effort to minimize these effects a tapered illumination of the parabolic reflector is employed. With the feed located at the parabolic focal point, typical ground noise contributions are 15 to 20°K when the antenna is pointing toward zenith.[1] Further improvement in reducing ground noise can be achieved by 1) having a small focal length to diameter ratio, say less than 1/4, for the conventional feed, 2) using a Cassegrain system, 3) using a horn-reflector antenna, or 4) laying

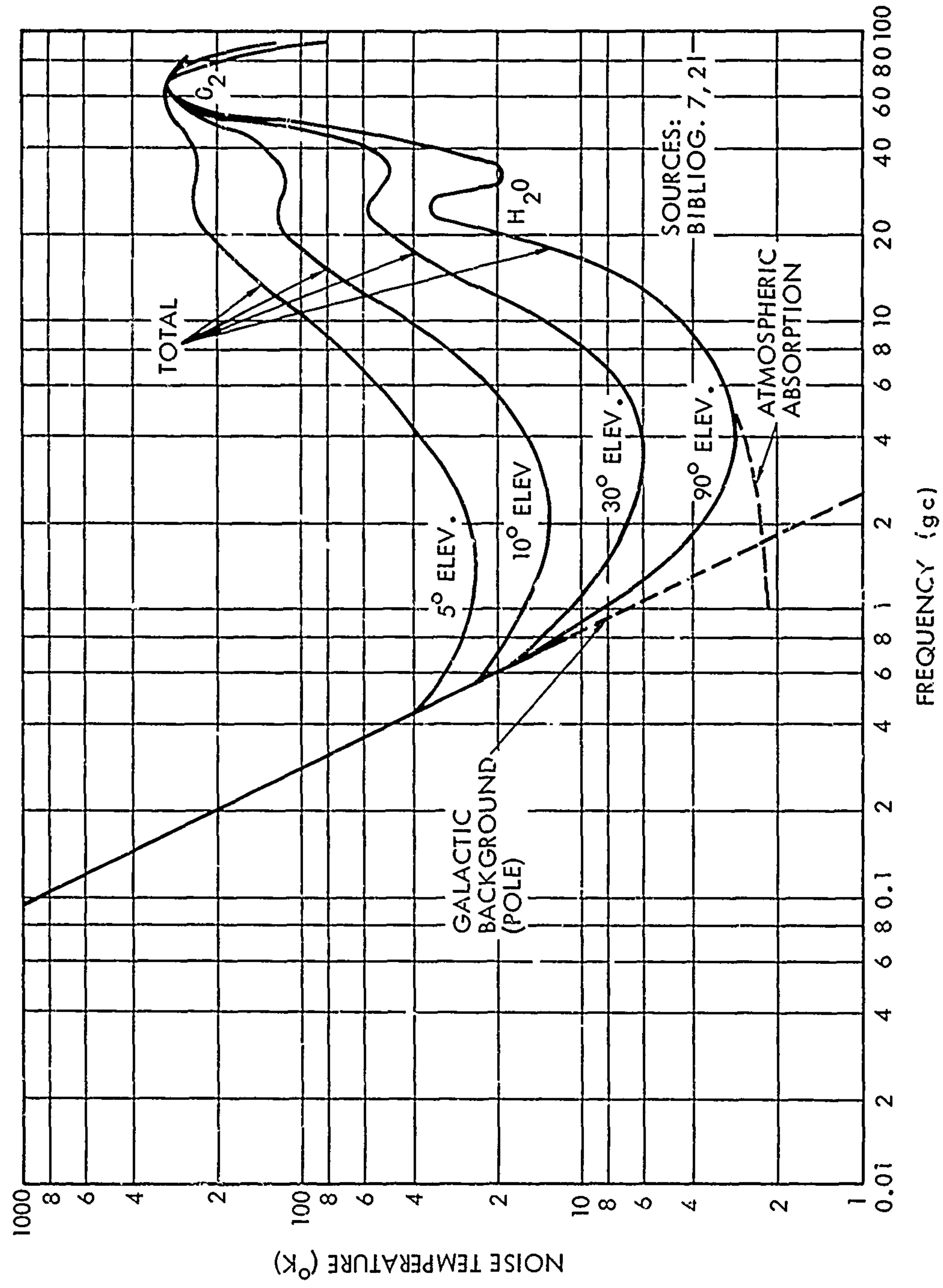


Figure 1. Sky Noise Temperature

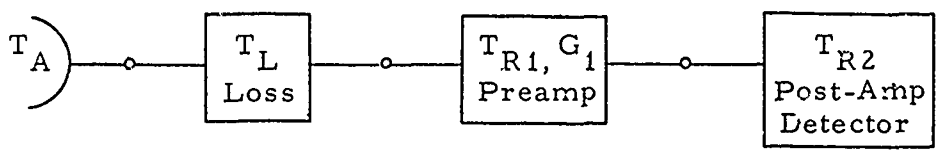
a very large screen with elevated edges in the vicinity of the antenna. Ground noise contributions can be reduced to about 2 to 10^{o}K by these means.

Some of the best reported results have been achieved using a horn-reflector antenna; however, this antenna is quite costly and has a large bulk-to-aperture ratio. To achieve a more compact design, consideration has been given to developing a "hybrid type" comprising features of both the horn-reflector and the Cassegrain antenna.[2] Another possible development in the more distant future might be to place a large parabolic antenna into orbit (possibly a stationary one). This would virtually eliminate ground and atmospheric noise, and an antenna noise temperature of far less than 1^{o}K could be attained at frequencies above about 5 gc where the galactic noise is small.

2. SYSTEM NOISE CONSIDERATIONS

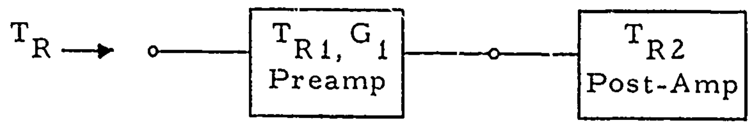
Once the system bandwidth and detection scheme have been chosen, the system sensitivity is then a function of the system noise level, which is expressed in terms of an equivalent thermal noise temperature.

Consider the simplified system shown below:



The antenna equivalent noise temperature is T_A. The receiver is composed of a preamplifier with a gain (G_1) and noise temperature (T_{R1}) and a post-amplifier and detector with a noise temperature (T_{R2}). Connecting the antenna to the receiver in a practical system requires some microwave circuitry such as diplexers, directional couplers, filters, rotary joints, and lengths of waveguide. These components are lossy and are lumped together as a loss ($L \geq 1$) at a temperature (T_L) which may be a source of thermal noise.

For cascaded systems the equivalent receiver noise temperature is given by



$$T_R = T_{R1} + \frac{T_{R2}}{G_1}$$

A low-noise receiver must have a low-noise preamp having enough gain to make the post-amplifier noise contribution small. For preamplifiers having a constant gain-bandwidth relationship, it may be necessary to cascade low-noise preamps to achieve adequate gain and bandwidth.

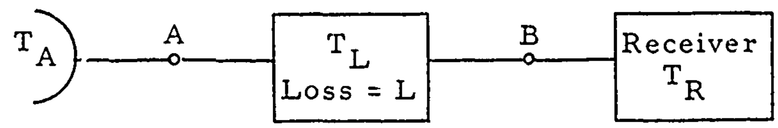
3. EFFECT OF LOSSY ELEMENTS

If the loss is reactive loss, there is no increase in noise power to the system due to the loss. The signal and noise from the antenna are assumed to be attenuated equally, thus preserving the S/N ratio through the loss. However, the system S/N ratio is degraded by no more than the magnitude of the reflected power loss in db. The degradation will be slightly less for practical antenna and receiver noise temperatures.

The worst case is when the loss is entirely resistive. Not only is the loss attenuating the input signal and noise, but it is also generating thermal noise.

4. SYSTEM NOISE TEMPERATURE

The equivalent system noise temperature defines the equivalent system noise power at a given point in the system. Given the signal power at that point, the S/N ratio of the system is determined. The equivalent system noise temperature must be clearly defined for a given point in the system. The following diagram and equation give the system noise temperature for point A and point B. The signal level then must be known at A or at B to determine the S/N ratio.



T_A = equivalent antenna noise temperature
T_L = equivalent loss noise temperature
T_R = equivalent receiver noise temperature
T_{SA} = equivalent system noise temperature at Point "A"
T_{SB} = equivalent system noise temperature at Point "B"
L = loss ratio ($L \geq 1$)

$$T_{SB} = \frac{T_A}{L} + \left(\frac{L - 1}{L}\right) T_L + T_R$$

$$T_{SA} = T_A + (L - 1)\, T_L + L\, T_R$$

$$= L\, T_{SB}$$

A receiver noise temperature of 100° K is assumed, and the equivalent system temperature is shown on Figure 2 as a function of loss for an antenna noise temperature of 25° K. Figure 3 shows the S/N ratio degradation in db normalized to a lossless system for losses from 0 to 1 db.

It is important to note that a 1-db loss causes a 2.55-db S/N ratio degradation for a 25° K antenna and 100° K receiver noise temperatures.

To fully realize a low-noise system, one must not only have low noise components but also use extreme care to minimize losses between the antenna and receiver.

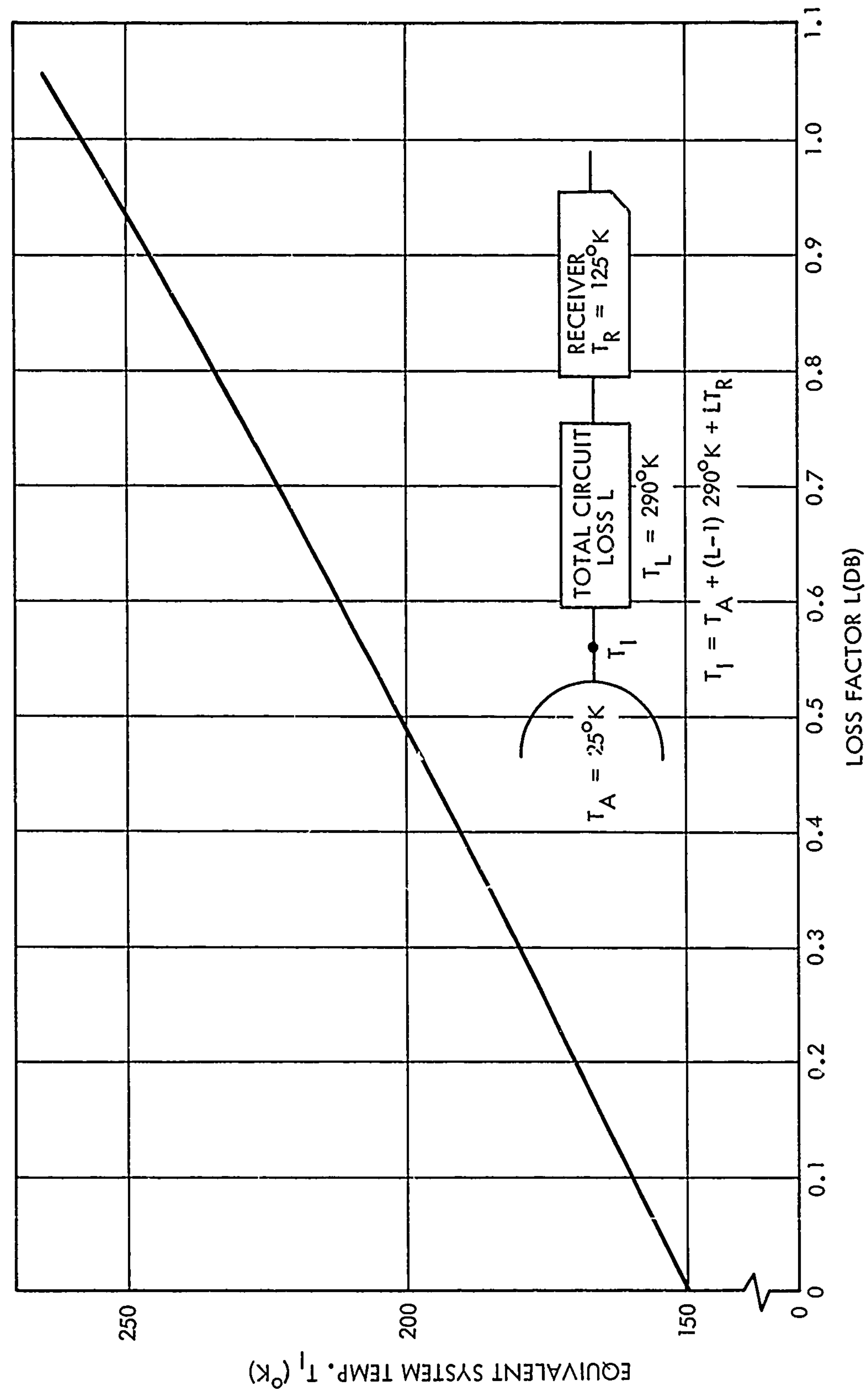


Figure 2. System Noise Temperature Degradation as a Function of Loss for 25^oK Antenna Noise Temperature

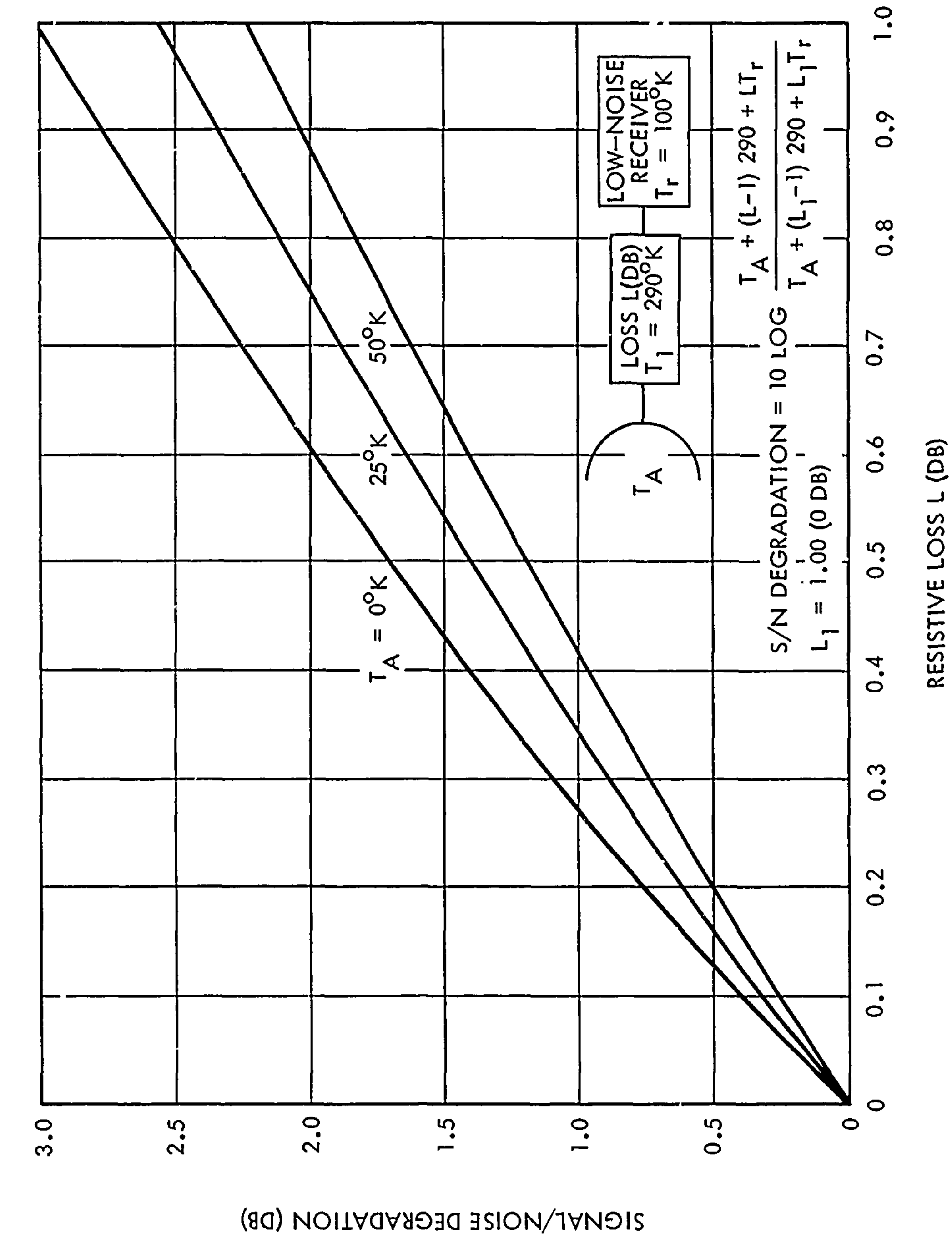


Figure 3. System Degradation by Diplexer Loss

C. LOW-NOISE DEVICES

1. MASER

The term MASER is an acronym standing for Microwave Amplification by Stimulated Emission of Radiation. Since this amplifying device is a fairly recent addition to the communication system's low-noise receiver field, a brief description of its major operating principles and limitations follows.[3,4]

a. Theory

The Maser is often called a quantum mechanical amplifier as its principle of operation is based upon quantum mechanics, which states that the energies associated with the rotation, vibration, or orientation of particles are quantized. These particles may only have discrete energy levels and may change levels while obeying the rules of conservation of energy and momentum. The Bohr frequency-energy relationship shows that a change in energy level is directly proportional to the frequency of the energy radiated or absorbed in the process. The following chart shows roughly the difference in energy levels for various forms of particle energy and the frequency range associated with this energy difference.

E2 - E1 (electron volts)	Frequency (mc)	Form of Particle Energy
4×10^{-7}	100	Orientation of nuclear magnetic moments in several thousand gauss
4×10^{-5}	10,000	Rotational and electronic paramagnetic energies
4	Visible light	Electron motional energy

Figure 4 gives a graphical presentation of the number of particles in each discrete energy state (E_1, E_2, E_3 . . .) of a maser material. Thermal equilibrium requires that the population distribution follow a Boltzmann's distribution curve shown as a dotted line. This line shows that in thermal equilibrium the lower energy states are more populated than higher energy states. The difference in energy states $E_3 - E_2$ and $E_3 - E_1$ correspond to discrete frequencies. When a particle goes

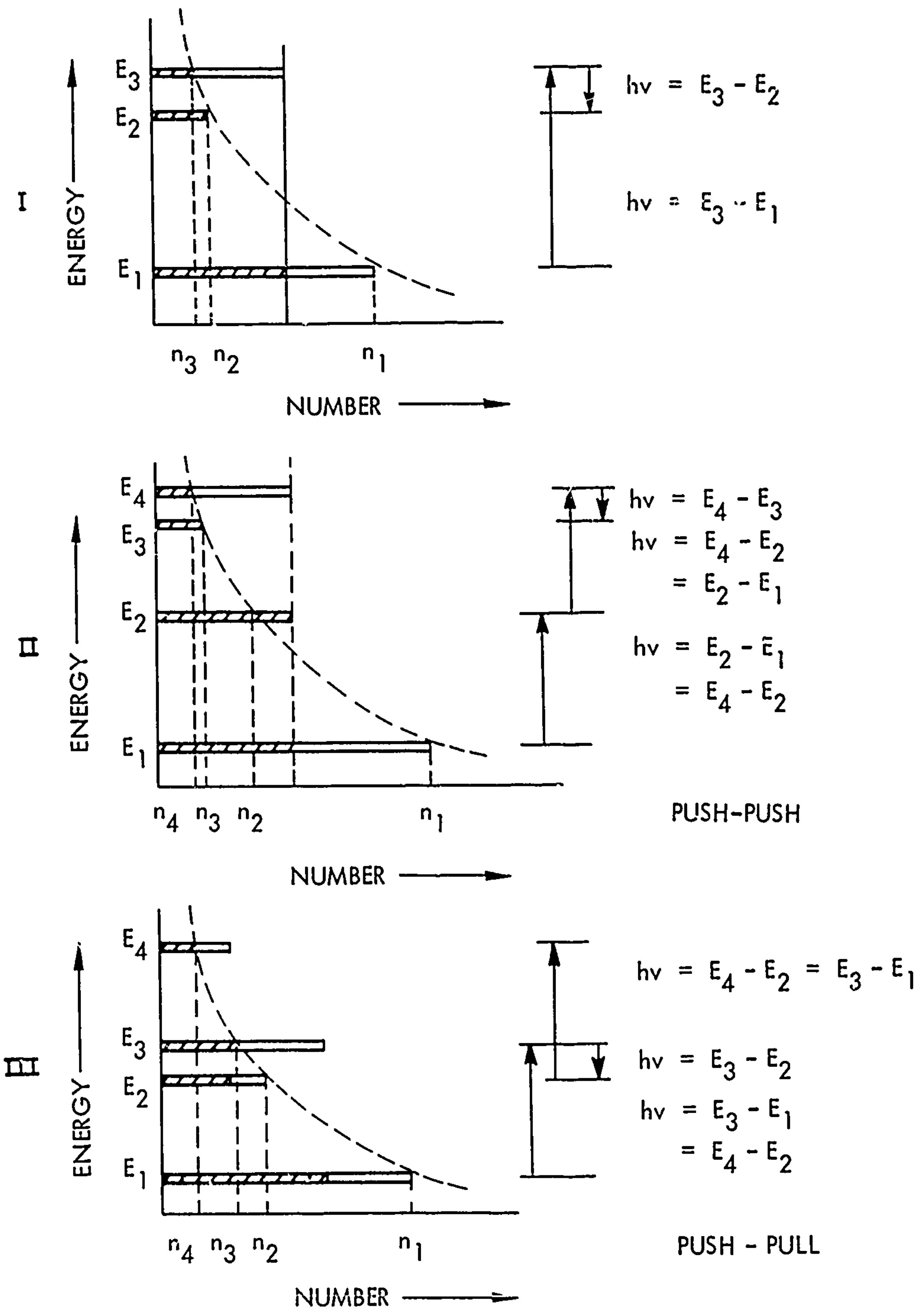


Figure 4. Maser Energy Levels and Pumping Schemes

from a lower to a higher energy state, it absorbs this difference in energy at that associated frequency. When a particle goes from a higher to a lower energy state, energy is emitted at a frequency proportional to the energy change. For purposes of illustration, we shall assume the energy diagrams are for energy states in a crystal. When energy at a frequency equal to $(E_3 - E_1)/h$ (h = Planck's constant) is coupled to the crystal, particles in state E_1 will absorb sufficient energy to move to the E_3 state. When the population of states E_1 and E_3 are made equal by the interaction with this energy, the probability that a particle will go from E_1 to E_3 is the same as the probability a particle will go from E_3 to E_1 due to the interaction with the "pumping" field. The E_3 state is said to be saturated with that particular pump frequency and energy levels. It is possible, therefore, to change the relative population of these energy states by the application of appropriate external energy sources. When the pump is turned off, the crystal returns to thermal equilibrium.

When an interaction field is present, the energy levels involved will tend to equalize their population. Therefore, if a "pump" creates an excess population in state E_3 over the population in state E_2 by raising particles to E_3 from E_1, then the number of particles in E_3 is greater than those in E_2. If a signal frequency $(E_3 - E_2)/h$ interacts with the crystal, it will tend to equalize the $E_3 - E_2$ population. Since there was an excess in E_3 over E_2, particles will lose energy and radiate it at the signal frequency. This is stimulated emission of radiation. The amount of radiation emitted is a function of the interaction field (signal frequency) and the excess particles in state E_3 over particles in state E_2.

For this model, it may be noted that the greater the $E_3 - E_1$ separation, the higher the pump frequency. The lower the signal frequency ($E_3 - E_2$ is small), the larger the excess population in E_3 over E_2 and the more gain and power output. For amplification, there must be a structure to allow the pump and signal frequencies to interact with the crystal, and the energy levels chosen must allow an excess of upper energy particles for interaction with the signal. The time a particle will stay in an excited state is called the relaxation time. This determines how much pump power it takes to keep the crystal "saturated."

Diagrams 4(II) and 4(III) illustrate possible techniques to pump the particles to a higher energy state. The diagram 4(II) uses the pump power to move particles from state E_1 to E_2 and also from E_2 to E_4. This is known as push-push pumping.

Figure 4(III) gives another scheme to obtain excess high-energy particles in the signal frequency bands by pumping particles from E_1 to E_3 and at the same time pulling particles from E_2 to E_4. This creates a larger excess population in E_3 over E_2 for signal amplification. This is commonly called push-pull pumping.

Our example was of a 3 and 4 energy level system. These systems may conveniently be operated for CW amplification. A beam maser is a 2-level maser. To achieve CW amplification with a 2-level maser, the particles must be excited at one location and then brought into the signal interaction structure for amplification.

The signal frequency induces coherent emission, but there is also emission at the signal frequency which is not coherent with the signal and appears as noise. The incoherent emission is a function of the life of the excited particles (relaxation time).

Microwave frequencies are not the only pump sources. Masers have been pumped optically by laser sources. By using push-push and push-pull or multiple frequency pumping, the pump frequency does not have to be greater than the signal frequency, which makes high signal frequency masers practical with microwave pump sources to several hundred gc.

b. Types

The most common types of maser amplifiers for low-noise operation are the beam maser, the solid-state cavity maser, and the traveling wave maser. The last two types are the most widely used as low-noise receiver preamplifiers.

1. Beam Maser

The most common beam maser is the ammonia beam maser. The ammonia beam maser consists of an ammonia source and the associated vacuum pump to keep the vapor pressure below 10^{-5} mm Hg, Diagrams 4(II) and 4(III) illustrate possible techniques to pump the particles to a higher energy state. The diagram 4(II) uses the pump power to move particles from state E_1 to E_2 and also from E_2 to E_4. This is known as push-push pumping.

Figure 4(III) gives another scheme to obtain excess high-energy particles in the signal frequency bands by pumping particles from E_1 to E_3 and at the same time pulling particles from E_2 to E_4. This creates a larger excess population in E_3 over E_2 for signal amplification. This is commonly called push-pull pumping.

Our example was of a 3 and 4 energy level system. These systems may conveniently be operated for CW amplification. A beam maser is a 2-level maser. To achieve CW amplification with a 2-level maser, the particles must be excited at one location and then brought into the signal interaction structure for amplification.

The signal frequency induces coherent emission, but there is also emission at the signal frequency which is not coherent with the signal and appears as noise. The incoherent emission is a function of the life of the excited particles (relaxation time).

Microwave frequencies are not the only pump sources. Masers have been pumped optically by laser sources. By using push-push and push-pull or multiple frequency pumping, the pump frequency does not have to be greater than the signal frequency, which makes high signal frequency masers practical with microwave pump sources to several hundred gc.

b. Types

The most common types of maser amplifiers for low-noise operation are the beam maser, the solid-state cavity maser, and the traveling wave maser. The last two types are the most widely used as low-noise receiver preamplifiers.

1. Beam Maser

The most common beam maser is the ammonia beam maser. The ammonia beam maser consists of an ammonia source and the associated vacuum pump to keep the vapor pressure below 10^{-5} mm Hg, a focuser for separating the upper-state molecules from the lower-state molecules, and a cavity resonator for interaction of the electromagnetic field with the upper-state molecules. The ammonia beam is passed through the focuser, and the emerging beam contains only molecules in the upper state.

The beam of upper-state molecules emerging from the focuser is injected into a cavity. Through the Q of the cavity and the power level, the lifetime of the upper-state molecules is adjusted to be equal to the time of passage of the molecules through the cavity. This interaction time is in the order of milliseconds. If the cavity is tuned to the transition frequency of the molecules, an applied signal field of the proper frequency will induce the molecules to downward transitions, and the energy radiated by the molecules when making the downward transition will amplify the signal.

The design of the ammonia beam maser is not critical. An array of very fine tubes is used as the ammonia source. Because of the large volume of gas released through the beam, a vacuum pump is not sufficient. The vapor pressure must be reduced through cooling the focuser down to liquid-nitrogen temperature to solidify the defocused portion of the beam.

Amplification could also be achieved in a nonresonant waveguide structure, but the length of the waveguide required for appreciable gain would be excessive. Because of the resonant properties of a cavity the effective interaction length can be large though the cavity is rather small. But the inherent positive feedback of a cavity which enables this decrease in the dimensions of the interaction structure also severely decreases the bandwidth for amplification to only a fraction of the molecular linewidth of ammonia. These extremely narrow bandwidths (less than 1 kc) are a great disadvantage for application of the ammonia maser as an amplifier, but they are desirable for use of the maser as a frequency standard or spectrometer.

2. Cavity Maser

The cavity maser is a solid-state amplifier employing either a reflection or transmission type cavity containing the maser material. The design of the solid-state maser is quite critical.

The signal and pump frequency must both interact properly with the maser material in the same cavity. Also the applied magnetic field should be homogenous throughout the material to avoid a linewidth broadening at the expense of gain. Tunability is achieved by varying the magnetic field.

For high efficiency, the ratio of energy in the active material to the energy in the cavity (or the "filling factor") should be close to unity.

It is difficult to satisfy all these requirements simultaneously. With a reflection type cavity, a circulator is needed to separate input from output, while a transmission type cavity will require an isolator in the output line to prevent noise or reflections from the output to enter the cavity and be amplified. Cavity masers, regardless of type, are inherently bilateral devices so that unilateral amplification characteristics must be achieved through the addition of an external unilateral device.

Though the bandwidths achieved with solid-state cavity masers are much higher than those of gas masers, the larger bandwidth of solids cannot be fully utilized with cavities because the internal feedback in cavities which is responsible for the high gain also severely reduces the amplification bandwidth to only a small fraction of the internal linewidth of the active material.

3. Traveling Wave Maser

The traveling wave maser is superior to the cavity maser in that it has a greater gain stability, a larger gain-bandwidth product, and a wider frequency tunability range.

A traveling wave maser consists of a slow wave structure, active material, nonreciprocal elements, magnets, and a pump source. A traveling wave maser is a two-port device and requires no circulator to separate the output from input port. A typical traveling wave maser will consist of a comb type, slow wave structure loaded with a chromium-doped, ruby crystal. As the signal travels along the slow-wave structure, it experiences an exponential growth. Since this growth is bilateral, nonreciprocal elements must be incorporated to pass the forward traveling wave and attenuate all reflected waves. Care also must be taken in matching to reduce regenerative effects. The nonreciprocal elements may be ferrite spheres imbedded along the slow wave structure.

The longer the interaction region, the more gain the maser has. Due to the broad bandwidths achievable with slow wave structures, the instantaneous bandwidth and tunability range of the traveling wave maser is determined by the maser materials.

c. Properties of Masers

1. Noise

Maser amplifiers are inherently low-noise devices because the major noise sources are thermal emission arising from circuitry and material losses, and spontaneous incoherent emission from particles in an excited state.

To reduce the thermal noise contribution, good circuitry design and cooling of the thermal noise sources are required.

The spontaneous emission noise is a function of the number of particles in the lower energy state. Only the excess particle population in the upper state over the lower state contributes to the gain or coherent emission, but all particles in the upper state may contribute to the spontaneous emission noise. The greater the ratios of excited-state population to the lower-state population, the lower the effective noise temperature. The minimum maser noise temperature, assuming all thermal losses are made insignificant, is

$$T_{e(min)} = 4.8 \times 10^{-5} f$$

f = frequency of signal in mc

Figure 16 shows the state of the art in equivalent noise temperature for maser amplifiers. Tables I and II give some typical maser characteristics and the theoretical minimum maser noise temperature as reported in the literature. Since most masers are cooled to liquid helium temperatures, the input line losses contribute significant thermal noise. By proper design, this loss may be made less than 0.1 db (7° K). A typical system noise contribution breakdown is shown in Figure 5.

Table I. Typical Maser Characteristics

	BTL C-Band	BTL S-Band	AIL S-Band	MEC S-Band (Dev.)
Signal Frequency (gc)	6	2.39	2.12 to 2.38	2.295 to 2.388
Electronic Tuning Range (mc)	300	60	400	100
Bandwidth (mc)	120 (Stagger tuned at 20 db gain)	17	20	20
Gain (db)	30	36	27 to 31	30
Reverse Loss (db)	50	>70	>70	>70
Noise Temperature (°K)				
Maser	0.5	0.8	0.8	1.9
Preamplifier	10.5 ± 2	9 ± 3	9.5 to 13.0	8
Dynamic Range (db)				
CW	70	64	75	100
Pulse	100 (0.5 Comp.)	95 (0.5 Comp.)	95 (3 Comp.)	
Refrigerator Temperature (°K)	1.5	1.8	1.8	4.6

Table II. Typical Maser Characteristics and Chart of Theoretical Minimum Maser Noise Temperatures

Manufacturer	Freq. (gc)	Gain (db)	Band-width (mc)	Noise Temp (^{o}K)	
CAVITY MASERS					
Hughes	9	20	10.5	25	Operates 1 to 100^{o}K
	3	30	0.05	2	
MIT	9.300	10.6	70		
JPL	0.960	20	0.6	39	
TRAVELING WAVE MASERS					
University of Michigan	9.65		130		
	35	14	10		
AIL	2.24	30	140	10	

THEORETICAL MINIMUM MASER NOISE

$$T_{e(min)} = 4.8 \times 10^{-5} f$$

f = frequency of signal in mc

f (gc)	$T_{e(min)}$ (^{o}K)
1	0.048
10	0.48
20.8	1.0

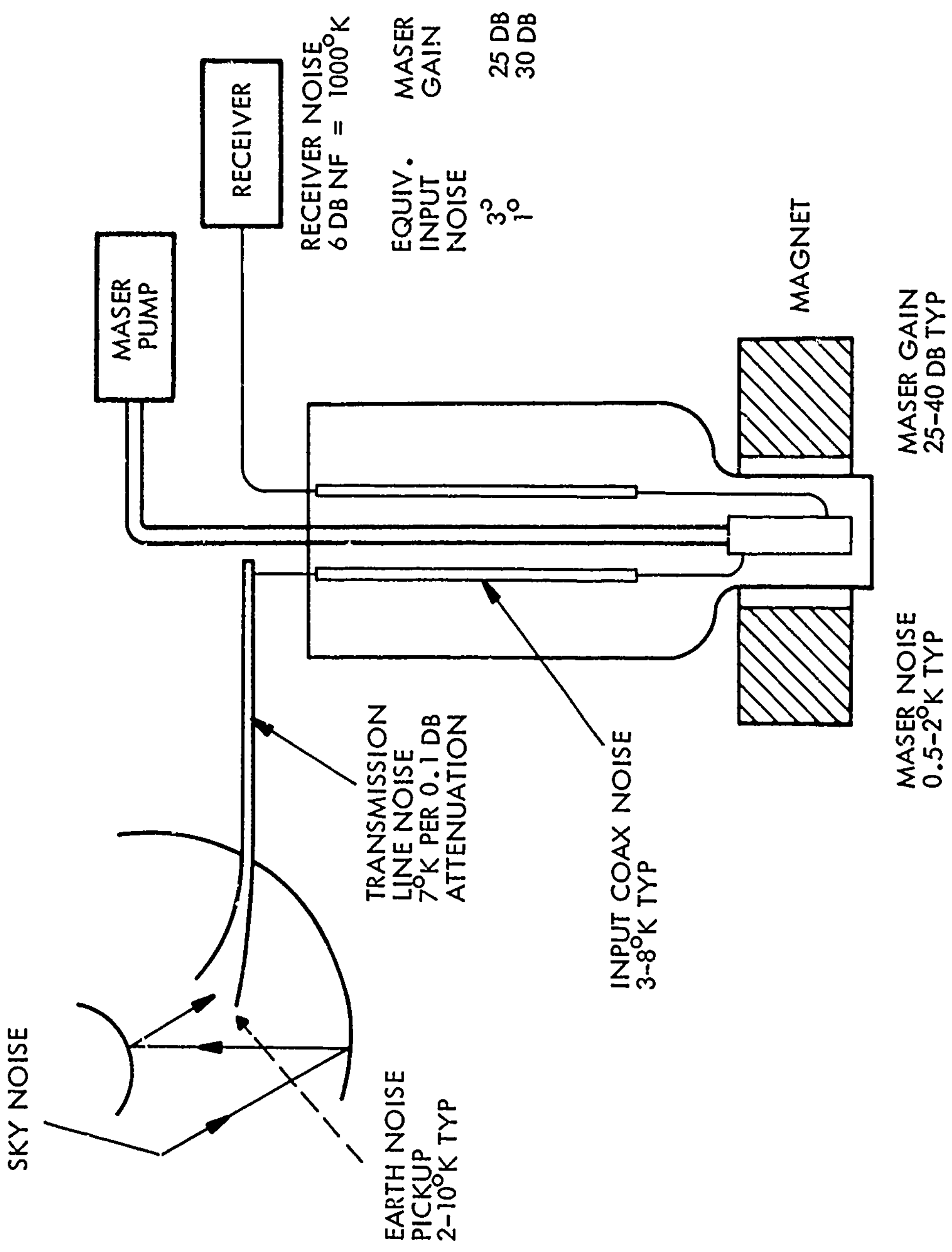


Figure 5. Breakdown of Noise Sources

2. Frequency Range

The maser can operate as a low noise amplifier in the frequency range from 100 mc up. Since its power output at the low frequency end of the spectrum is very low, it is a practical amplifier in the 2-gc region and above.

The upper region of usefulness extends to the light region and covers many technologies. Microwave techniques will be practical to several hundred gc in maser design. Higher power pump sources in this frequency range will be required to overcome the additional losses encountered in the circuits which will be many wavelengths long due to practical construction techniques.

The traveling wave maser is the most tunable of all the masers and has the broadest bandwidth. Tuning is achieved in a solid-state maser by varying the transition levels in the crystal with a magnetic field (the Zeeman effect).

3. Gain and Bandwidth

The gain of a maser is a function of the excess population in the upper state and the length of interaction with the maser crystal. The gain of a maser can be well over 20 db, sufficient to mask second-stage noise.

The bandwidth of a maser is determined by the internal line width of the maser material or crystal and by the interaction circuitry. Gas masers have very narrow internal linewidths and therefore very narrow bandwidths, on the order of kilocycles per second. The major use of a gas maser is in low-noise systems other than communication receivers.

The cavity maser is limited in bandwidth by the microwave circuitry to a fraction of the internal linewidth of the material.

The traveling wave maser bandwidths are now approaching the material limitations. In the 4-gc region, it is now possible to achieve 100-mc bandwidth. In the near future, it will be possible to obtain 200-mc bandwidth at all usable maser frequencies and with some material development, eventually 400- to 500-mc bandwidth in the 4- to

10-gc region will be possible. For a solid-state maser, the voltage gain times the percentage bandwidth is nearly constant.

Future work on materials and broadband coupling circuits will yield larger gain-bandwidth products.

4. Stability

Gain - The gain of cavity masers is very critical and depends strongly on pump power and coupling into the cavity of the signal and pump powers. The traveling wave maser is much more stable and the gain between several amplifiers may be matched conveniently. The JPL S-band maser system gain stability versus time was reported as[6]

Time	Gain Stability (db)
10 seconds	±0.015
10 minutes	±0.06
18 hours	±0.10

Phase - The phase stability of a maser is closely linked to both the gain stability and the bandwidth of the amplifiers. The traveling wave maser is again the most stable with a typical number of $\pm 1^{\circ}$.

The gain and phase stability is sufficient for most applications at present and with some development work could be further improved for future application.

5. Power Output

For maser action to take place, the interacting particles must be sufficiently decoupled from each other to exhibit discrete energy levels. When particles are coupled more strongly, the discrete energy levels broaden out into an energy band structure. The power output is a function of the number of interacting particles. The weak coupling between required particles limits the output power achievable. Also, the output power is a function of pump and signal frequencies as they relate to the excess population achievable in a given maser material. A gas maser utilizing molecular vibrations as energy states has much more power available per interacting particle than a solid-state maser utilizing spin states. However, at X-band, 40 db more power out is obtained with a solid maser compared to a gas maser because the number of interacting particles in a crystal lattice is far greater than the interacting molecules in the gas.

For linear operation, the maser action can use only a small fraction of the excess population to result in coherent emission. Output powers in the order of one microwatt are now achievable. With more work on maser materials, it should be possible to raise this power, but the maser will always be a low-power device. The total power available from the energy transitions used is small. For low-noise communication systems, the power output is not a serious problem, as the maser has sufficient gain to determine the signal to noise in the system, and other types of amplifiers can amplify the received signal to higher output power levels.

6. Cooling Techniques and Associated Equipment

Cooling is required for a maser for several reasons. The pump power to saturate the crystal is a function of the time an excited particle will stay in the excited state (relaxation time). By cooling, the required pump power is lowered. The gain of the maser is determined by the excess of higher energy particles over the lower energy. It is thus possible to achieve higher gain by cooling.

Cooling a maser sharpens the energy distribution curve and permits maser action to be observed above the thermally induced transitions. Since the major noise source at present is thermal, the noise contribution from the material and circuitry losses is minimized, resulting in a very low-noise amplifier.

The majority of masers are cooled by liquid helium to several degrees Kelvin. Masers have been built to operate at about 100° K and eventually may operate at room temperature with no cooling. However, for very-low-noise operation in the next seven years, it appears that the cooled systems will be used. It is possible to cool a maser with liquid neon (27° K) and achieve an 8 to 10° K amplifier noise temperature and a system temperature of 15 to 20° K at S-band.[6]

Present cooling techniques to achieve low temperatures are more advanced than the present hardware to support these techniques. The main problem with closed-cycle, liquid-helium refrigerators seems to be in the short life of the gas compressors.[7] Present life of a compressor to operate the coolers is 250 to 1000 hours. Then new rings must be installed with the possibility of contaminants entering the rest of the system and requiring a complete system purge. The current activity in developing new coolers will undoubtedly result in a compact, reliable, long-life cooling system suitable for cooling masers on antennas in remote areas.

d. Summary of 1970 Projection

Bandwidth - Will be limited by maser material -- 500-mc bandwidth at 8 gc.

Noise - Noise temperature reduced by better circuit design and improved lower loss materials.

Power Out - Limited by maser material. Possibly 10 to 100 microwatts.

Cooling - Efficient, reliable, compact, closed-cycle refrigeration units.

2. LOW-NOISE TRANSISTORS

In the past several years, tremendous progress has been made in high-frequency low-noise operation of transistors. Transistors are well known and only brief remarks will be made concerning the noise mechanisms and circuit aspects of low-noise transistor amplifiers.

a. Noise

The noise sources in semiconductors are fairly well understood. The two main sources of noise in high-frequency operation are thermal noise and shot noise. Thermal noise is developed in the bulk resistivity of the material. In practice this is mainly the base spreading resistance. The shot noise is caused by the random motion of charged carriers. It is a function of the dc current and the bandwidth. The collector and emitter shot noise generators are partially correlated by the α of the transistor.

At lower frequencies 1/f noise (flicker or scintillation noise) and generation-recombination noise become important. The 1/f noise causes the 3 db per octave rise in the noise power per unit bandwidth, as the frequency goes lower. The generation-recombination noise is negligible for good transistors.

In general for low-noise operation, the base spreading resistance (r_b') and the collector current (I_{co}) should be as low as possible and the common emitter current gain (h_{FE}) and the alpha cutoff frequency (f_α) should be as high as possible.

Figure 6 (top) shows a typical noise figure versus frequency curve. The 3 db per octave rise on the low frequency end is due to the 1/f noise and the 6 db per octave rise on the high frequency end is due mainly to the loss of gain of the transistor. The corner frequency, the frequency where the increase in noise power equals the mid-frequency noise power, at the high frequency end is generally approximated by

$$f_{CH} = f_\alpha \sqrt{1 - \alpha_o} \qquad (1)$$

a better approximation is[3]

$$f_{CH} = f_\alpha \sqrt{1 - k\alpha_o}$$

where $k = 1 - \dfrac{2\, r_e}{R_g + r_b'}$

r_e = emitter resistance

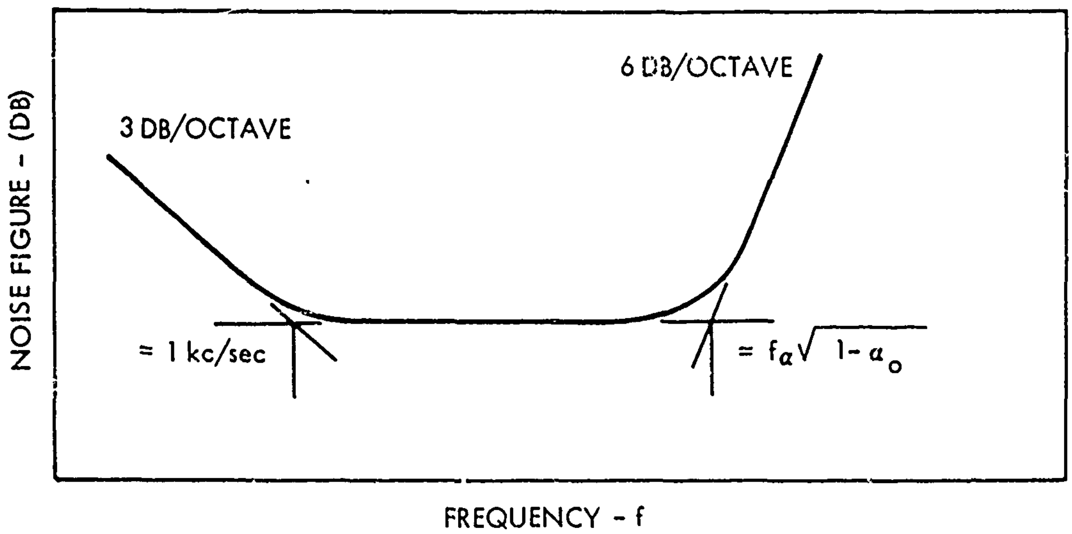
R_g = generator source resistance

r_b' = base spreading resistance

α_o = forward current transfer ratio (common base)

f_α = alpha cutoff frequency.

The lower corner frequency is hard to specify analytically. Figure 6 (bottom) shows that the higher frequency transistors may have higher 1/f corner frequencies and higher plateau noise power levels than good low-frequency transistors. It is possible for the 1/f noise



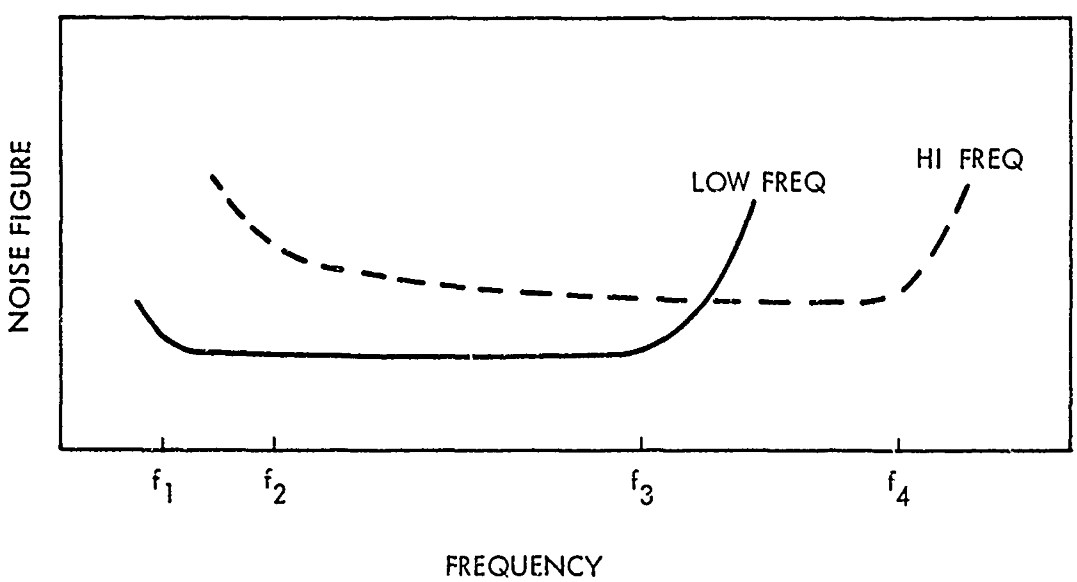


Figure 6. Typical Noise Figure Versus Frequency Curves

corner frequency to be in the order of several mc but more generally is in the 1-kc region.

Excessive 1/f noise is an indication of defects in transistor fabrication. The possibility of the high 1/f noise power at low frequencies cross modulating into a higher frequency band is still a controversial subject but remains a possible source of high-frequency noise.

b. Materials

1. Silicon

Silicon transistors have shown a great improvement in noise performance the last several years. Silicon has inherently a high operating temperature range and power handling ability. This reduces the temperature control needed in extreme environments over systems employing germanium transistors. Photo-etch techniques can be employed in fabricating silicon units which can increase their useful maximum frequency to 10 gc, their seeming practical upper frequency limit. Silicon units currently are not as low noise as germanium units but in general have a larger usable low-noise bandwidth up to the 50-mc region. In the next several years great improvements can be expected in silicon transistor noise and frequency performance.

2. Germanium

The germanium transistors are the lowest noise (above 1 mc) and highest frequency transistors. They are harder to photo-etch than silicon. Tremendous improvements over the present performance are not expected for germanium transistors as they are quite close to their theoretical performance.

3. New Materials

Gallium arsenide transistors have been built with a f_t in the 500- to 600-mc region. Material contamination is still a large problem with gallium arsenide. Concentrated effort on material control and fabrication techniques could produce an order of magnitude improvement in frequency performance of gallium arsenide transistors.

c. Tabulated Performance[8, 9, 10, 11]

Figure 7 shows the noise performance of present transistors. Off-the-shelf transistor cost versus performance is shown below:

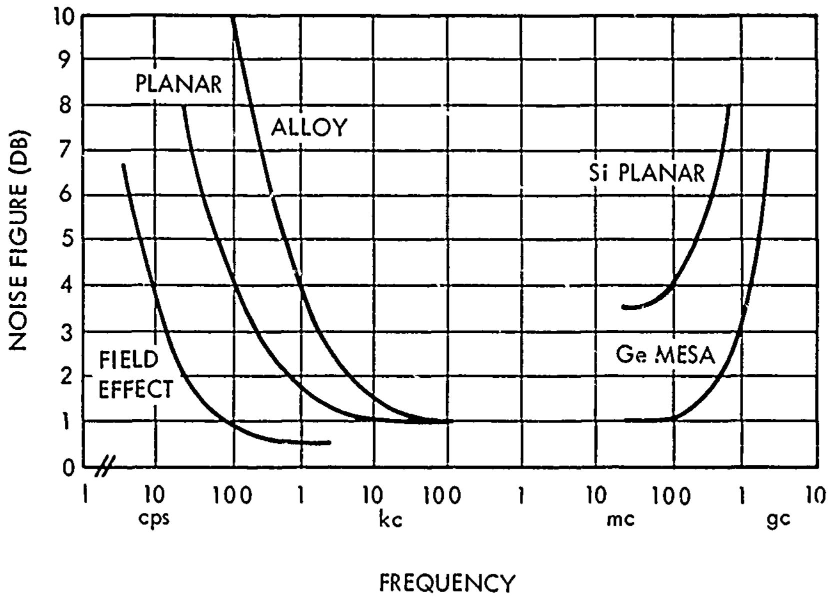
Noise Figure (db) vs Cost

Frequency(mc)	Several Dollars	$50.00	$300.00
450	4.0	2.5	
1000	9.0	6.0	
3000			8

3. PARAMETRIC AMPLIFIERS

Parametric amplification is the term applied to a nonlinear or time-varying energy storage and transfer process in which energy is transferred from a source or "pump" to the signal to be amplified through a nonlinear or time-varying element. Although the basic principle of operation has been known since early in the 19th century,[12, 13] it was not until 1957 that the first microwave parametric amplifier was proposed[14] and demonstrated.[15] This first microwave parametric amplifier used ferrite material as the nonlinear element. Subsequently low-noise parametric operation was attained with semiconductor PN junctions, and with electron beams as the nonlinear energy storage element; then the parametric amplifier field became a very active one.

Some of the main qualitative features of the parametric amplifier are an outcome of the Manley-Rowe relationships between power and frequency in a nonlinear element.[16] Usually the frequencies of interest are: 1) the input signal frequency, 2) the pump frequency (which is usually greater than twice the signal frequency), and 3) the sum or difference frequencies, or both, of the signal and pump. If power is allowed to flow at the difference frequency (the idler), an effective negative resistance is generated with the possibility of unlimited gain and oscillation. If the output is taken at the input frequency we have the usual negative-resistance reflection type of amplifier. If the output is taken at the idler frequency the device is termed a lower-sideband upconverter. If power is allowed to flow at the sum frequency (and not allowed to flow at the idler), there is no negative resistance and



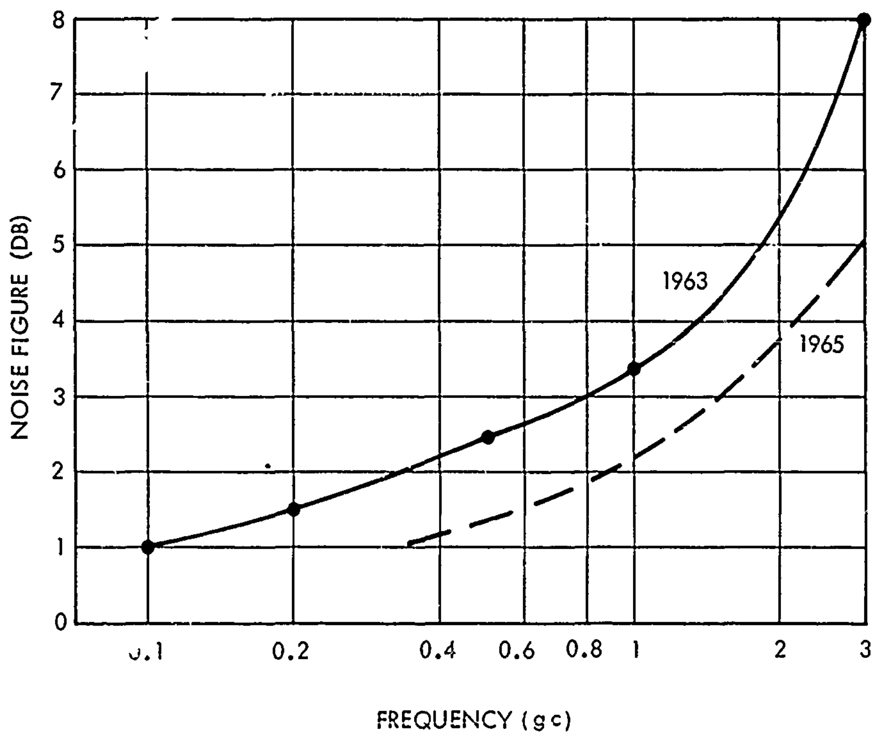


Figure 7. Transistor Noise Figure State of the Art -- March 1963

gain is limited to the ratio of the sum frequency to the input frequency when the output is taken at the sum frequency; a device operating in this manner is termed an upper-sideband upconverter. The features of these various types of parametric amplifiers are tabulated in Table III.

Another important distinction in types of parametric amplifiers is whether they operate in the degenerate or nondegenerate mode. In the degenerate mode the pump frequency is exactly twice the signal frequency and the amount of gain is dependent on the phasing between the pump and signal. In the nondegenerate mode the pump frequency is different from twice the signal frequency and gain is independent of phasing between signal and pump.

a. Diode Parametric Amplifiers

The most widely used type of parametric amplifier uses a highly doped PN junction or varactor diode as the nonlinear element. Since the varactor diode operates with virtually no dc currents flowing, shot noise is eliminated and the main noise source is thermal noise from the spreading resistance of the junction. A figure of merit of the varactor diode called the cutoff frequency is defined in terms of the spreading resistance and the junction capacitance:

$$f_{cutoff} = \frac{1}{2\pi r_s C}$$

The noise performance depends on the quality of the varactor diode, the degree of coupling between the input signal and the diode, and the ratio of idler to signal frequencies. An expression for the single channel noise temperature is

$$T_{noise} = T_{diode} \frac{\left(\frac{\tilde{Q}}{r}\right)^2 + 1}{\frac{\tilde{Q}^2}{r} - 1} \qquad (2)$$

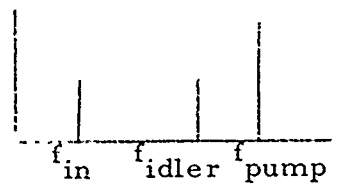
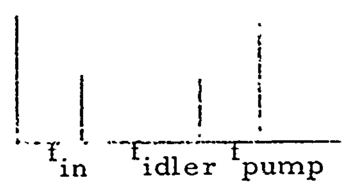
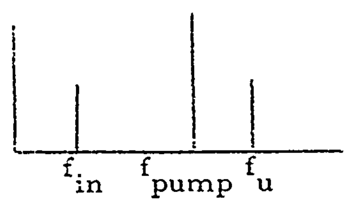
where

$$r = \frac{f_i}{f_s} = \text{ratio of idler to signal}$$

$$\tilde{Q} = \frac{C_1}{C} \cdot \frac{f_{cutoff}}{f_s} = \text{dynamic Q}$$

$$\frac{C_1}{C} = \text{fractional pumped capacitance variation}$$

Table III. Main Types of Parametric Amplifiers

Type	Frequency Diagram	Output Frequency	Effect on Signal Spectrum	Stability	Gain
Reflection		f_{in}	Noninverting	Potentially Unstable	Unlimited
Lower Sideband Upconverter		$f_{idler} = f_{pump} - f_{in}$	Inverting	Potentially Unstable	Unlimited
Lower Sideband Upconverter		$f_u = f_{in} + f_{pump}$	Noninverting	Stable	$G_{max} = \frac{f_u}{f_{in}}$

A plot of this equation is shown in Figure 8. Equation (2) implies that noise performance is improved by having a high cutoff frequency diode with large pumped capacitance variation, by choosing an optimum idler to signal frequency ratio ($r \approx \tilde{Q} - 1$), and by cooling the diode.

The present state of the art in low-noise performance of room temperature paramps indicates attained noise temperatures ranging from 60° K at L-band to 270° K at X-band.[17] These noise temperatures refer to the device alone and do not include contributions from circulators or feed lines. This present-day performance is plotted in Figure 9 where it is superimposed on optimum curves computed from Equation (2). The data appear to be shifted about 50° K above a computed curve corresponding to a diode having f_c = 150 gc.

By reducing the temperature of the varactor diode, it is possible to improve the noise performance of the parametric amplifier. With thermoelectric cooling, the diode temperature has been lowered to about 200° K and the resulting double-channel noise temperature was about 100° K as compared with a noise temperature of 170° K for the same amplifier operated at room temperature.[18]

A great reduction in noise temperature is achieved by cooling the varactor to liquid nitrogen temperature or even to liquid neon or helium temperatures. The best result at liquid nitrogen temperature has been the 20° K noise temperature reported by BTL.[19] Quite recently, parametric amplifiers have been operated successfully at liquid helium temperatures. A noise temperature of 10°K at L-band was obtained at Lincoln Labs,[20] and a 2° K noise temperature at C-band was achieved at BTL.[21] It appears that when operated at the same temperature, masers and paramps have equivalent noise performance. Data on cooled paramps together with computed noise performance for cooled paramps are shown in Figure 10. Liquid-nitrogen-cooled parametric amplifiers have evolved from the laboratory stage and are now operational. A standby amplifier in the BTL Telstar communication link, and the Relay communication link liquid-nitrogen-cooled paramp are among several units presently in the field.

Another important means of improving the noise performance of the parametric am ''fier is to develop varactor diodes with higher

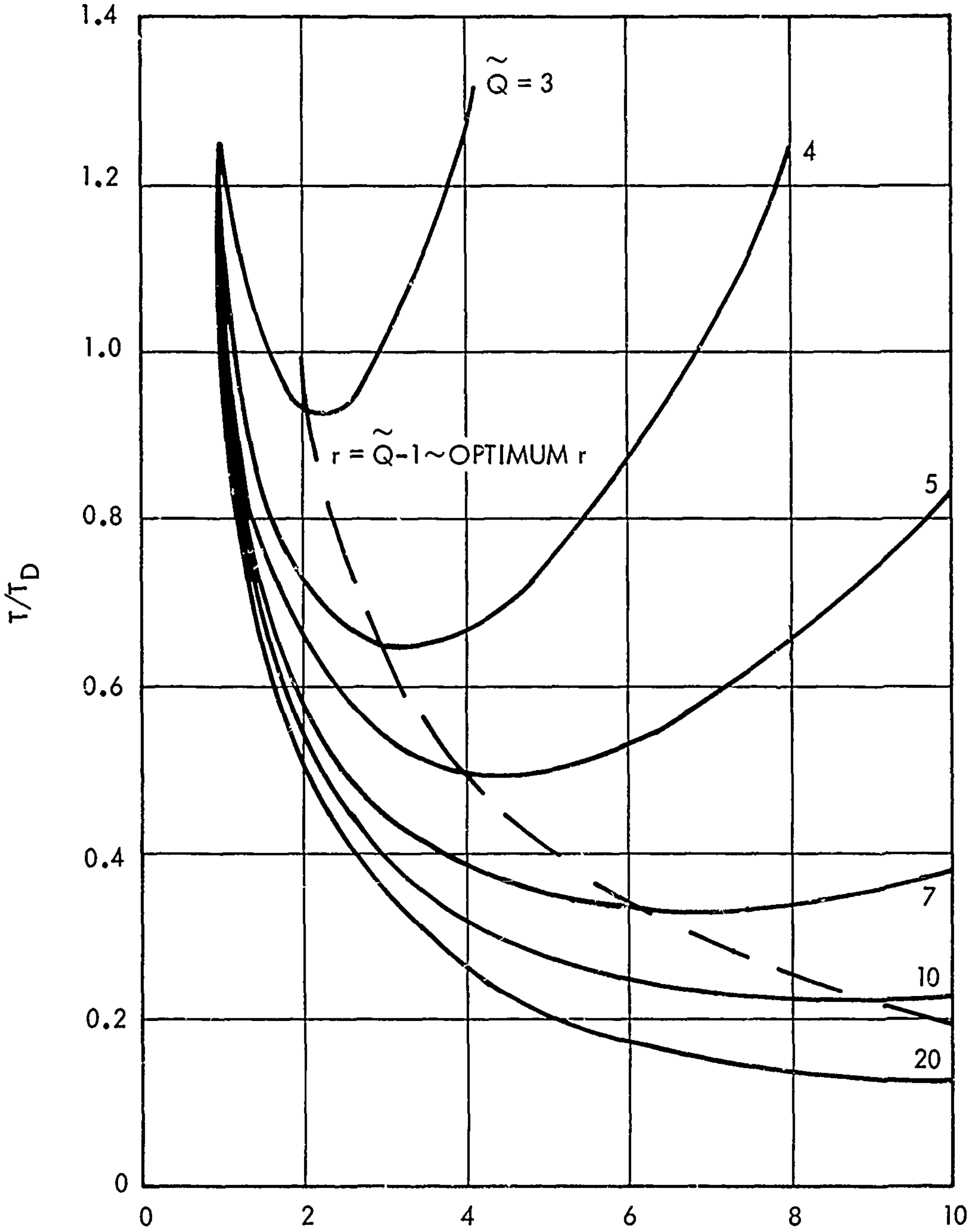


Figure 8. Noise Temperature of Nondegenerate Reflector Type Diode Parametric Amplifier

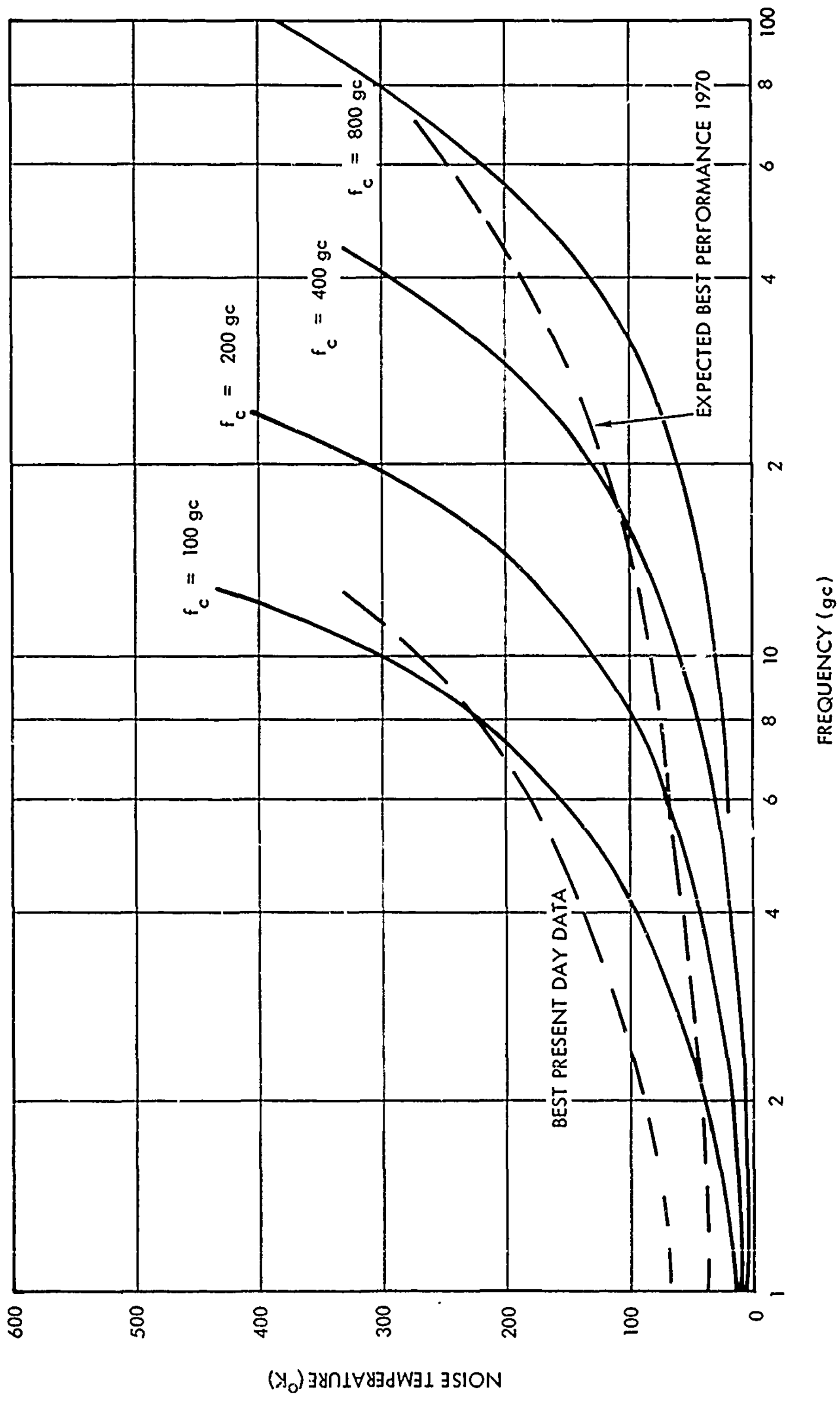


Figure 9. Uncooled Nondegenerate Diode Parametric Amplifier Noise Temperatures: Best Present-Day Data Projected Temperatures: Computed Values

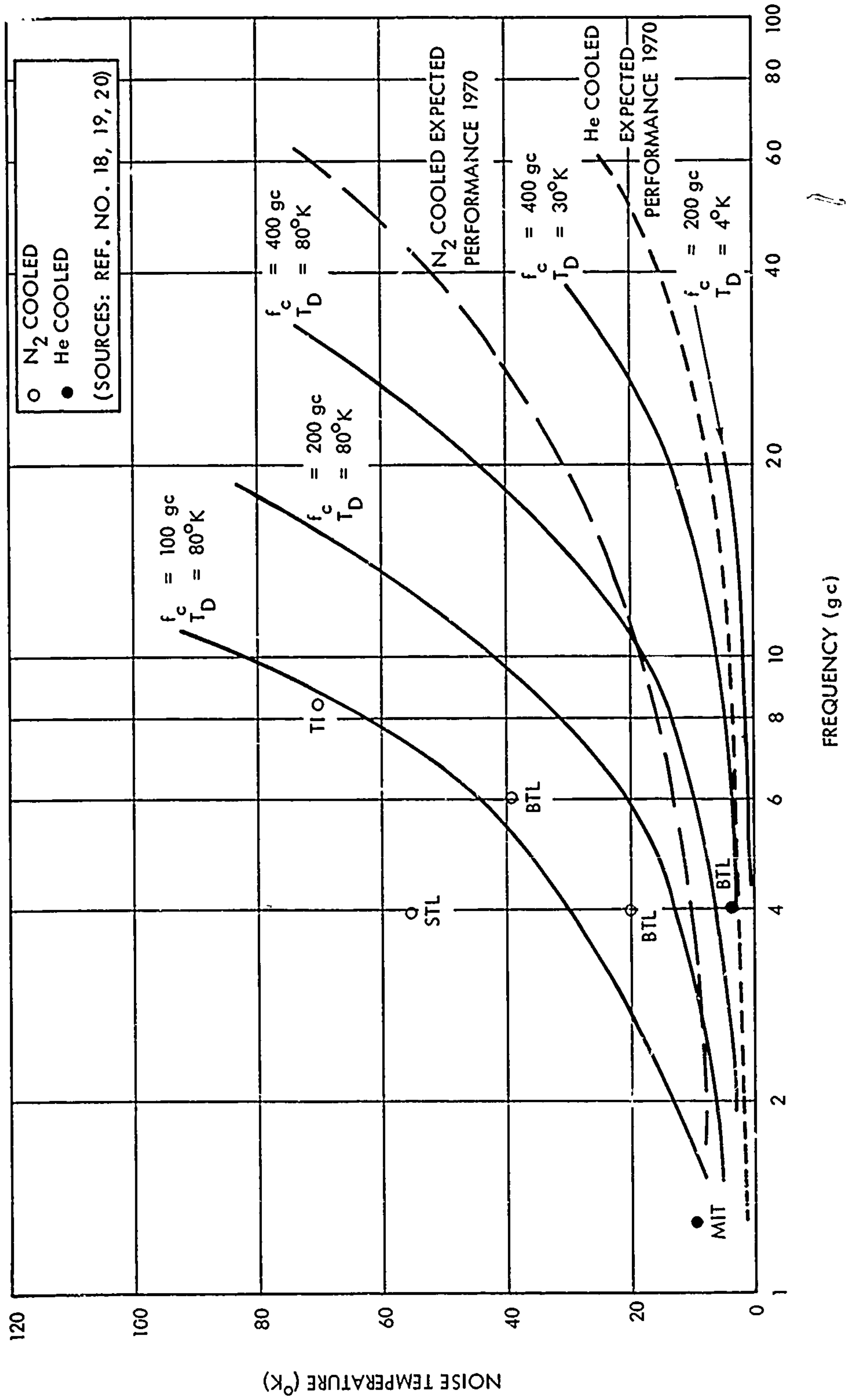


Figure 10. Cooled Nondegenerate Diode Parametric Amplifiers Noise Temperature: Present-Day Data Projected Performance: Computed Values

cutoff frequencies. The best present readily available varactors have cutoff frequencies at about 200 gc at reverse breakdown voltage. A few diodes have demonstrated cutoff frequencies beyond 500 gc.[22] A high cutoff frequency becomes especially important when the signal frequency is high, i.e., X-band or higher. This shows up in the computed noise temperature curves with diode cutoff frequency as a parameter which are plotted in Figures 9 and 10. The semiconductor materials in use or under investigation as varactor elements are: germanium, silicon, gallium arsenide, and indium antimonide. The estimated ultimate achievable cutoff frequency for a silicon varactor diode is 600 gc, for gallium arsenide 1000 gc, and for indium antimonide 2000 gc.[23]

Besides the main performance capability of low-noise operation, other capabilities such as frequency range, bandwidth, tunability, stability, and dynamic range are of interest.

It has been shown that the upper frequency limit is determined mainly by diode cutoff frequency; when the signal frequency becomes greater than about one-fifth the cutoff frequency, noise performance becomes severely degraded. Also as the signal frequency becomes high, the packaging techniques become very significant. The tendency has been to build the PN junction right into the microwave waveguide when dealing with signal frequencies at K-band and beyond. Thus far, 30 and 35.5 gc have been the highest frequencies at which parametric amplifiers have been operated. These were degenerate amplifiers.[24, 25] There is no lower frequency limit for parametric amplifiers and an upconverter to amplify signals in a band of 2 to 50 cps having a 0.16-db noise figure has been developed.[26] Below 500 mc, however, a simpler device such as a transistor usually gives adequate noise performance.

In a parametric amplifier having single tuned circuits, the fractional bandwidth is a function of the signal circuit Q, the idler circuit Q, and the gain:[27]

$$BW \sim \frac{1}{G^{1/2}} \cdot \frac{1}{Q_{signal} + \frac{Q_{idler}}{r}} \tag{3}$$

In the negative resistance reflection paramp, the idler circuit usually has a much higher Q and is the main limitation on bandwidth. The use of multiresonant filter circuit techniques can increase the single tuned bandwidth by a factor of three or more. A very thorough filter analysis has indicated that functional bandwidths of single diode paramps of up to 30 percent are attainable at reasonable gain.[28] The traveling wave parametric amplifier employing several varactor diodes to form a ladder network is theoretically capable of even greater bandwidths; however, the complexity of this device and the variability of the varactor diodes have prevented it from coming into widespread use. Some of the best bandwidths obtained for nondegenerate amplifiers using a single diode have been 450 mc at S-band with 13-db gain,[29] 900 mc at C-band at about 13-db gain;[30] and 900 mc at X-band with 10-db gain.[29]

There have been a great demand and an interest lately in having relatively narrowband paramps easily and rapidly tunable over a wide frequency range. Since the signal circuit can be made quite broadband, it usually is fixed-tuned. The idler circuit which then determines the amplifying frequency can be mechanically or electronically tuned, or alternatively the idler circuit can be fixed tuned and the pump frequency varied to tune the amplifying frequency. Methods of electronically tuning the idler circuit are: changing the bias on the varactor diode providing amplification, having a separate varactor diode as part of the idler circuit and varying the bias on it, or having a ferrite or YIG as part of the idler circuit and changing the magnetic field applied to this element. A method of rapid, wide-range electronic tuning by means of changing pump frequency is to employ a BWO rather than the more usual klystron as a pump source. Some examples of achieved tunability performance follow. An upper sideband upconverter achieved tunable gain over the range 100 to 1800 mc by changing the pump frequency over the range 8500 to 6800 mc (the output frequency was constant at 6800 mc). However, the noise performance was not good over this entire range.[31] A lower sideband upconverter exhibited 10- to 13-db tunable gain over the range 800 to 1150 mc. Tuning was achieved by varying the pump frequency. The output frequency was constant at 4037 mc.[32] A wide tuning range of 2100 to 3100 mc was obtained by changing the pump frequency and mechanically tuning the pump circuit.[33]

A negative resistance reflection type paramp programmed to cover 20 frequency channels over a 1-gc range at X-band has been developed. Tuning was obtained by using a K-band BWO pump source with programmed AFC and power control.[34] A 1-gc electronic tuning range at X-band was also demonstrated at a constant pump frequency with a YIG sphere in the idler circuit as a tuning element.[35] A similar approach is being pursued for S-band paramps.

Diode parametric amplifiers, especially those operating in the negative resistance mode, are intrinsically sensitive and unless care is taken do not exhibit high stability for either gain or phase. The main factors tending to vary the gain and phase characteristics are changes in pump power and frequency due to drift or ripple, and changes in diode temperature. The magnitude of these effects has been estimated.[36] By controlling the diode temperature and by employing closed-loop control on the pump, adequate stability can be attained for most applications. When high gain is required, stability can also be enhanced by employing two cascaded parametric amplifiers, each one having moderate gain. A degenerate paramp in a radio astronomy application employing a phase-locked pump and a closed-loop power control attained a stability of better than 1 percent over a period of several hours.[37] In a novel method of achieving stability, the varactor-diode rectified current (which has to be kept well below 1 microamp for low noise operation) is sensed and a correcting voltage is applied to the pump. A stability of 0.1 percent at 25-db gain over a period of 1/2 hour has been achieved by this method.[38]

For applications which require a linear input-output relation, the dynamic range of the diode parametric amplifier extends from a low-signal power limit imposed by its noise performance to the high-power limit which occurs where the small signal approximations cease to hold. This point is dependent on the diode characteristics and on the magnitude of pump power; it usually occurs at signal powers between -35 and -20 dbm.[27] The diode parametric amplifier begins to saturate at greater signal power levels than the maser or the tunnel diode amplifier.

Now that helium-cooled laboratory parametric amplifiers have demonstrated noise temperatures of a few degrees Kelvin, comparable to the maser, it is to be expected that a large effort will be made to develop operational units covering the 1- to 10-gc range. Closed-cycle helium refrigerators similar to those used for masers will be further developed, especially with regard to improving the life of the contamination-free compressor. Where very good, but not the ultimate, low-noise performance is required, the nitrogen-cooled paramp will play an important role. Also, cooling systems using other inert gases such as neon which boils at 27°K will probably be developed.

To improve the noise performance at X-band and beyond, it is anticipated that a great deal of effort will be made to provide sizeable quantities of diodes with high cutoff frequencies, at say up to 1000 gc for gallium arsenide diodes or even 2000 gc for indium antimonide diodes. The curves in Figures 9 and 10 give anticipated noise performance for both cooled and uncooled diodes with cutoff frequencies up to 800 gc.

It is expected that emphasis will be placed on further development of amplifiers having the pump frequency lower than the signal frequency. This pumping method appears especially attractive for signal frequencies above X-band.

Another area of anticipated development is in the elimination of the ferrite circulator. By pumping two varactor diodes in phase quadrature and having the diodes spaced a quarter-wave length at the idler frequency, it is possible to achieve nonreciprocal amplification.

To make the parametric amplifier more attractive for airborne or satellite application, it will be necessary to increase the lifetime and reduce the size, weight, and power requirements of the pump source. Solid-state microwave sources have progressed very rapidly recently and their widespread application as parametric amplifier pump sources may be expected.

b. The Electron-beam Parametric Amplifier

An electron beam can support both longitudinal oscillating waves, called space charge waves, and transverse oscillating waves, called cyclotron waves. Each of these waves can in turn be classified into slow waves or fast waves, depending on whether their propagation velocity is less or greater than the beam velocity. The excitation of slow waves lowers the beam energy and the excitation of fast waves raises the beam energy. A consequence of this is that beam energy can be used as the energy source for amplification of a signal on an electromagnetic propagating structure properly coupled to the slow wave, as is done in the ordinary traveling wave amplifier. On the other hand, extended coupling between a propagating structure and the fast beam wave will exhibit a periodic exchange of energy rather than an amplification process. This effect has a very useful application in that it provides a means of removing noise from the fast beam wave by transfer to the propagating structure. It is possible to obtain parametric amplification of fast and slow waves of both the space charge and cyclotron types. However, the convenient method just mentioned of reducing beam noise is available only for the fast wave, and in practice successful operation has been attained only by using the fast cyclotron wave.[39] A diagram of the Alder beam type parametric amplifier is shown in Figure 11. The input and output Cuccia couplers act as though they were circulators. The input coupler carries out fast beam noise and couples the input signal to the fast cyclotron wave. A pump signal is applied to a quadrupole structure and as the cyclotron wave passes through the quadrupole structure it is amplified parametrically. The output coupler couples the amplified signal out at the beam and to the receiver. It also aids in establishing stable unidirectional behavior by coupling any receiver noise or reflections to the electron beam which is then carried to the collector.

The main sources of noise in the electron beam parametric amplifier (EBPA) are losses in the couplers and a residue of noise remaining in the beam. The best reported data on the noise performance of the EBPA follows. EBPA's operated at 425 mc and at 780 mc had a double channel noise temperature of 80° K of which 45° K was due to coupler loss, and 35° K due to beam noise.[40] A 1300 mc EBPA had a double channel noise temperature of 35° K.[41] An EBPA operated at 4137 mc exhibited a double channel noise temperature of 58° K of which 33° K was due to coupler loss and 25° K due to beam noise.[42] The noise temperature for this latter device peaked rather sharply

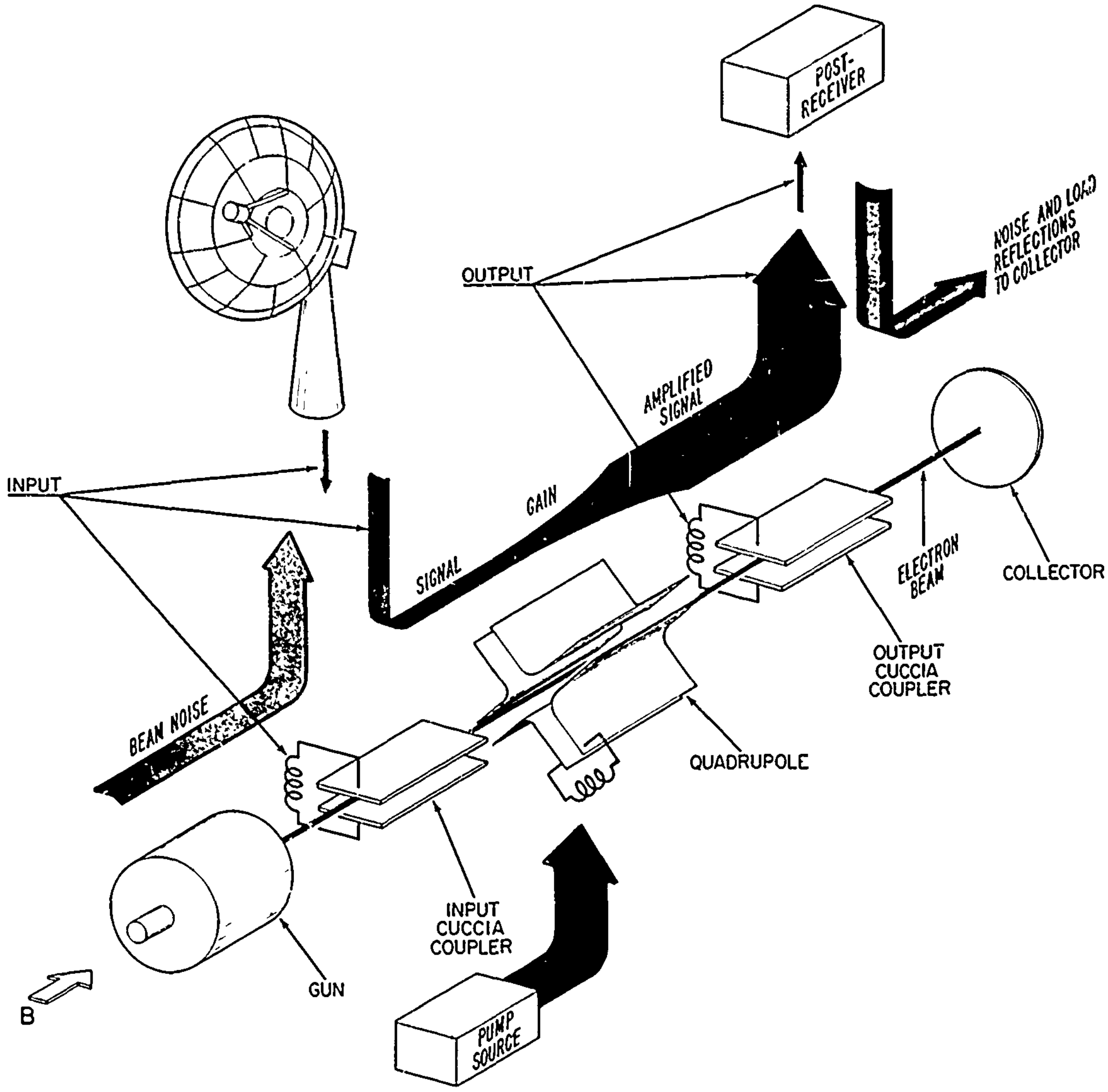


Figure 11. Adler Beam Type Parametric Amplifier

with frequency; the bandwidth for a noise temperature of less than 80° K was about 5 mc. The noise performance of the EBPA is summarized in Figure 12.

While the EBPA has applications where single channel noise performance applies, it is especially attractive in areas such as radio astronomy, or in a phase-locked synchronous pumping system, where the double channel noise performance applies. This is so because until recently the EBPA was a degenerate or near degenerate device. Early in 1963 a nondegenerate EBPA was developed with the signal frequency and the cyclotron frequency at 400 mc, the pump frequency at 2000 mc, and the single channel noise temperature at 170° K.[43] Undoubtedly there will be an improvement in noise temperature. For example, one direct improvement could be obtained by cooling the idler termination.

The EBPA has many desirable features. Its bandwidth is independent of gain and is quite large, being limited mainly by the bandwidth of the couplers. The gain response can be made almost flat over a 10-percent fractional bandwidth. However, the noise bandwidth is somewhat smaller. If desired the tube can be designed to have a narrower, say 20- to 30-mc bandwidth which can be electronically tuned over a 10-percent frequency range. Tuning is accomplished by varying both the pump frequency and the applied magnetic field. In comparison with the diode parametric amplifier, the EBPA has greater phase and gain stability. Computed gain variations with changes in pump power indicate that the EBPA is more stable by a factor of about 6 at 20-db gain.[44]

The main factors affecting the phase stability of the EBPA are variations in the magnet current and the quadrupole voltage. The sensitivity to magnet current is about 4° per milliamp and the sensitivity to quadrupole voltage changes is about 0.06° per millivolt.[45] The dynamic range for linear amplification of the EBPA is about the same as for the diode paramp, i.e., it begins to saturate at input signal levels of about -20 to -35 dbm. Further work to develop the nondegenerate EBPA will probably continue at a fast pace. There is some feeling that the noise performance of the nondegenerate tube can be improved by raising the cyclotron frequency to half the pump frequency.[41] Extension

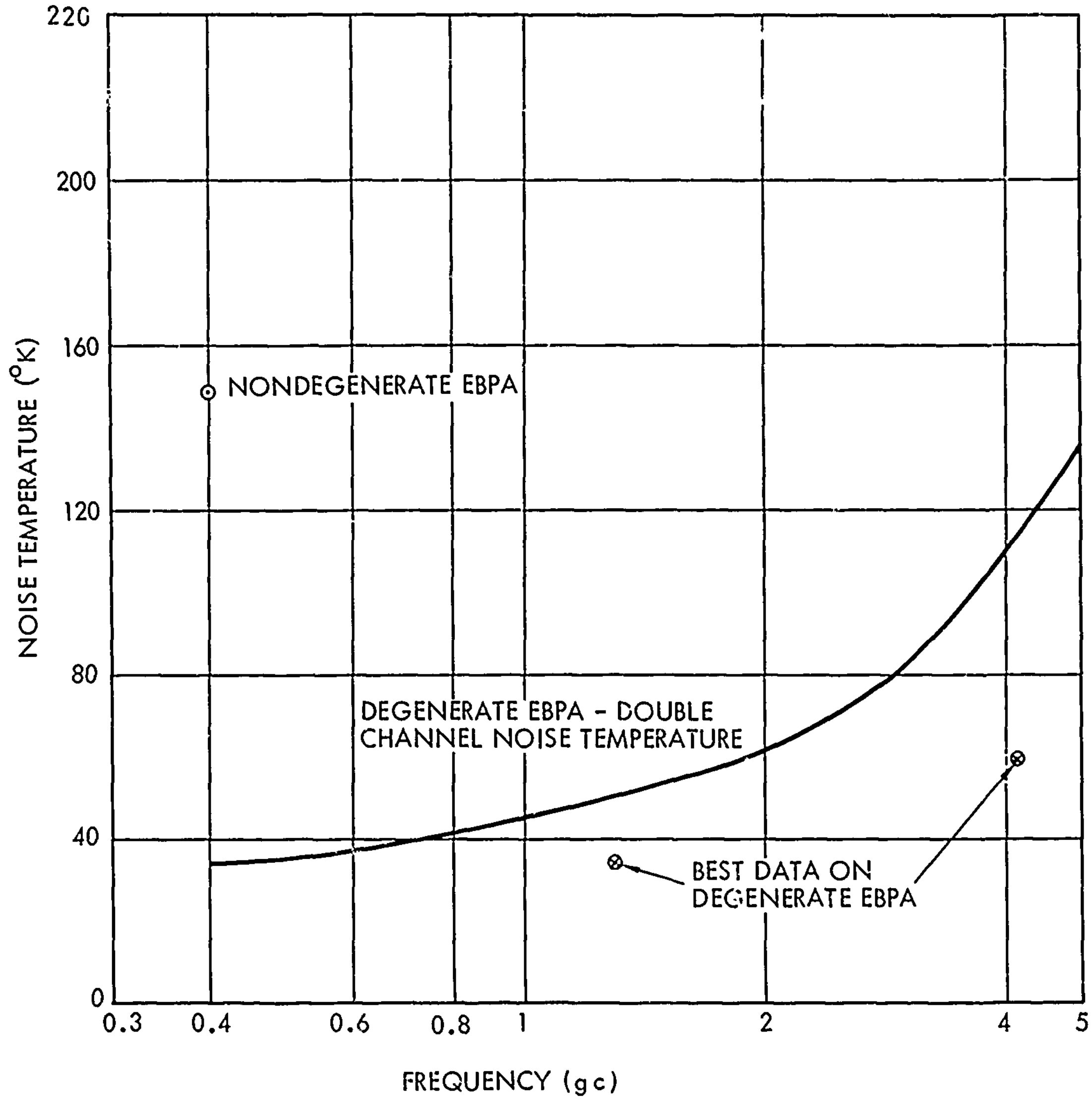


Figure 12. Electronic Beam Parametric Amplifier Noise Temperature Data

of the operating frequency for the degenerate EBPA to X-band or beyond will probably occur. The use of super-conducting magnets to provide the large magnetic fields over extended areas at these frequencies is anticipated.[41]

4. LOW-NOISE TUNNEL DIODE DEVICES

The tunneling phenomena in PN junctions was discovered in 1958 in Japan,[46] and a year later the first low-noise amplifier using a tunnel diode was demonstrated. In a tunnel diode, the doping of the semiconductor materials is so high that when a junction is formed the upper valence band edge in the P-type material is at a higher energy than the lower conduction band edge in the N-type region. When a small forward voltage is applied, electrons from the P-type conduction band are able to pass through the junction potential barrier by virtue of a quantum-mechanical tunneling process and occupy available states in the P-type valence band. As the applied bias is increased, the P-type conduction electrons find themselves opposite the "forbidden band gap" and current decreases. A further increase in applied forward bias again results in increased current as the ordinary forward conduction mechanism comes into play. The behavior of diode current with bias voltage just described together with the resulting "n" shaped I-V curve is diagramed in Figure 13. A point contact junction between a conductor and highly doped semiconductor exhibits a similar behavior. Negative slope in the I-V curve implies that a tunnel diode biased in the region will exhibit a negative resistance. By properly coupling to this negative resistance, unlimited gain can be obtained, even at microwave frequencies.

At frequencies above a few tens of kilocycles, 1/f noise becomes insignificant[48] and the main sources of noise in the tunnel diode are shot noise and thermal noise. An expression for the noise figure of a tunnel diode amplifier is given by:[49]

$$F = \frac{1 + K_d}{\left(1 + \frac{R_s}{R}\right)\left(1 - \frac{f}{f_c}\right)} , \qquad (4)$$

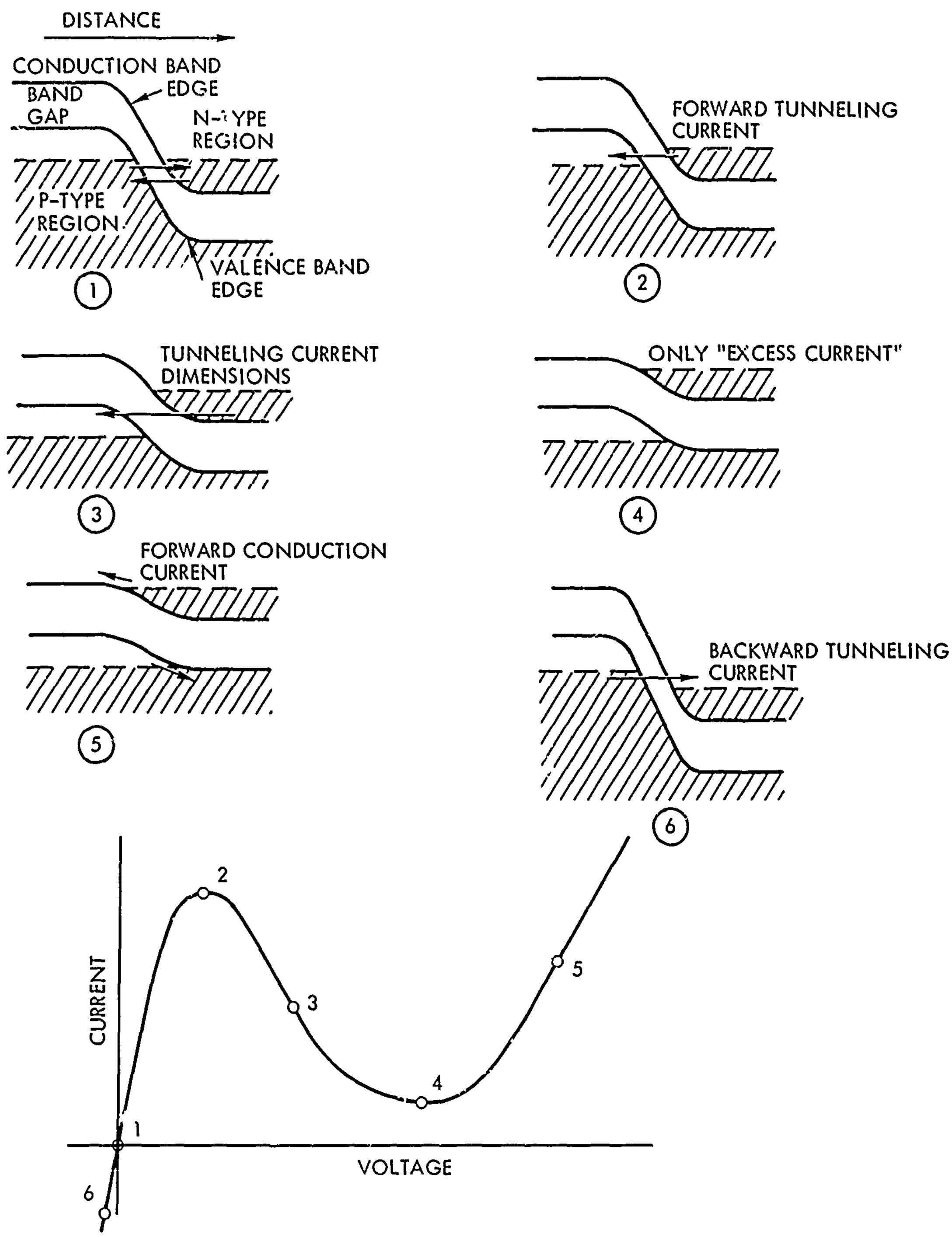


Figure 13. Tunnel Diode I-V Curve

where, $K_d = 20I_oR$ = shot noise factor

I_o = current at bias point

R = magnitude of negative resistance

R_s = diode spreading resistance

f_c = resistive cutoff frequency

Inspection of Equation (4) indicates that to obtain a low-noise figure it is desirable to have a small as possible spreading resistance, a small f/f ratio, and a small shot noise factor, K_d. Equation (4) is plotted in Figure 14 together with data on some noise figures obtained to date. Present day noise figures run from about 3 db at UHF frequencies to about 5 db at X-band. The shot noise factor, K_d, is largely determined by the semiconductor material. Germanium diodes have a K_d of about 1.25 and gallium antimonide diodes have a value of about 0.8. The cutoff frequency is also somewhat determined by semiconductor material through the available negative resistance, R, which occurs in the expression for cutoff frequency:

$$f_{cutoff} = \frac{1}{2\pi RC}\left(\frac{R}{R_s} - 1\right)^{1/2}$$

The cutoff frequency of most present tunnel diodes ranges from 2 to 30 gc. Occasionally a few diodes have exhibited cutoff frequencies as high as 100 gc.

In contrast to the parametric amplifier the upper frequency limit on the tunnel diode amplifier is not determined solely by its diode cutoff frequency. This is so because of stability considerations. While the parametric amplifier may exhibit a negative resistance at a few discrete frequencies, the tunnel diode presents a negative resistance to all frequencies below f_{cutoff} and the possibility of having free or relaxation oscillations is greatly increased. One condition for avoiding these oscillations is that the self-resonant frequency of the diode be greater than the loaded diode cutoff frequency. This self-resonant frequency is given by:[49]

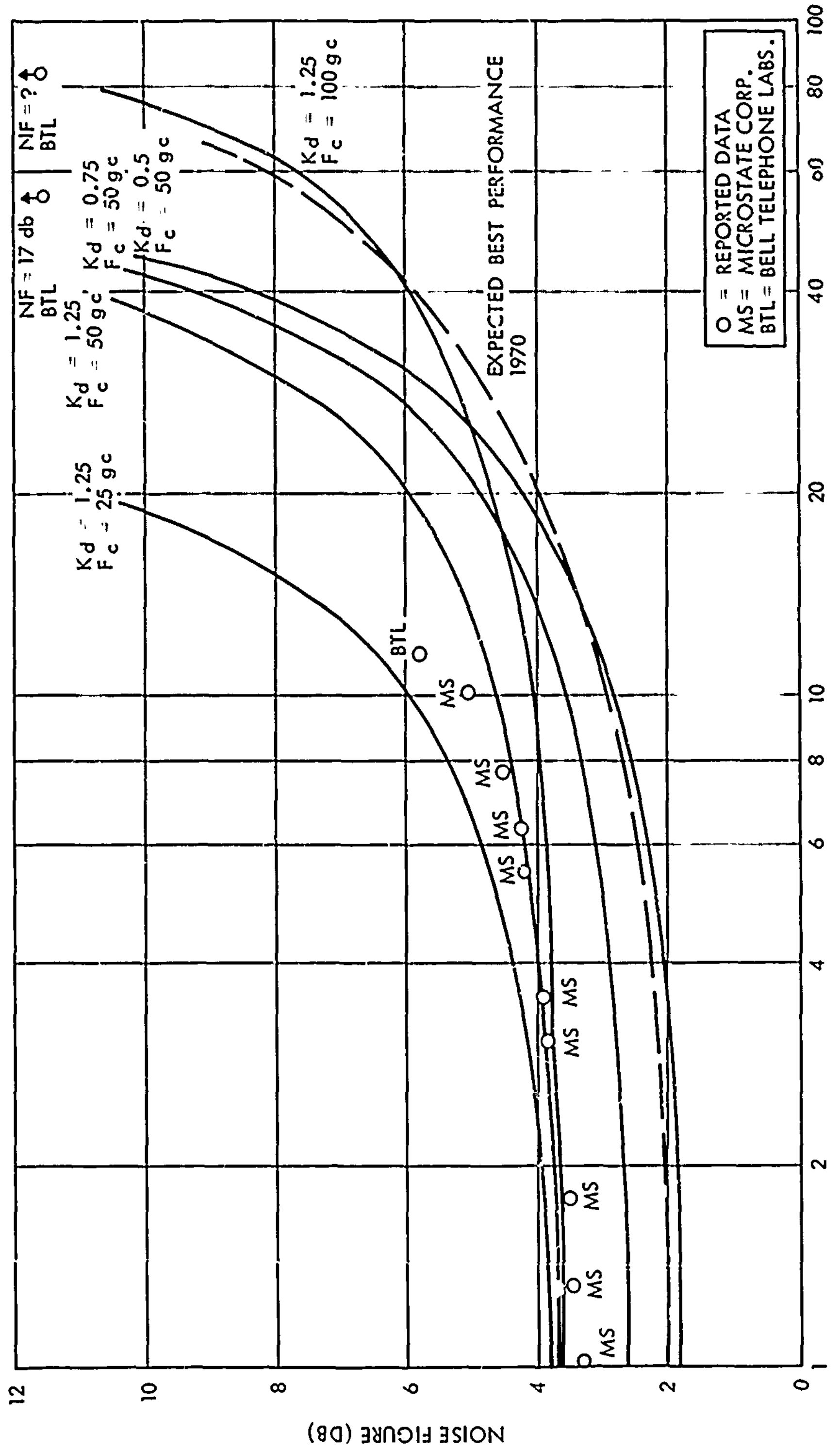


Figure 14. Tunnel Diode Amplifier Noise Figures: Data and Computed Curves

$$f_{\text{self-resonance}} = \frac{1}{2\pi}\left(\frac{1}{LC} - \frac{1}{R^2C^2}\right)^{1/2}$$

where,

L = diode lead inductance

C = junction capacitance

R = magnitude of negative resistance

For typical tunnel diodes $f_{\text{self-resonance}}$ ranges from 1 gc to about 20 gc.[49] By fabricating the tunnel diode junction right into the waveguide structure, an increase in the usable operating frequency can be obtained. By using this technique, tunnel diode amplification has been observed at 85 gc[50] and oscillations have been observed at over 100 gc.[51]

The instantaneous bandwidth capability of the tunnel diode amplifier is greater than that of parametric amplifiers. For the circulator coupled amplifiers, 10 to 15 percent bandwidths are common.[52] In the hybrid coupled device using two tunnel diodes, bandwidths extending over an octave have been obtained.[52] For many applications a wide tuning range rather than an instantaneous bandwidth is desired. By employing a YIG resonator in the tunnel diode circuitry and changing the magnetic biasing field on this element, electronic tuning over an octave frequency band has been accomplished.[53]

The capability of the tunnel diode amplifier for handling received signals at moderate power is somewhat limited. A tunnel diode amplifier using one diode usually begins to depart from linear gain at a signal level of about -40 dbm or less, and is completely saturated at an input of -25 dbm.[54]

In regard to phase and gain stability, the tunnel diode amplifier does not have the problems associated with pump variations as does the parametric amplifier and the main sources of variations in gain and phase are caused by changes in bias voltage and tunnel diode temperature. By controlling the bias voltage and employing a temperature compensating element in the tunnel diode amplifier circuit, adequate stability for most applications can be obtained. TD amplifiers have exhibited a differential phase change between two amplifiers in order of $\pm 1^{\circ}$ per day and an absolute gain stability of ±1 percent per 8 hours.[23]

Another useful low-noise application for tunnel diodes is as a mixer or converter element in converting RF signals to IF frequencies. In contrast to mixers using nonlinear resistance diodes, the tunnel diode mixer can have conversion gain. The local oscillator signal can be either from an external generator as usual or the tunnel diode in the mixer itself can generate its own local oscillator signal. Converters with an output IF of 30 mc have been built with the input frequency ranging from about 150 mc to over 1300 mc.[55, 56, 57] The double channel noise figures for these tunnel diode converters has been in the 2.5-to 5.0-db range,[55, 56] which indicates that a 1-1/2 to 2-db improvement in noise figure over the conventional mixer can be expected.

Because of simplicity, light weight, and low power consumption, it is anticipated that the tunnel diode amplifier will find widespread application as a moderately low noise preamplifier in airborne and space vehicle receivers. A tunnel diode amplifier is planned for the Mariner "C" and "D" space probe vehicles,[23] and has already undergone and passed the shock and vibration qualification tests. Temperature variations in the space environment do not rule out the use of the TD amplifier. With temperature compensation the tunnel diode amplifier operates with negligible change in gain and noise figure in an ambient temperature range of -30° C to $+60^\circ$ C.[23]

In projecting the possible future improvement in noise performance of the TD amplifier it seems that a noise figure of about 2 db or a noise temperature of about 170° K will be attained. Since shot noise rather than thermal noise is the main source of TD amplifier noise, cooling the diode to liquid nitrogen temperatures can only give a modest improvement, say 25 percent, in low-noise performance. The shot noise contribution is mainly a function of the semiconductor material. Work is being done on developing indium and gallium antimonide mixed crystals. It is hoped that these crystals will exhibit a shot noise factor of $K_d \sim 0.5$.[23] Figure 14 indicates that for a K_d of 0.5 and cutoff frequency of 50 gc, one can expect noise figures of 2.0 db for amplifiers through S-band, 2 to 5 db at C-band, and about 3 db at X-band.

5. HIGH-FREQUENCY TRIODES

Moderately low-noise performance at frequencies below about 1 gc can be obtained from the high-frequency triode. To operate at high frequencies these tubes are usually of planar construction, with very close interelement spacings. This is not considered a new low-noise device, most of the development having taken place in the early 1950's, and yet unless one goes to a parametric amplifier it is difficult to obtain better low-noise performance at video or UHF frequencies.

A plot of noise figure vs frequency is shown in Figure 15 for some of the best low-noise triodes.[58, 59] The noise figure is seen to range from a few tenth's of one db at 10 mc to about 6 db at 1000 mc. Optimum source and load impedences are required to obtain such performance. A cascade or grounded grid circuit configuration is usually employed with those triodes. The obtainable bandwidth is about 2 to 10 percent of the signal frequency when optimized for low noise.[58]

6. TRAVELING WAVE TUBES

The principles and design of traveling wave tubes will not be discussed as they are well established in electronic engineering. Suffice it to say that they are inherently broadband and high gain amplifiers.

A former noise theory for traveling wave tubes was based on the assumption that the electrons were accelerated to a few volts as soon as they emerged from the cathode of the tube. The electrons would then be practically uncorrelated, and there would exist a minimum shot noise setting the lowest noise figure for traveling wave tubes at 6 db. However, by correcting the theory to account for the correlation introduced by the action of the potential minimum near the cathode on the beam electrons, there exists no theoretical minimum noise figure. To achieve low-noise tubes, grids are inserted in the cathode region to control the potential minimum of the tube.

The main sources of noise in the traveling wave tube are the shot noise and thermal noise generated by the lossy helix structure. The shot noise is reduced by the low potential area near the cathode and the thermal noise reduced by the use of lower loss traveling wave

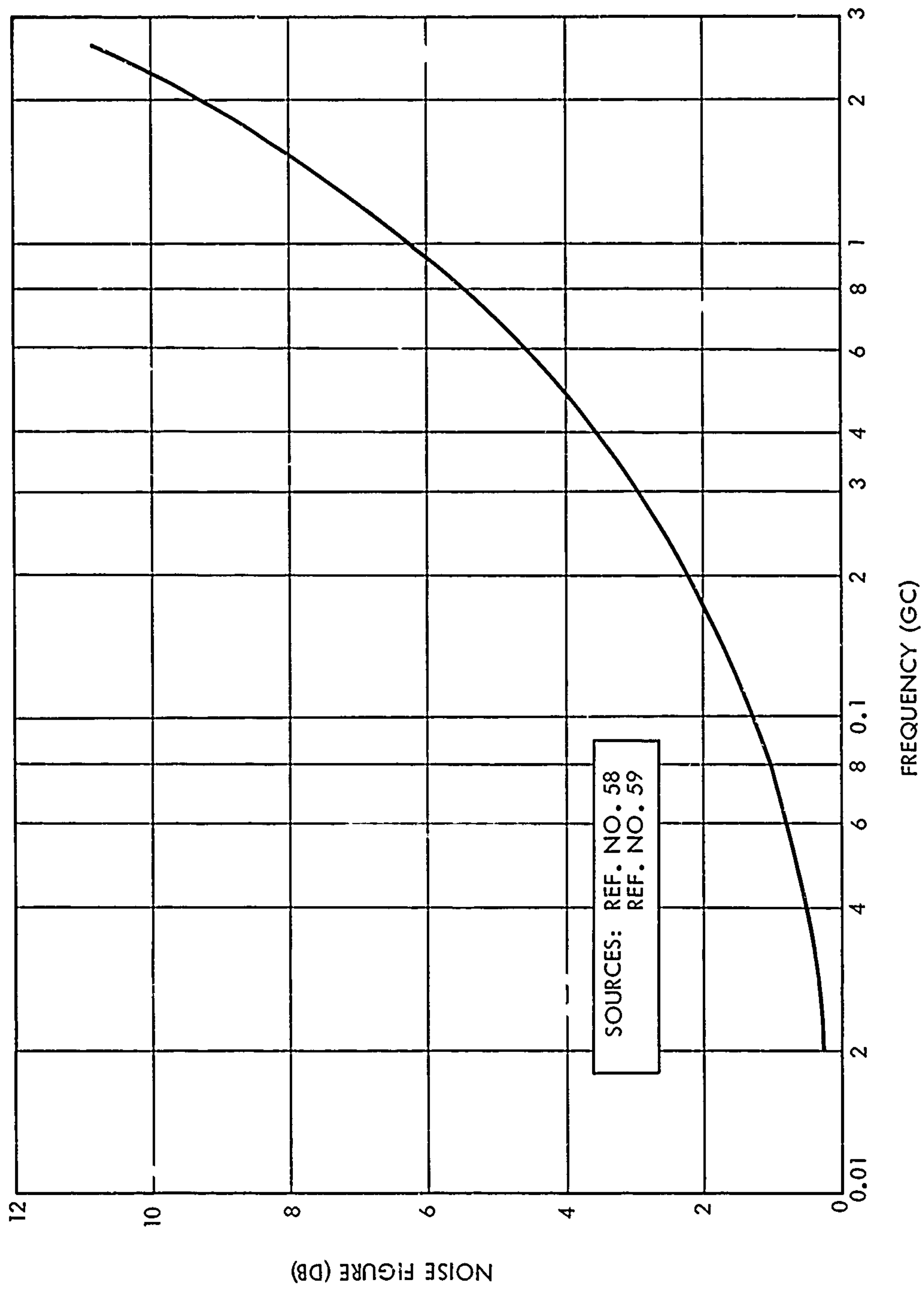


Figure 15. High Frequency Triode Noise Figure

structures and by cooling. Tubes have been cooled to liquid nitrogen temperatures, and the improvement in noise figure agrees with the theory.

Traveling wave tubes require strong focusing magnetic fields utilizing either a solenoid, permanent magnet (PM), or periodic permanent magnets (PPM). Electrostatic focusing may also be used.

Tube manufacturers are now concentrating their efforts on packaging traveling wave tubes for Mil-Spec and space environments. Also low-noise, high-power tubes are being developed suitable for satellite low-noise receivers. Some manufacturers are incorporating the power supplies within the tube package.

Traveling wave tubes are going to be very competitive in the millimeter wave region. Table IV gives some typical data on "off-the-shelf" tubes.[10] Figures 16 and 17 graphically show the state of the art now and projected 1970 performance.[60, 61] In addition to the noise figure, such features as simplicity, stability, high-gain, and broadband characteristics must be considered in evaluating the traveling wave tube as a low-noise amplifier.

Table IV. Low-Noise Traveling Wave Tubes

Manufacturer	Freq (gc)	NF (db)	Gain (db)	Power Out(mw)	Focus	Helix Voltage (volts)
GE	2 to 4	5	20	1	PM	300
MEC	7 to 11	9	30		PPM	
WJ	0.25 to 0.5	3.5	25		1000 gauss	50
	1 to 2	4	25		1000	140
	2 to 4	4.3	25		1000	175
	2.3 to 2.7	3.4	25		1000	175
Huggins	3.4 to 3.6	5.5	25		Solenoid	

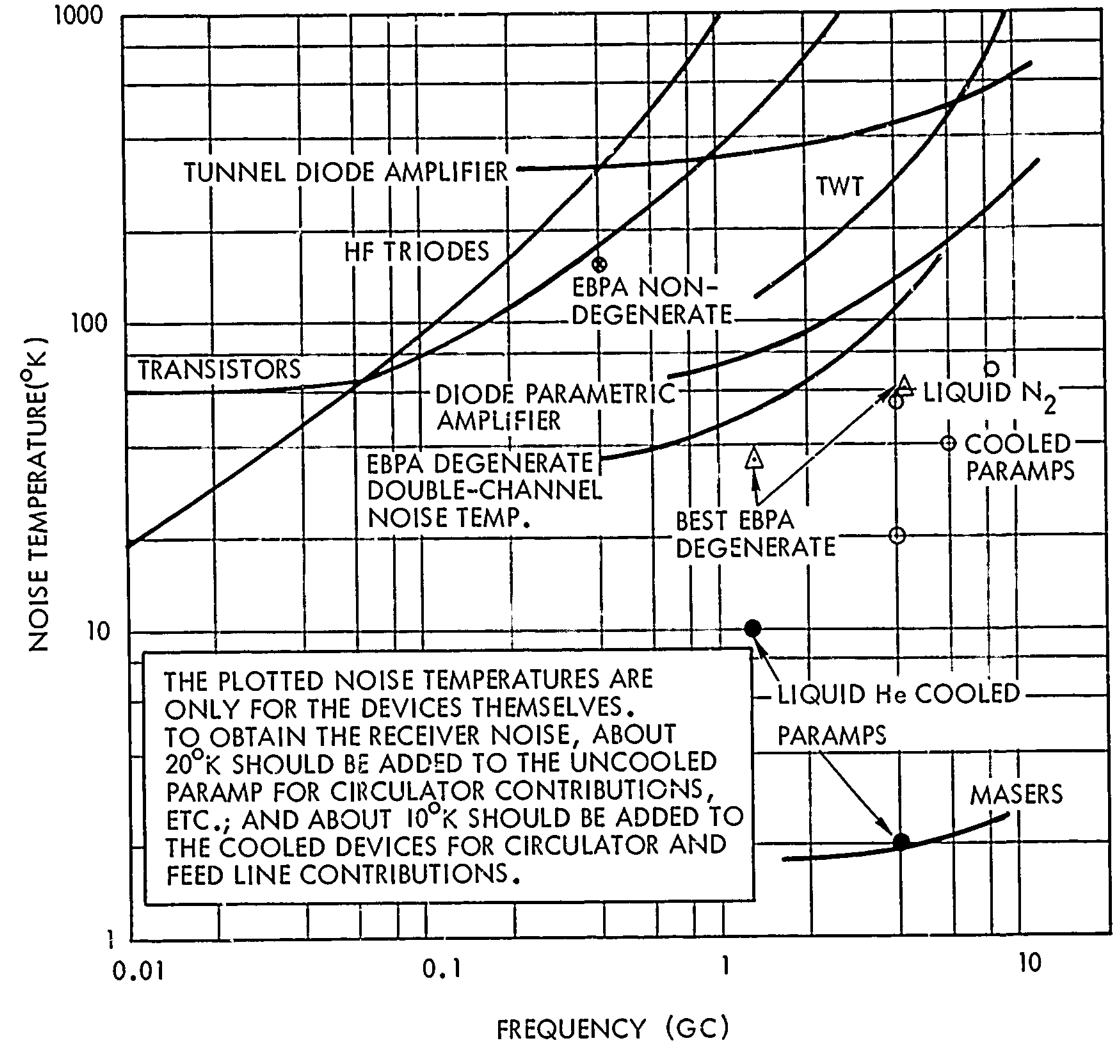


Figure 16. Present Day Status of Low-Noise Devices

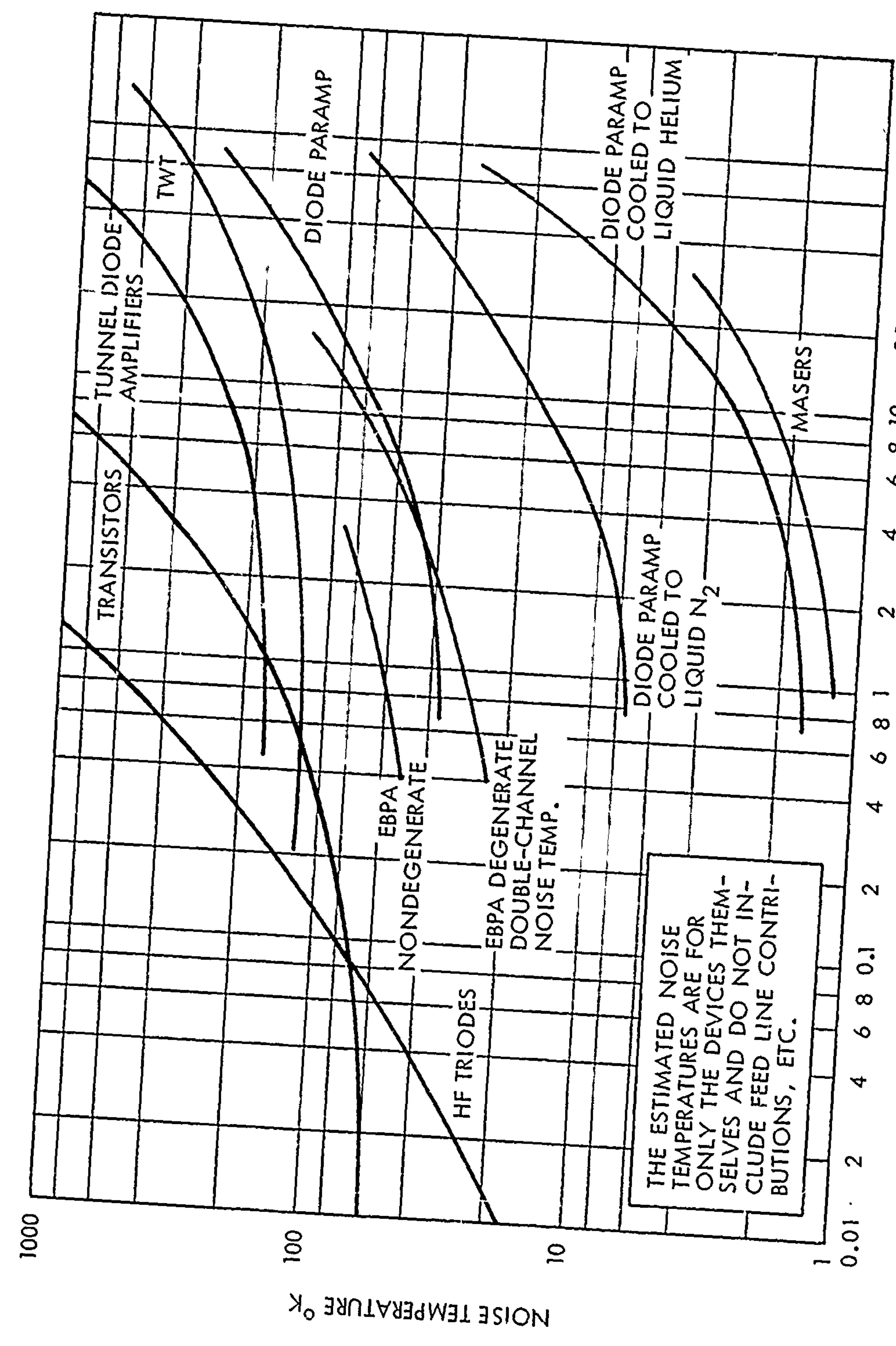


Figure 17. Estimated Low-Noise Device Performance in 1970

D. CONCLUSION

The various low-noise devices which have thus far been discussed separately are compared to each other in the following charts. Figure 16 is a frequency plot of present-day device noise temperatures and Figure 17 shows the projected low-noise device performance to be expected by 1970. These noise temperatures refer only to the device itself and do not include contributions due to input feed lines, circulators, etc.

A comparison of the various low-noise devices with regard to capabilities other than low-noise performance is tabulated in Table V.

In a ground-based receiver, where the design goal is to achieve the best possible noise performance, and size, weight, complexity, and cost are secondary factors, the helium cooled maser and parametric amplifier are the most suitable low-noise devices. At present, the helium-cooled parametric amplifier has not been incorporated into an operational system, and it is therefore difficult to predict for which applications the maser will be preferred over the parametric amplifier and vice versa. With either device the antenna noise and transmission line losses become very significant, and careful design to reduce these contributions is necessary.

In a missile or space vehicle other considerations such as weight and power consumption become extremely important. From this point of view the transistor amplifier below about 1 to 2 gc and the tunnel diode amplifier or converter for the higher frequencies are the most desirable moderately low-noise devices for space vehicle use. A desirable feature of the tunnel diode for space vehicle application arising from the heavy doping concentrations is that it is more resistant to radiation than are varactor diodes or transistors.

In applications where the space vehicle receiver needs better noise performance than that obtainable from the tunnel diode amplifier, the parametric amplifier can be employed. It is anticipated that further development to reduce the pump power requirements and also to improve the efficiency of solid-state pump sources will make the parametric amplifier more attractive for space vehicle applications by reducing its power consumption.

Table V. Chart of Low-Noise Device Characteristics

Device	Frequency Range * (mc to gc)	Maximum Fractional Bandwidth (%)	Electronic Tuning Range (%)	Max. Input Level for Linear Gain (dbm)	Gain Stability %/hr
Traveling-Wave Maser	10 to 100	10	10	-60	1/18
Diode Parametric Amplifier	10 to 14	15	25 (reflection type) 150 (upconverter)	-20	1/8
Beam Parametric Amplifier	350 to 6	15	15	-20	1/24 **
Traveling-Wave Tube	200 to 30	70	-	-20	-
Triode	10 to 1	10	-	0	-
Tunnel Diode Amplifier	10 to 12	70 (hybrid coupled) 15 (circulator coupled)	70	-40	2/24
Transistor Amplifier	10 to 3	>100	-	-10	-

*Only operation above 10 mc is considered.

**Estimated, rather than reported, data.

It is probable that by 1970 there will be in existence several low-noise devices which are unknown at the present time. Seven years ago, the only choices for the microwave receiver first stage was the crystal mixer or, at lower frequencies, the triode; there were no solid-state masers, no parametric amplifiers, no tunnel diode amplifiers, and no low-noise traveling wave tubes. Any projection made at that time as to the present low-noise capability of microwave receivers could not help but be somewhat conservative. Unless advances in the low-noise field are becoming "saturated," it may well be that our projections to 1970 will also prove to be somewhat conservative.

REFERENCES

1. I. I. K. Pauliny-Toth, Proc. IRE, Vol 50, 1962, p. 2483.

2. S. R. Jones and K. S. Kelleher, 1963 IEEE International Convention. [inc]

3. J. F. Price and A. K. Dunlap, "Masers and Lasers," STL Tech. Library Research Bibliography, No. 45, March 1963, pp. 1-178.

4. J. F. Price and A. K. Dunlap, "Masers and Lasers," A Bibliograph ASTIA AD 274 843, pp. 1-161.

5. August W. Rihaczek, "Low-Noise Amplifiers," STL Report No. GM-60-9110.61-66.

6. R. W. DeGrasse, Microwave Electronics Corp., Personal conversation. [inc]

7. M. S. Sellinger, Air Products and Chemicals, Inc., Personal conversation. [inc]

8. H. F. Cook, Texas Instruments, Personal converation. [inc]

9. H. F. Cook, "Causes of Noise," IEEE National Convention 1963 Presentation.

10. Misc. Manufacturer's Published Data. [inc]

11. H. Miyahira, STL, Unpublished results. [inc]

12. M. Faraday, Phil. Trans. Roy. Soc., Vol 121, 1831, p. 299.

13. Lord Rayleigh, Phil. Mag., Vol 16, 1883, p. 50.

14. H. Suhl, Phys. Rev., Vol 106, 1957, p. 384.

15. M. T. Weiss, Phys. Rev., Vol 107, 1957, p. 317.

16. J. M. Manley and H. E. Rowe, Proc. IRE, Vol 47, 1959, p. 2115.

17. Composite data from many published and unpublished sources. [inc]

18. M. Uenohara and R. Wolfe, IRE Trans. on Electron Devices, Vol ED-8, 1961, p. 521.

19. M. Uenohara, BTL, Unpublished data. [inc]

20. C. Blake, et al, submitted for publication to Appl. Phys. [inc]

21. D. C. Hanson, H. J. Fink, and M. Uenohara, 1963 Solid-State Circuits Conference Digest, 1963, p. 54.

22. P. Mullen, Solrac Corp., Unpublished data. [inc]

23. A. Kestenberg, Microstate Electronics Corp., Conversation. [inc]

24. B. C. DeLoach, Proc. IRE, Vol 48, 1960, p. 1323.

25. G. W. Fitzsimmons, Proc. IRE, Vol 50, 1962, p. 2490.

26. J. R. Biard, Proc. IEEE, Vol 51, 1963, p. 298.

27. L. A. Blackwell and K. L. Kotzebue, Semiconductor Diode Parametric Amplifiers, Prentice Hall, 1961.

28. G. L. Matthaei, IRE Trans. on Microwave Theory and Techniques, MTT-9, 1961, p. 23.

29. B. J. Vincent and K. Johnson, 1962 Solid-State Circuits Conference Digest, 1962.

30. C. E. Nelson, Micromega Corp., Unpublished report. [inc]

31. G. P. Shepherd and P. G. Kiely, Proc. IRE, Vol 49, 1961, p. 966.

32. G. L. Matthaei, Proc. IRE, Vol 49, 1961, p. 1703.

33. M. Grace, Proc. IRE, Vol 49, 1961, p. 1940.

34. Hughes Aircraft Co., Unpublished data. [inc]

35. Texas Instruments Inc., Unpublished data. [inc]

36. K. Johnson, Microwave Journal, December 1961.

37. R. F. Clark, et al, Proc. IRE, Vol 50, 1962, p. 1541.

38. B. J. Robinson, et al, Proc. IRE, Vol 48, 1962, p. 1648.

39. R. Adler, et al, Proc. IRE, Vol 47, 1959, p. 1713.

40. R. Adler, et al, Proc. IRE, Vol 48, 1960, p. 256.

41. C. B. Crumly, Zenith Corp., Conversation. [inc]

42. A. Ashkin, Proc. IRE, Vol 49, 1961, p. 1016.

43. R. Adler and G. W. Hrbek, IEEE Trans. on Electron Devices, ED-10, 1963, p. 1.

44. "The Electron-Beam Parametric Amplifier," Zenith Publication. [inc]

45. R. Miller, STL Report No. 5101-0027-MU-000, 1961.

46. L. Esaki, Phys. Rev., Vol 109, 1958, p. 603.

47. K. K. N. Chang, Proc. IRE, Vol 47, 1959, p. 1268.

48. "Tunnel Diode Manual," General Electric Publication, 1961.

49. "Semiconductor Diodes," Microstate Electronics Corp., Brochure.

50. C. A. Burrus and R. Trambarulo, Proc. IRE, Vol 49, 1961, p. 1075.

51. C. A. Burrus, Proc. IRE, Vol 48, 1960, p. 2024.

52. "Low Noise Tunnel Diode Amplifiers ..." Microstate Electronics Corp., Brochure.

53. Watkins-Johnson Inc., Unpublished results. [inc]

54. R. Trambarulo, 1962 Solid-State Circuits Conference Digest, 1962.

55. F. Sterzer and A. Presser, Proc. IRE, Vol 49, 1961, p. 1398.

56. L. E. Dickens and C. R. Gneiting, IRE Trans. on Microwave Theory and Techniques, Vol MTT-9, 1961, p. 99.

57. J. C. Greene and E. W. Sard, Proc. IRE, Vol 49, 1961, p. 350.

58. "Ref. Data for G. E. Planar Triodes," General Electric Publication, 1962.

59. L. H. Von Ohlsen, "Small Signal Performance of the 416B Planar Triode Between 60 and 400 Mc," BTL.

60. W. Robinson, Watkins Johnson, Conversation. [inc]

61. Elliot Rudy, Microwave Electronics Corp., Conversation. [inc]

BIBLIOGRAPHY

1. R. Adler, et al, Proc. IRE, Vol 47, 1959, p. 1713.

2. R. Adler, et al, Proc. IRE, Vol 48, 1960, p. 256.

3. R. Adler and G. W. Hrbek, IEEE Trans. on Electron Devices, ED-10, 1963, p. 1.

4. A. Ashkin, Proc. IRE, Vol 49, 1961, p. 1016.

5. J. R. Biard, Proc. IEEE, Vol 51, 1963, p. 298.

6. L. A. Blackwell and K. L. Kotzebue, Semiconductor Diode Parametric Amplifiers, Prentice Hall, 1961.

7. C. J. Buczec, STL Report No. 8949-0001-NU-000, 1961.

8. C. A. Burrus, Proc. IRE, Vol 48, 1960, p. 2024.

9. C. A. Burrus and R. Trambarulo, Proc. IRE, Vol 49, 1961, p. 1075.

10. K. K. N. Chang, Proc. IRE, Vol 47, 1959, p. 1268.

11. R. F. Clark, et al, Proc. IRE, Vol 50, 1962, p. 1541.

12. H. F. Cook, "Causes of Noise," IEEE National Convention 1963 Presentation.

13. B. C. DeLoach, Proc. IRE, Vol 48, 1960, p. 1323.

14. L. E. Dickens and C. R. Gneiting, IRE Trans. on Microwave Theory and Techniques, Vol MTT-9, 1961, p. 99.

15. L. Esaki, Phys. Rev., Vol 109, 1958, p. 603.

16. M. Faraday, Phil. Trans. Roy. Soc., Vol 121, 1831, p. 299.

17. G. W. Fitzsimmons, Proc. IRE, Vol 50, 1962, p. 2490.

18. M. Grace, Proc. IRE, Vol 49, 1961, p. 1940.

19. J. C. Greene and E. W. Sard, Proc. IRE, Vol 49, 1961, p. 350.

20. D. C. Hanson, H. J. Fink, and M. Uenohara, 1963 Solid-State Circuits Conference Digest, 1963, p. 54.

21. J. V. Jelley, Proc. IEEE, Vol 51, 1963, p. 30.

22. K. Johnson, Microwave Journal, December 1961.

23. S. R. Jones and K. S. Kelleher, 1963 IEEE International Convention. [inc]

24. J. M. Manley and H. E. Rowe, Proc. IRE, Vol 47, 1959, p. 2115.

25. G. L. Matthaei, IRE Trans. on Microwave Theory and Techniques, MTT-9, 1961, p. 23.

26. G. L. Matthaei, Proc. IRE, Vol 49, 1961, p. 1703.

27. R. Miller, STL Report No. 5101-0027-MU-000, 1961.

28. I. I. K. Pauliny-Toth, Proc. IRE, Vol 50, 1962, p. 2483.

29. P. Penfield, Jr., and R. P. Rafuse, Varactor Applications, The MIT Press, 1962.

30. J. F. Price and A. K. Dunlap, "Masers and Lasers," STL Tech. Library Research Bibliography, No. 45, March 1963, pp. 1-178.

31. J. F. Price and A. K. Dunlap, "Masers and Lasers," A Bibliograph ASTIA AD 274 843, pp. 1-161.

32. Lord Rayleigh, Phil. Mag., Vol 16, 1883, p. 50.

33. August W. Rihaczek, "Low-Noise Amplifiers," STL Report No. GM-60-9110.61-66.

34. B. J. Robinson, et al, Proc. IRE, Vol 48, 1962, p. 1648.

35. G. P. Shepherd and P. G. Kiely, Proc. IRE, Vol 49, 1961, p. 966.

36. F. Sterzer and A. Presser, Proc. IRE, Vol 49, 1961, p. 1398.

37. H. Suhl, Phys. Rev., Vol 106, 1957, p. 384.

38. R. Trambarulo, 1962 Solid-State Circuits Conference Digest, 1962.

39. M. Uenohara and R. Wolfe, IRE Trans. on Electron Devices, Vol ED-8, 1961, p. 521.

40. B. J. Vincent and K. Johnson, 1962 Solid-State Circuits Conference Digest, 1962.

41. L. H. VonOhlsen, "Small Signal Performance of the 416B Planar Triode Between 60 and 400 Mc," BTL.

42. M. T. Weiss, Phys. Rev., Vol 107, 1957, p. 317.

Appendix XXI. SOLID-STATE TRANSMITTERS

1. INTRODUCTION

Semiconductor techniques have advanced to the point where it is now practical to use them for UHF and microwave transmitters in space and airborne telemetry applications. Solid-state transmitters have many advantages over those using thermionic tubes:

- Long operating life
- No heating (cathode) power required
- Operation at low voltages
- Small size and weight
- Extreme ruggedness
- Higher DC to RF efficiency.

Because semiconductors used in transmitters operate at temperatures close to ambient, the circuits are relatively cool, contributing greatly to increased circuit-component reliability.

The limitations of semiconductor devices that are presently being minimized include:

- Variations in characteristics among units of the same type
- Temperature sensitivity
- Power output limitations at high frequencies
- High-frequency-cutoff limitations.

The purpose of this appendix is to discuss semiconductors used in solid-state transmitters, and solid-state transmitter design. The objectives are to explain briefly the characteristics of semiconductors, the current design approach in transmitter design using semiconductors exclusively, the limitations of this approach, and the future trends in hardware technology and transmitter capabilities. This discussion is based primarily upon the application of semiconductors to transmitters operating in the frequency range from approximately 50 mc to X-band.

2. SEMICONDUCTOR DEVICES

a. General

Semiconductors, from which the active elements used in solid-state transmitters are made, are materials whose electrical resistivities fall between those of insulators and of conductors. By way of comparison, the resistivities of conductors are of the order of 10^{-6} ohm/cm; of insulators greater than 10^{6} ohm/cm; and of semiconductors of 10 ohm/cm. Two members of the carbon family, germanium and silicon, in their crystalline form, are classified as semiconductors and are used in solid-state devices. More recently, gallium-arsenide has become popular as a solid-state device. A bimetallic compound, the gallium-arsenide is used primarily for diode and varactor applications at present. Research is continuing on the gallium-arsenide, which is not considered practical for transmitter applications in the present state of the art.

From the viewpoint of life expectancy and satisfactory operation, the ambient temperature is probably the most important single factor. In this respect, silicon devices are preferable over germanium. Silicon units operate satisfactorily at temperatures of up to about 175°C, whereas operation with germanium is probably impossible above a junction temperature of about 90 to 100°C. Both types can be operated at low temperatures, e.g., -55°C. A comparison of operational parameters of both types shows that silicon devices are the more suitable for spaceborne operation. This judgement is based primarily on the temperature and radiation effects on the semiconductor material. Therefore, this appendix makes reference to silicon devices only.

b. Transistors

The most common transistor is a three-terminal device. It is the active element used for solid-state transmitter applications. In contrast to the vacuum tube, the transistor is a low-voltage device (typically 10 to 70 volts), and can handle less power than vacuum tubes. The power gains of VHF power transistors are also limited to about 15 db or less, as compared to about 15 to 30 db for thermionic devices.

The performance of power transistors at high frequencies is seriously limited by several factors:

Collector-to-base capacitance

Base spreading resistance

Power dissipation

Finite transit time of minority carriers through the base

Conflicting requirements of small base width for high frequency operation and large width requirements for power handling capability.

The characteristics of high-frequency transistors can be summarized by the following parameters:

- f_{max} = maximum frequency of oscillation
- f_t = gain-bandwidth product
- r'_b = base-spreading resistance
- C_c = output collector capacitance
- P_d = total device dissipation
- Thermal resistance

The transistor has a power gain of unity at the maximum frequency of oscillation (f_{max}). The power gain increases at 6 db/octave as the frequency is decreased from f_{max}. The various parameters of the transistor are related to f_{max} in the following manner:

$$f_{max} = \sqrt{\frac{f_t}{8\pi r'_b C_c}}$$

As noted earlier, f_{max} defines the high-frequency tuned amplifier's upper performance limit, and may be considered one of the most important parameters in the design of transistor RF amplifiers. The maximum available gain at any other frequency can then be determined from

$$\text{Gain (db)} = 20 \log \left(\frac{f_{max}}{f_{operating}} \right)$$

For the low-level stages, i.e., transistors capable of 500 mw and below, the present state of the art for silicon transistors is about f_{max} = 1.5 gc. As the power handling capability of the transistor is increased, f_{max} falls off quite rapidly. At present, transistors capable of handling more than 10 watts dissipation have an $f_{max} \cong 300$ mc. For this reason, power transistor amplifiers are presently operated below 200 mc when power gain, power output, and collector efficiency are all considered.

A curve representing power output versus frequency for a single transistor is shown in Figure 1. This represents what may be considered the present state of the art in transistor technology.

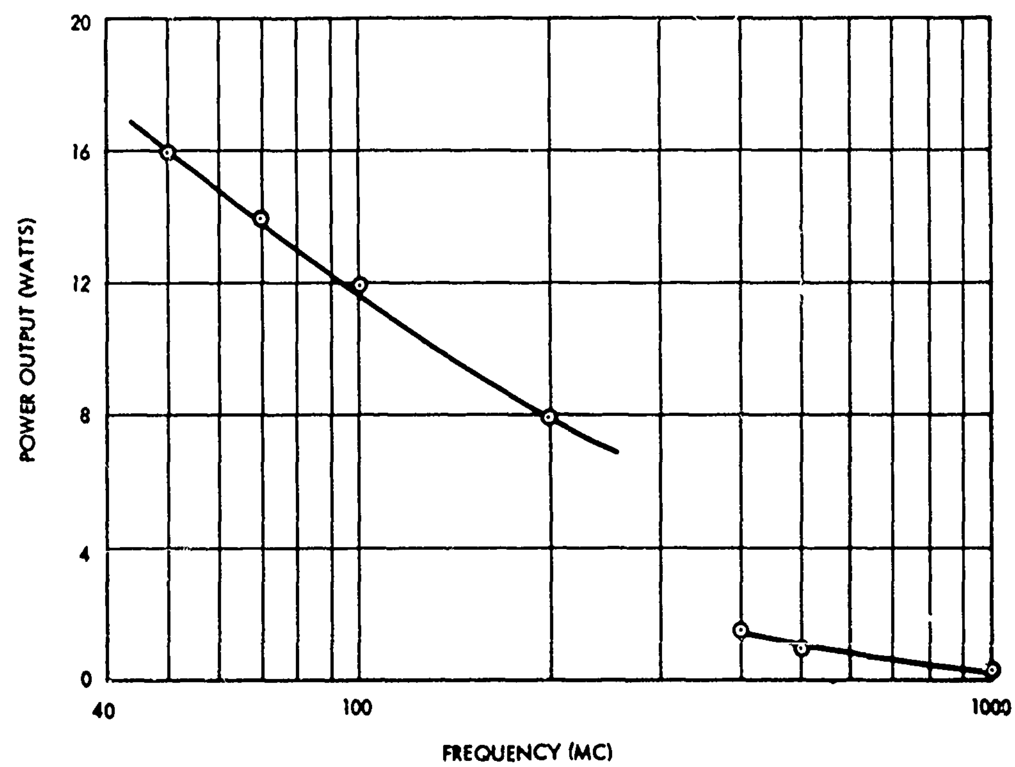


Figure 1. Single Transistor Performance, State of the Art 1963

c. Varactors

The varactor is basically a p-n junction diode that is back-biased in its operation. The junction diode, which consists of a single p-n junction, is similar to its vacuum-tube counterpart in that it is characterized by the property of allowing current to pass easily in only one direction, within a given operating voltage range. If a positive potential is applied to the anode, the diode is said to be forward biased, and will exhibit a fairly low value of forward resistance. On the other hand, a voltage of the opposite polarity provides a reverse bias, and very little current flows. Should this reverse voltage exceed a certain value, called the breakdown (V_B), normal operation ceases. This breakdown voltage is also referred to as the avalanche or Zener voltage. When this voltage is exceeded, reverse current increases rapidly. Avalanche breakdown occurs when the electric field across the junction produces ionization because resulting high-energy carriers collide with valence electrons. Zener breakdown appears to be a "field emission" phenomenon, the strong electric field in the junction region pulling carriers from their atoms.

When a reverse bias is applied, the majority carriers leave the vicinity of the junction, thus leaving behind an area that is "depleted." The thickness of the space-charge region widens with increasing reverse bias. This is a capacitive effect, because C = dq/dv, and the phenomenon is referred to as a transition, depletion, or space-charge capacitance. It is this capacitive effect that is optimized for varactor diodes.

The varactor can be defined as a semiconductor diode with a useful, nonlinear, reverse-bias capacitance. Varactors can be used at frequencies up to and including the microwave region. Their useful frequency range is limited primarily by a series-equivalent resistance. The incremental capacitance of the varactor is expressed as

$$C(v) = \frac{k}{(V + \phi)^{\gamma}}$$

where k = a constant dependent on the varactor geometry

v = the reverse voltage applied

ϕ = the "contact" potential of the p-n junction

γ = nonlinearity exponent derived from doping geometry (1/3 for linear grading and 1/2 for abrupt).

Most manufacturers specify the breakdown voltage, V_b, C_{min} (capacity at breakdown voltage), the nonlinearity exponent, and the series equivalent resistance.

Another equation used in characterizing varactors is the cutoff frequency,

$$f_c = \frac{S_{max} - S_{min}}{2\pi R_s} = \frac{1}{2\pi R_s C_{min}}$$

The cutoff frequency may be considered a figure of merit for the varactor. It relates the series resistance element and the nonlinearity of junction capacitance.

For transmitter applications, the varactor is used for efficient frequency-multipliers and modulators. They are capable of handling watts of power as indicated in Figure 2. This figure simply shows that a single varactor is capable of handling the indicated power levels; it does not imply that for every point on the curve an all solid-state transmitter can be built to deliver the indicated power levels.

3. TRANSMITTER DESIGN

a. General

Solid-state transmitters and vacuum tube transmitters are similar in many respects. Both require a frequency generating source (e.g., a crystal oscillator), a modulator, power amplifiers, frequency multipliers in some cases, and a dc power source. If we consider each of these requirements as a component, then it is only the arrangement of the components and the magnitudes of power or voltages required from these components that differ between solid-stage transmitters and their vacuum-tube counterparts.

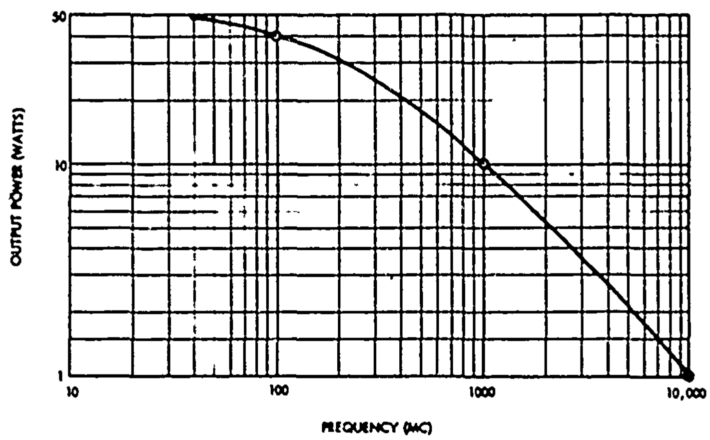


Figure 2. Single Varactor, the State of the Art

b. Oscillators

For space and airborne transmitter applications, the stability of the output frequency is usually controlled by a crystal oscillator. Crystals are presently manufactured to operate from the audio range up to approximately 200 mc. Properly designed, crystal-controlled oscillators will oscillate only at the mechanically resonant frequency or its multiples, which are called overtones. The performance characteristics of a crystal depend on the particular mode of vibration selected, on the crystal's dimensions, and on how it is cut with respect to the different faces. Crystals operating in the thickness-sheer mode can be made to operate on odd overtones, commonly used ones being third-mode or fifth-mode. Seventh- and higher-overtone oscillation is possible but seldom used at the present time. From about 10 mc, crystals are usually operated in overtone modes. Most telemetry transmitters use overtone crystals.

The stability of transistor crystal oscillators can be maintained within ±0.0025 percent over the temperature range of about -35 to + 85°C. For most applications, this stability is quite satisfactory. If better stability is required, the temperature of the air surrounding the crystal unit is maintained at some higher value than the highest ambient temperature expected. More sophisticated oscillator design can also obtain better stability.

Power output levels from transistor oscillators are usually maintained at about 0 dbm. The signal is then amplified by a buffer amplifier stage that provides isolation as well as amplification.

c. Modulation

The most popular method of modulation now used for solid-state transmitters is angle modulation. Another method is amplitude modulation. It is not as easily applied to solid-state transmitters as it is to vacuum tube transmitters because of the limited-amplitude dynamic range of power transistors for optimum operation. The dynamic range of varactor multipliers is also limited.

The two practical forms of angle modulation are phase modulation and frequency modulation. The degree of modulation achieved is primarily limited by the bandwidth of the varactor multipliers. If the operating frequency of the transmitter does not require the use of varactor multipliers, the bandwidth limitation becomes similar to that of vacuum tube transmitters. Varactor multiplier bandwidth will be discussed later. Figure 3 is a block diagram of a typical phase-modulated transmitter.

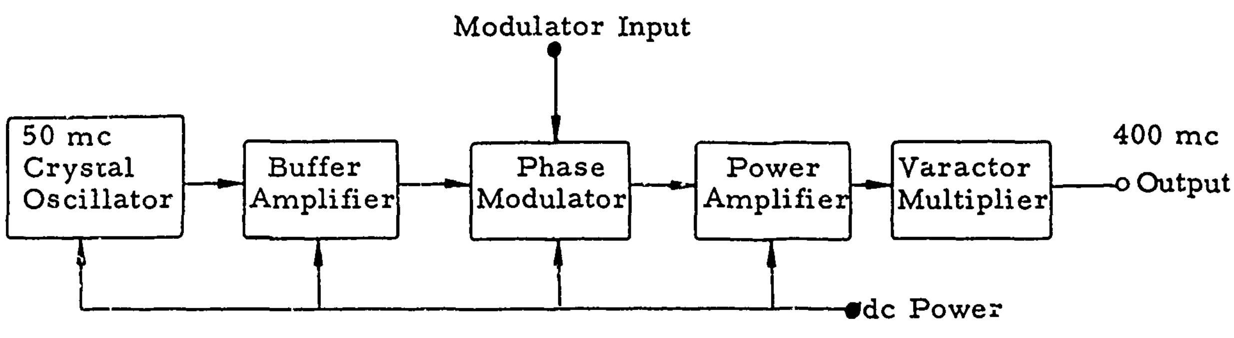


Figure 3. Typical Phase-Modulated Solid-State Transmitter, Block Diagram

Amplitude modulation is very difficult to apply to solid-state transmitters if 100-percent modulation is to be achieved with optimum performance. For single sideband operation, it may be possible to use a high-power, parametric upconverter. The efficiency of such a system is a function of the upconverter efficiency, which may range typically from 10 to 80 percent. A block diagram of a typical single sideband transmitter is shown in Figure 4.

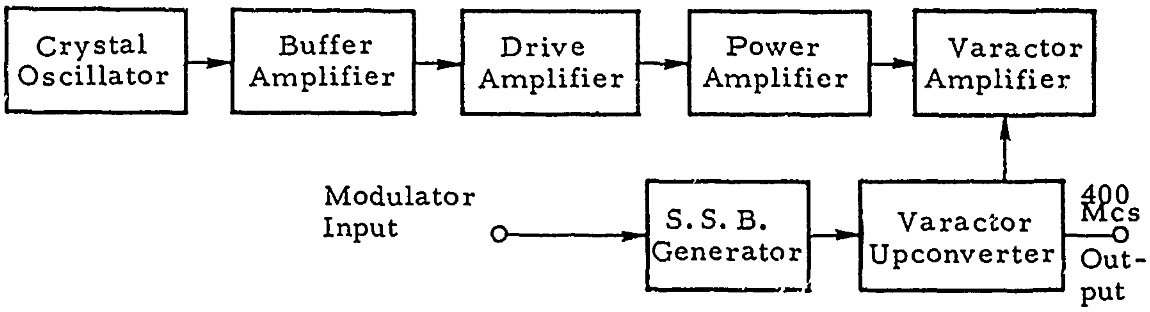


Figure 4. Typical Single Sideband Solid-State Transmitter, Block Diagram

d. Transistor Power Amplifiers

The efficiency of most transmitters is determined by the power amplifier stage. Transistor power amplifiers are presently operated below 200 mc if 2 watts or more of power output are required from the transmitter. A typical power transistor circuit configuration for operation at 100 mc would employ two or three transistors connected in parallel. At present, a power output of approximately 24 to 36 watts can be obtained with a collector efficiency of about 55 percent. The power gain is about 7 db.

If the operating frequency of the transmitter is to be below 200 mc, the transistor amplifier can serve as the output stage. (Of course, if the power output required is a few milliwatts, the transistor could be used in the output stage as high as 2 to 3 gc at present.) The transmitter design will then resemble its vacuum-tube counterpart, two basic differences being the lower power gain (less in transistors) and lower dc supply voltage required for the transistor amplifiers.

Vacuum tubes can withstand dc voltage and power dissipation transients that exceed their ratings without causing permanent damage to the device; transistors cannot. Power supply design requirements are more stringent for transistors in this respect. Also, transistors must be designed into a circuit that can present a low thermal resistance from the power transistors to the heat sink (normally the skin of the missile or spacecraft).

In general, more care is required in designing transistor power amplifiers than thermionic power amplifiers. However, once designed properly, the transistor amplifier offers a higher MTBF and increased reliability.

e. Varactor Frequency Multipliers

Because transistors are not presently capable of generating much power above 200 mc, and because crystal oscillators operate below 200 mc, efficient frequency multipliers are necessary when transmitters are to be operated in the microwave region.

The varactor frequency multiplier is presently the most efficient means of generating higher frequencies. It is also capable of handling many watts of power.

Present techniques in varactor multipliers utilize at least two varactors in a symmetrical circuit. When operated in this manner, the multipliers can handle as much as 40 watts at 100 mc. More output power could be obtained, but the conversion efficiency would decrease.

In doubler circuits, the conversion efficiency is typically 1 to 3 db, depending on the quality of the varactor and the operating frequency.

When operated as a tripler, the efficiency may vary from 1 to 5 db, again a function of the varactor quality and the operating frequency.

Frequency quadruplers can operate with from 2 to 6 db loss. The power-handling capability would decrease by a factor of almost 1/2 as compared to cascaded doublers.

For high-power operation, it is not advisable at present to use multipliers higher than the quadrupler. Cascaded low-order multipliers offer higher conversion efficiency as well as higher power handling capabilities.

Typical bandwidths of varactor doublers at present are about 10 percent or less. For frequency triplers, the bandwidth is about 7 percent or less.

Varactor multipliers are nonlinear devices in operation. Hence they are sensitive to the drive power applied. The dynamic range is quite limited and this limits the type of modulation to either low percentage amplitude modulation or to angle modulation.

A typical transmitter utilizing varactor multipliers is shown in block diagram form in Figure 5 below.

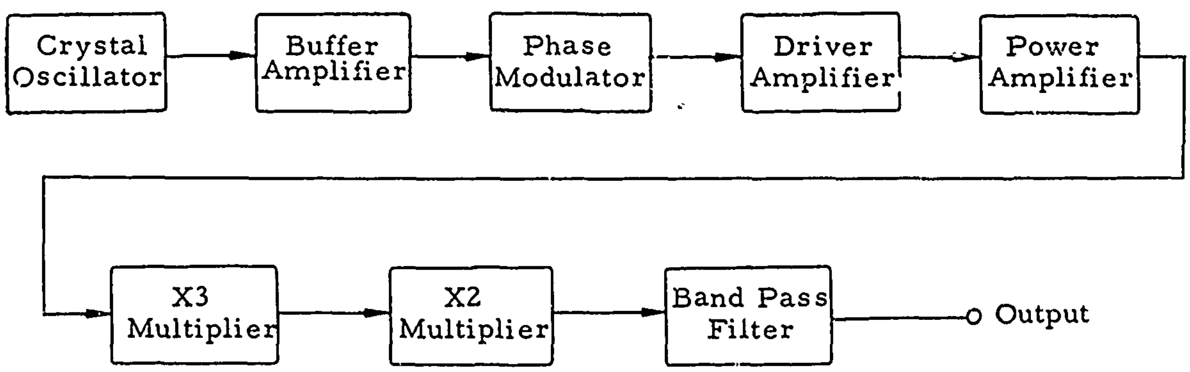


Figure 5. Typical Solid-State Transmitter Using Varactor Frequency Multipliers, Block Diagram

The use of frequency multipliers as output stages of the transmitter is perhaps the major difference in design philosophy between solid state transmitters and thermionic transmitters.

It should be emphasized that it is not by choice that multipliers are used in the output of solid-state transmitters. They are the only means by which high power can be generated at microwave frequencies. If power transistors capable of generating the necessary power at the operating frequency are available, this is obviously the more desirable approach from the standpoint of higher overall efficiency as well as offering less stringent requirements for the output filter and the output termination.

f. Summary of Transmitter Design

Solid-state transmitters for airborne and missileborne applications can be designed to operate from the low RF frequencies up into the high microwave frequencies. Transistors are used as the active elements in generating the RF power up to about 200 mc for high-power applications. Above 200 mc, the present approach is to use varactor frequency multipliers.

Angle modulation (either phase or frequency) is most suitable for solid-state transmitters. Single sideband modulation can be used by utilizing a parametric upconverter at the output of the transmitter. Amplitude modulation can be used if high distortion can be tolerated, and/or if the percentage of modulation is kept low.

The present state of the art in solid-state transmitter power output is summarized in Figure 6. As semiconductor devices improve, changes can be expected in design philosophy, with subsequent increases in overall transmitter efficiency and power output.

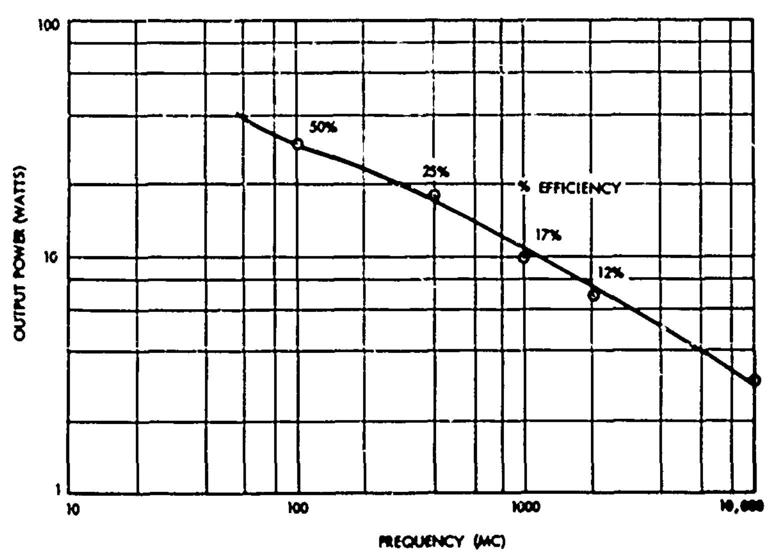


Figure 6. Solid-State Transmitters, State of the Art

4. FUTURE TRENDS IN SEMICONDUCTOR DEVICES

The frequency and power limitations of solid-state transmitters are directly related to semiconductor technology.

For transistors, the short-circuit common-emitter current, gain-bandwidth product (f_t) has a theoretical upper limit of about 10 gc. Since germanium transistors have already been built with f_t as high as 4 gc, and silicon transistors with f_t of 2 gc, we can expect improvement of f_t by only a factor of two or three. Of course, VHF power transistors have f_t of only 200 to 300 mc at present. It is predicted that the upper limit of power transistors could also be extended by a factor of three.

Varactor diode manufacturers feel that the fundamental upper limit on silicon varactor operation is about 600 gc. Improvement in packaging is anticipated to reduce the thermal resistance of present varactors only a little. However, because higher cutoff frequencies are anticipated, higher power output levels and higher conversion efficiencies can be expected.

5. FUTURE TRENDS IN TRANSMITTER DESIGN

With improvements in transistors and varactors, the design philosophy of solid-state transmitters may change by 1970. The major change expected would be in design of the last stage of the transmitter as a function of operating frequency. At present, it is not possible to use transistors as the last stage of a transmitter above 200 mc with more than 2 watts output. As has been stated previously, the use of transistors in the final stage offers several advantages:

- Output filtering and termination of the transmitter is less critical
- The transmitter DC to RF efficiency is increased
- High power varactor multipliers with their associated problems are not necessary.

Unfortunately, power transistors are not expected to operate with many watts of power at microwave frequencies in the foreseeable future. The design of transmitters at 1 gc and below may be able to utilize power transistors in the output stage of the transmitters; e. g., power outputs of 1 watt at 1 gc, or 10 watts at 400 mc can be expected.

High-power varactor multipliers will still be necessary for operation above 1 gc if one or more watts of power are required. But because of improvements in the varactors anticipated, the overall transmitter efficiency and power output may be increased.

Circuit techniques will be advanced to the point where varactor frequency multipliers may operate more efficiently for higher-order multipliers, eliminating the interface problems now encountered in cascading many multipliers. The bandwidth of frequency multipliers should also be improved so that more sophisticated systems could be designed utilizing all solid-state transmitters.

The parametric upconverter will become more popular for systems requiring amplitude modulation. Circuit design in this area will also be advanced to reduce the bandpass filter requirements. Solid-state transmitter output power versus operating frequency curves, as predicted for 1970, are shown in Figure 7.

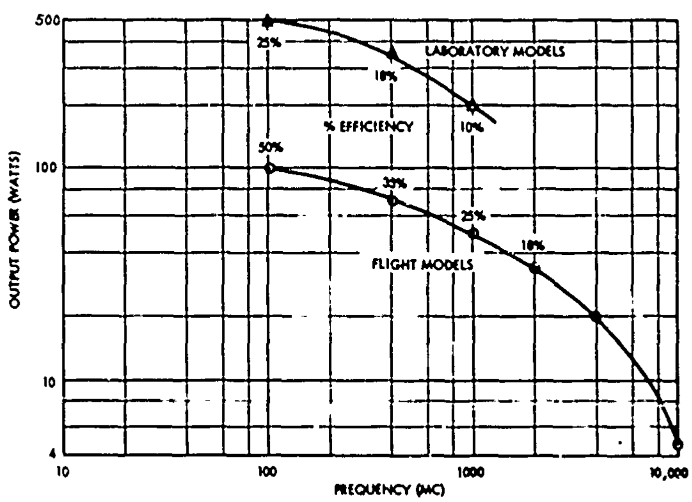


Figure 7. Solid-State Transmitter, State of the Art 1970

6. SUMMARY

High-power, solid-state transmitters are expected to operate with greater efficiency by 1970. In general, varactor frequency multipliers will still be used at the higher frequencies and for the highest output power levels. As an example, the 500 watts predicted at 100 mc for the laboratory model is based upon a 1-kw transistor amplifier at 25 mc, and the use of varactor frequency multipliers. In the flight models, it is assumed that a transistor amplifier is used at 100 mc, followed by varactor frequency multipliers up to X-band.

The transmitter design engineer may be able to use transistors at 400 mc if the desired output power is 10 watts or less. This appears to be a more desirable method of designing a transmitter whenever possible. Wider bandwidths and greater overall efficiency can be expected.

As a basis of comparison, the following chart shows present and predicted state of the art in transmitter power levels and efficiencies. The improvements are primarily based upon the power-transistor upper frequency-limit of operation.

	Early 1963		1970	
Frequency	Power Out	Efficiency (DC to RF)	Power Out	Efficiency (DC to RF)
100 mc	30 w	50%	500 w	25%
			100 w	50%
400 mc	15 w	25%	350 w	17.5%
			70 w	35%
			10 w	45%
1 gc	10 w	16.6%	200 w	10%
			50 w	25%
			1 w	40%
2 gc	7 w	11.7%	35 w	17.5%
			0.7 w	28%

Appendix XXII. ELECTRONICALLY SCANNED ANTENNAS

1. INTRODUCTION

Increasing demands upon radar performance have been occasioned by an increasing traffic of satellites, space probes, and missiles. Larger precision antennas, greater receiver sensitivity, more radiated power, and improved data handling are required to meet these demands.

One problem that the radar designer has been forced to consider is the need for large precise antennas with faster scanning capabilities. Considerable effort is being expended in the development of array antennas which incorporate a large number of elements, interconnected to radiate coherently. These coherent radiators form a beam which may be "steered" or positioned rapidly by electronic means over a wide angle without physical motion of the antenna. Typical applications in which various types of array are being sucessfuly employed indicate the versatility of this approach. Fixed-beam arrays with electronic lobing are currently being used for fire-control radar, multiple beam arrays for surveillance, and large electronic scanning radars are being employed in both ground and mobile multiple-target detection and tracking systems.

This versatility of the "inertialness-beam" antenna is probably the most significant advantage of the phased array; but it is obvious that this extreme flexibility can only be fully realized if the array is complemented with an equally well designed computing facility that is fully cognizant of the antenna's capability and limitations. In addition, an integrated transmitter/receiver system which can take full advantage of the array characteristics must be employed.

To enunciate the advantages and disadvantages of the array radar system, with emphasis on the steerable-beam mode of operation, the following review of major system design criteria is undertaken.

2. DESCRIPTION

Although the many techniques, combinations, and embellishments will not be discussed here in great detail, it will be useful to point out some of the basic building blocks for a phased array.

The block diagrams of Figure 1 show the fundamental versions of the electronically steerable array. The figure shows the simplest arrangements of the basic techniques to produce a single beam at the terminal point which may be scanned by varying the appropriate parameters. These systems show methods of beam scanning by frequency, phase, and time. They are here also presented in the order of fundamental increasing bandwidth capability that can be achieved. For a phased array radar, one or more of these methods can be combined in a wide variety of fashions to form a two-dimensional matrix yielding an array with scanning capability in both elevation, θ, and azimuth ϕ.

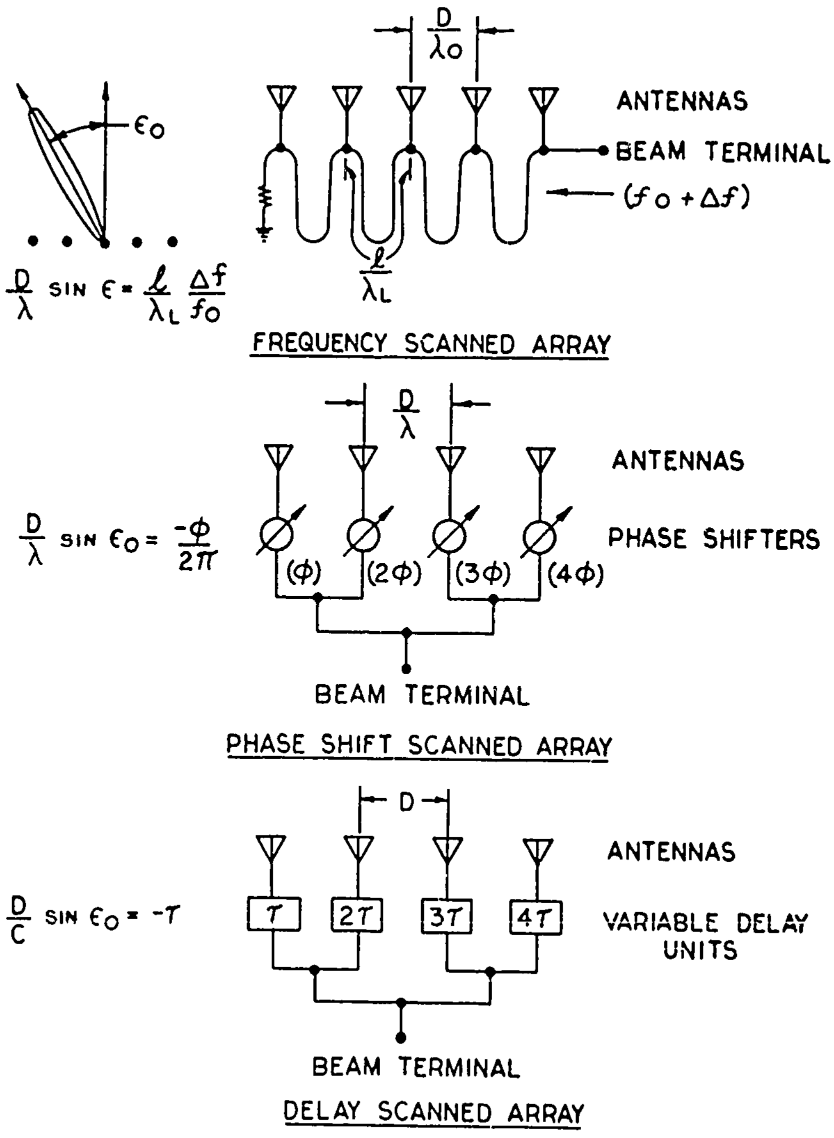


Figure 1. Fundamental Versions of Electronically Steerable Arrays

The complexity of any such radar is proportional to the number of antenna elements. To obtain an indication of the required complexity, the following relation may be used as a rough estimate giving the number of elements N in a planar array for half-power beamwidth θ_o and ϕ_o (in degrees) in azimuth and elevation, respectively.

$$N = \frac{10^4}{\theta_o \phi_o} \tag{1}$$

Thus, for a 1 deg x 1 deg pencil beam, the number of elements required would be on the order of 10,000. This may be reduced by a factor of 2 or 3 by appropriate distribution of the elements or excitation.

However, at the wider angles of scan, the number of elements required to realize a given system sensitivity increases rapidly because the beam width or gain of a single flat array is limited by the decrease in projected area at angles off boresight. There then comes a point (at approximately ± 60 deg or so for a 1-deg beam) where it is more practical to build additional arrays with faces pointed at different angles. The other obvious approach is to mechanically reposition the single array. Both techniques are currently being employed.

The amount of angular coverage for a single planar array is a function of two factors: element spacing and the radiation pattern of individual elements.

To illustrate the first, consider the simpler case of a linear array of elements regularly spaced by a distance d, in wavelength. A phase shift of d sin ε is required between adjacent elements in order to place the main beam at ε degrees from boresight. The beam may be theoretically directed anywhere ±90 deg from broadside since an endfire lobe may be obtained when

$$d\,(1 + \sin \varepsilon) = \lambda \tag{2}$$

This requires d = $\lambda/2$ or half-wavelength spacing of the element. Since as stated before, the beamwidth is a function of total aperture size, the designer will try to place the elements as far apart as possible consistent with the required coverage, side lobe suppression, operational bandwidth, and other pertinent parameters.

The other factor, element pattern requires that the individual elements have a radiation pattern that is as broad as the desired array coverage in a given plane. Much work in this area over the past years has enabled the designer to now obtain scan limits up to ± 60 deg from boresight without difficulty, and ± 90 deg is achievable.

In light of the above discussion, an array radar which would scan a complete hemisphere (comparable to the more conventional surveillance radars using a dish antenna) would require at least four planar antennas.

Although one might consider arrays on spheres or cylinders, intensive investigations on these configurations show that limitations on the percentage of surface that can be used to produce a beam at any given instant is such that the total number of elements required is rather independent of geometry for a given surveillance volume.

Thus in summarizing the configurations one should bear in mind that for typical two-dimensional arrays the antenna elements will probably number in the thousands and more than one face may be required to completely scan a large volume.

3. SYSTEM CONSIDERATIONS

a. General

The most common uses on any radar system are:

- The detection of an object's presence
- The resolution of multiple targets or objects one from another
- The measurements of orbit or trajectory parameters of the target such as position, range, velocity, etc.

Here we will examine some of the advantages and limitations of the array radar based on the above criteria. In particular, an indication of the ways in which the array may lead to enhanced performance in detection, resolution, and measurement accuracy will be made. Also ways in which the array radar may be inferior to the conventional radar will be pointed out. Considerations of reliability and cost are briefly discussed.

The following functional relation, one of several forms of the radar equation, will be useful in making the desired comparisons. Thus for a search radar and uncooperative targets

$$\frac{E_r}{N_o} = f\left(t_s, \frac{1}{\Omega}, P\frac{1}{R^4}, A_e, \sigma, \frac{1}{T_e}, \frac{1}{L}\right) \tag{3}$$

where

E_r = energy received from the target

N_o = noise power density

t_s = search time

Ω = solid angle to be searched every t_s seconds

P = average power

R = range

A_e = effective receiving area

σ = radar cross section of the target

T_e = effective system temperature referred to the antenna

L = combined loss factor for atmosphere, scanning, safety factors, etc.

A cursory examination of the above relation indicates simply that high power, large apertures, and low system noise are needed. However, a consideration of the radar's uses, environment, and an examination of each term of Equation (3) will indicate areas in which the array configuration can be an asset.

b. Detection

The ratio E/N_o bounds the signal-to-noise ratio obtainable with a given system. A typical design minimum is $E/N_o = 25$. Minimum signal-to-noise ratio is also a function of the desired probability of detection and probability of false alarm. Array radars may achieve a probability of detection of approximately 90 percent with a 10^{-10} probability of false alarm by exploiting the virtue of extreme beam mobility in a target "re-check" procedure. This will reduce the required E/N_o by a factor of two (3 db).

Investigation of programmed search rate Ω/t_s is suggested by the extreme scan versatility of the array radar's inertialess beam. To see how this operational capability will enhance probability of detection, consider the case of orbiting vehicles.

The angular rate of an orbiting vehicle decreases with increasing altitude. To illustrate, consider the simplest case of an earth satellite in a circular orbit. An idealized case including only the gross effects will be used for the purpose of illustration.

For the special case of the two-body problem the "vis-viva" integral gives orbital speed in geocentric coordinates thus:

$$\dot{s} = \frac{\mu}{r} = \frac{1}{(h + 1)}^{1/2} \tag{4}$$

where

$\dot{s}$ = orbital speed normalized to speed at unit radius ($\dot{s}_o = 25.9 \times 10^3$ ft/sec)

μ = sum of mass constants of the 2 bodies = 1

r = distance between mass centers in earth radii

h = altitude above earth surface in earth radii

For an observer at the earth's surface the maximum angular rate will occur for an overhead pass and is given by

$$\dot{\theta}_{max} = \frac{\dot{s}}{h} + \left(\frac{h + 1}{h}\right) \dot{\phi} \tag{5}$$

where

$\dot{\theta}_{max}$ = maximum angular rate for observer on equator counter-rotating to satellite

$\dot{\phi}$ = earth's angular rate at equator

The second term in the above equation accounts for the counter-rotation of an observer located on at the earth's equator (worst case). Combining Equations (4) with (5) and introducing constants (earth's angular rate, unit conversion etc.) $\dot{\theta}_{max}$ in deg/sec is obtained thus

$$\dot{\theta}_{max} = \frac{a\,0.071}{h\,(h+1)}\;1/2 + 0.0417\;\left(\frac{h+1}{h}\right) \qquad (6)$$

Keeping in mind that range (R) increases with altitude (h), it can be seen from Equation (6), that to detect low altitude vehicles with high probability requires a very fast scanning radar, i.e., Ω/t_s must be large.

On the other hand, a fast scanner will be range limited by power because, to maintain a given probability of detection and given small t_s, power must increase as R^4. The desirability of a suitable scan-time versus range program becomes apparent if the radar is required to survey orbiting targets at all altitudes.

Consequently zoned-search techniques, which may be easily employed by array radars, can maintain a high probability of detection over a large surveillance volume for a given power, thus providing more detection-per-watt due to a better scan versatility.

Power handling capacity is generally limited by either voltage breakdown for pulse radars or temperature rise allowable for CW systems. The demands for more power to extend the detection or tracking range of radar systems increase with the natural trend in the development of far ranging space vehicles and weapons systems. By using the integrated or active arrays, which can employ a large number of relatively small transmitters, voltage breakdown can be avoided. This offers the possibility of extending the power output of radars to the point where the primary power available becomes the limiting factor. Arrays using integrated transmitter/antenna elements are capable of extending the average power output of large arrays into the tens of megawatts. It is considered very reasonable, for example, to transmit 20-kw peak power or 400 watts average per element or group of elements with a small transmitter module. Temperature rise per element is also reduced by these techniques. Although cooling of many transmitters can become a problem, small efficient cooling systems are rapidly being developed for many applications and it is not anticipated that cooling will be a limiting factor here.

The phased array may capitalize on another factor of the radar equation (Equation 3) to enhance detection, namely aperture area (A_e). Here the array does not require mechanical motion, and has the advantage of a simpler geometry (e.g., a planar configuration).

Fixed antennas are easier to fabricate and to maintain with given precision than are movable structures. Large changes in bending moments do not occur as with a corresponding steerable aperture. Therefore, precision and shape are not a function of beam pointing direction. Also, the absence of moving equipment with associated changes due to wear, loads, concentricity, etc., may extend the validity of the original calibration over a longer time period. The simpler geometries have the feature that they can be determined to conform to a given accuracy and shape without the use of extremely complex or specialized instruments. Further, the array being the primary radiator or aperture is less sensitive to surface tolerances by a factor of two or more over antennas, employing one or more reflecting surfaces. That is, tolerances in the array surface contribute only once to distortion of the wave front as compared to the double contribution of each reflecting element of a quasi-optical antenna. Fabrication costs are thus reduced, and operation at shorter wavelengths is possible for a given surface tolerance. These advantages are particularly true when considering large apertures, and the array can indeed be made very large.

c. Resolution

The ability of the radar to discriminate between targets depends primarily upon the following target characteristics:

- The difference in radar cross section (σ) of the targets
- The separation between targets in range and angle
- The signal-to-noise ratio returned from each target (a function of scintillation, polarization, etc.).

The question of resolution accuracy then becomes not only a matter of single target measurement but of the amount and kind of targets the radar must resolve. When employed against, or in cooperation with, a wide variety of targets, the versatility of the array radar may show the array to have more optimum overall resolution characteristics. It is not clear that the phased-array radar can compete in angular resolution with conventional systems of comparable aperture. It must be remembered that the electronic-scan array is required to detect targets off the optimum boresight axis with consequently wider beams. Therefore, the angular resolution cannot be expected to be superior to the more conventional radar with a comparable aperture. However, the ability of a radar to resolve in range is limited only by peak power obtainable and effective bandwidth. As has already been shown the array can offer advantages in attaining large peak powers, but it is necessary to resort to delay-scanning techniques to obtain the large bandwidths required. Hughes Aircraft currently has under development a high-power, delay-scanning pulsed radar capable of 40 to 50 percent bandwidths. These bandwidths permit the use of pulse lengths in the millimicrosecond region with consequent range resolution in inches. Finally, there is no reason why the doppler resolution capabilities of arrays should be superior to conventional radars.

d. Measurement Accuracy

Improved resolution capability also implies increased measurement accuracy and thus the preceeding remarks under "resolution" also apply here. Only when the radar is faced with a large number of highly maneuverable targets, such as in testing penetration of weapons systems, and must make measurements on many targets in a short time, is the array flexibility a definite advantage. When measurements must be made from a mobile mount (e. g., a ship) the inertialess beam is also a distinct advantage.

e. Reliability Considerations

In a large array the active and passive components might number in the range of tens of thousands to millions. Obviously, the complexity of such a system points to the maximum employment of solid-state devices or passable elements to achieve high reliabilities. This approach is certainly now feasible for the antennas, receiver and computer modules. However, the higher transmitter powers desired are currently obtained only by the use of vacuum tubes. Although it might be expected that the multitude of components, which include vacuum tube modules, would result in a less reliable system than simpler conventional radars, the situation may in fact prove to be more satisfactory when considering "on-the-air" time or overall performance.

To evaluate the situation consider the typical failures of modulators, transmitters, receivers, beam forming, and steering circuits, which in a conventional radar can cause catastrophic failure. In the phased array, the same effects can cause only localized degradation. For example, in a typical array, a 10-percent random failure of transmitters would result in a reduction in range of only 5 percent and negligible side lobe degradation. If some components are common to all receiving and transmitting channels, redundancy techniques comparable to the conventional radar can be employed. Thus, the attainable reliability, although it may be more complex to evaluate, can. by careful consideration in a well-designed array, virtually eliminate the possibility of catastrophic failure.

A recognized disadvantage of the array is that the great number of components present a corresponding problem detection and maintenance. This problem is eased by the following:

- Amount of maintenance can generally be predicted in advance.
- Maintenance of simpler modules can be reduced to an efficient routine.
- Effective techniques have already been proven in large computing facilities which have a similar problem.

Finally, histories on large array radar systems employing several thousand electron tube elements indicate operational reliabilities exceeding 0.99.

f. Cost Factors

It is apparent that the array radar can become quite complex and therefore costly. Unfortunately, little is known of the ultimate cost of large array radars, due to the fact that they are becoming operational equipment.

One comparison that could be a useful factor when considering a multiple target capability is the system cost per target. Studies have indicated that single arrays can be made to handle thousands of targets. Figure 2 shows a qualitative relation between system cost and number of targets for conventional radar and a phased-array system.

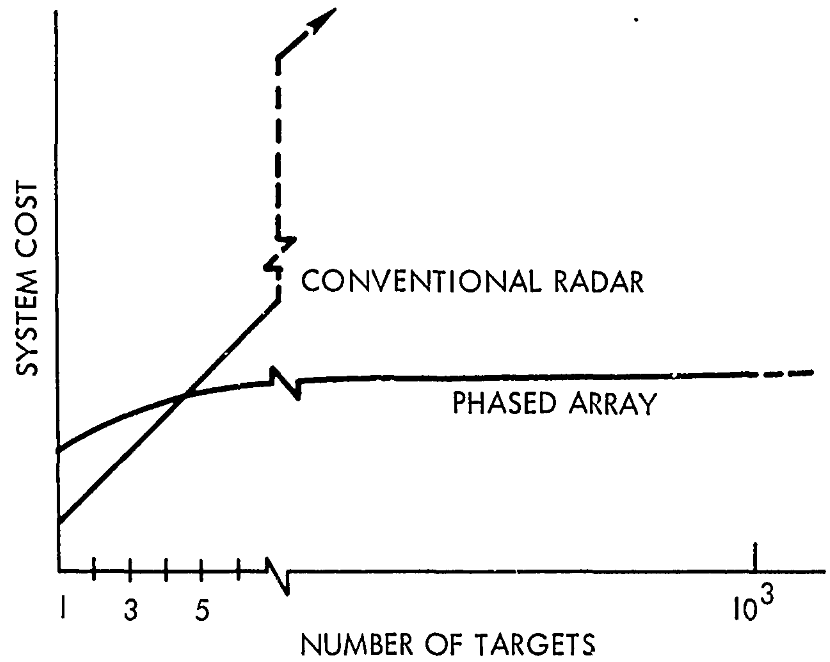


Figure 2. Radar System Cost Versus Number of Targets

The significant point here is that, when handling large numbers of targets, the use of a single phased-array system shows a distinct advantage in cost over a system employing the required multiplicity of mechanically steered antennas. Present information indicates the crossover to be in the neighborhood of four to five targets.

Since industry is lacking extensive construction and mass-production experience necessary for generalized costing predictions on arrays, quantitative examples must be based on a few specific systems. For a large (several thousand element) array, one might expect the cost to run on the order of 5,000 dollars per element, or a system cost of 20 to 50 million dollars. Mass-production techniques might be employed to reduce costs of potential modular highly redundant systems such as the array.

One final consideration worth noting is in conjunction with the modular nature of the phased array system. A well-planned array radar system is amenable to gradual growth in capability by the simple expediency of adding modules. This is a distinct advantage of the phased array in applications where demands will be ever increasing.

4. CONCLUSIONS

Table I shows the main advantages and disadvantages of arrays in comparison with two types of conventional or "dish" radar systems. For precise angle tracking or resolution, the steerable dish currently appears to have a greater potential.

A numerical comparison of dish-array tradeoffs from Reference 1 will be even more enlightening. Assume the problem of tracking a low cross section ballistic vehicle in the shape of a narrow cone with a half angle of 15 deg. The on-axis tip contribution for $f = 5$ gc and $\alpha = 15$ deg is $\sigma \simeq 4 \times 10^{-7} m^2$. An FPS-16 could barely see this target at all, so assume, further, a hypothetical loss-free, 85-foot dish radar with pulse duration of 50 μsec, peak power of 5 megw, frequency of 5 gc, receiver gain of 59.8 db, S/N of 2 db for a single pulse, etc., which would result in 64-db better performance than that of the FPS-16. This hypothetical dish radar could track the above target tip at 184 nmi. Compare this with a hypothetical array radar operating at the same frequency. Assuming 10^5 elements in the receiving array with a monopole gain of 5 db, the overall maximum receiving gain would be only 55 db. If 10^4 log-periodic transmitting elements were to be employed with individual gains of 10 db each, the total maximum transmitting gain would be 50 db. The 85-ft dish gain, at the same frequency, would have close to 60 db gain so, at this point, the array is 15 db down from the dish for a two-way path. Assume, further, 10^4 transmitters of 50 kw peak power output, each, for the array, making a total array peak power of 500 megaw. The net advantage of the array over the dish is 5 db, enabling it to track the same vehicle at 206 nmi. If an MTBF of 2000 hrs can be realized for each 50-kw transmitting tube, the 10^4 element array would require 4×10^4 tuber per year, or about 100 tubes per day. At 500 dollars per tube, the total annual transmitting tube cost would be 20 million dollars. The system would, of course, require plug-in modules insensitive to element position of power supply variations. Further calculations of array-dish radar trade-offs indicate that arrays are attractive only in the region of roughly 200 mc to 2 gc for multiple beam or two-dimensional scanning applications.

Electronically steerable phased-array or similar types of radars might be advantageous at low frequencies for coping with:

- Detection and tracking of a variety of multiple targets
- Demands for multiple mode of operation or function such as simultaneous search and track
- The additional requirements for high power, wide bandwidth and large apertures for lunar or deep-space data acquisition
- Specialized requirements involving mobile platforms or blast hardness.

From an engineering and scientific standpoint, the feasibility of the phased array has been established beyond any reasonable doubt; however, their potential for global range instrumentation is somewhat more questionable unless unforeseen developments should enable either technological simplification or greatly reduced cost and increased reliability per element. The possibility of this can be deduced from the parallel case of large-scale digital computer development, and their widespread availability and use today.

REFERENCES

1. R. C. Hansen, "Whither Arrays," The Microwave Journal, April 1963 April 1963, pp 24-28.

Table I. Advantages and Disadvantages of Various Radars

Combination Detection and Tracking Radar (Phased Array)		Tracking Radars With Acquisition Radar		Tracking Antenna Alone	
Advantage	Disadvantage	Advantage	Disadvantage	Advantage	Disadvantage
1. Multitarget tracking capability 2. Electronic beam scan (scan from farthest point within 10 μ sec) 3. High peak power 4. Controlled side lobes. 5. Good hardening 6. Good jamming resistance. 7. Good probability of detection with low false alarms 8. Versatility 9. High data rate for good accuracy. 10. Can stand 20% element failure without severely reducing accuracy. 11. Good planned growth capability.	1. Very high initial cost 2. Beamwidth and gain are functions of scan angle 3. Maintaining large number of components 4. Heavy 5. Uses high prime power	1. Same as under tracking radar except cost would be higher. 2. Detect unknown targets	1. Tracking radar needed for each target 2. Long range tracking requires large antenna w/ difficult mech. tolerance 3. High inertia 4. Can be confused by multiple targets or false targets 5. Detection of unknown or "surprise" targets would be difficult	1. Low cost 2. Proven design techniques 3. Can operate over wide frequence range. 4. Capable of hemispherical coverage 5. High accuracy for slower tracking rates.	1. Time lost in acquiring target 2. Can track only one target (some track more if target is within beam) 3. Long-range tracking requires large antenna w/difficult mech. tolerance 4. High inertia 5. Requires time to reacquire lost target 6. Can be confused by multiple targets or false targets 7. Unable to stand sustained high power 8. Detection of unknown or "surprise" targets would be difficult

Appendix XXIII. THE EFFECTS OF FLAME PLASMA ON RF COMMUNICATIONS

Flame ionization effects on electromagnetic propagation have been apparent on virtually all ballistic missile flight tests. In many cases, especially with smaller liquid-propellant vehicles, these effects are not severe enough to impair progress of a flight test program. In other cases, particularly with the larger solid-propellant vehicles, flame ionization effects have been very serious and have required drastic adjustments to data acquisition procedures. Future generations of vehicles of both types can be expected to pose even greater problems. Measures presently used to reduce or avoid unwanted flame ionization effects may prove to be either inadequate or prohibitively costly for these future cases.

Much theoretical and experimental work has been done to arrive at a better understanding of the flame plasma and its effects on electromagnetic propagation. Much flight-test data analysis has been directed toward a definition of the precise nature of these effects on actual range instrumentation. This work has been directed primarily toward the immediate needs of the missile programs and, as such, has been reasonably effective. It has been possible to recommend airborne antenna designs, ground instrumentation site locations, operating procedures, and other means for limiting desirable effects to the extent that useful communication with the missile can be maintained.

There is no means available as yet for accurate prediction of effects that are expected from larger vehicles of the future. The extremely high cost of both vehicles and instrumentation systems demands a detailed knowledge of the precise effects of the flame plasma on precision tracking and telemetry instrumentation to assure, even in program planning phases, that reliable communication can be maintained. To accomplish this, it will be necessary to review comprehensively the present state of knowledge, identify those areas where present knowledge is insufficient, devise means for filling these gaps in knowledge, and develop techniques for effective use of this knowledge in missile system instrumentation design.

It is not intended here to summarize the flame plasma problem comprehensively but rather to provide adequate background of the situation.

Current flame plasma knowledge can be summarized under three headings:

Basic flame ionization processes

Theoretical analysis of propagation

Radio interference observations on missile flight tests.

These are discussed briefly in the following paragraphs.

Sea level experiments with large rocket engine exhausts have confirmed that varying degrees of ionization occur along the length of the exhaust. Three regions are of interest:

The exhaust gases from the combustion chamber through the expansion section, past the exit plane until the gases become shock heated

The shock heated region which, if it acts as a flame holder, may also be accompanied by afterburning

The afterburning region (This is a superficial layer of fuel-rich gases that mix with the ambient atmosphere and ignite.)

Ionization near the exit plane is independent of environmental conditions. It is a function of mixture ratio or propellant composition and of flame temperature. The main efforts in this region have been toward determining its electromagnetic properties (electron density and collision frequency) and the sources of ionization. The latter are not well known, but experiments show that ionization levels are too high to be attributed to processes in thermodynamic equilibrium. Conjectures as to source include chemi-ionization; decay of molecules in long-lived metastable energy levels with subsequent electron emission; thermionic emission from high-molecular-weight graphitic-like particles with low work function; and thermal ionization of alkali metal impurities with slow recombination rates. The last mechanism has been experimentally investigated on an Atlas engine with the finding that traces of sodium contaminant in the fuel cause significant increase in ionization. There has not yet been sufficient experimentation to identify the magnitudes of the other ionization sources, hence, the effects of contaminants cannot be pinpointed. Similar experiments with Minuteman engines also have not been conclusive, but the relatively high degree of ionization may be attributed either to the presence of aluminum in the propellant composition, or to the presence of potassium as an impurity. Two Air Force programs are being conducted to clarify this problem: 1) a theoretical and experimental study by Stanford Research Institute to identify the role of Al_2O_3 particles, and 2) a measurement program by the University of Utah to identify the role of potassium.

Shock-induced ionization also is being investigated by the University of Utah. The conical shock wave issuing from the lip of the expansion section complicates microwave measurements made during sea-level testing. Ionization of this region must be understood for proper interpretation of experimental results and extrapolation of these results to vacuum conditions.

Afterburning is believed to be a significant source of ionization. Preliminary research in this area has led to the current 2-year Air Force research program at Stanford Research Institute.

There also has been much work done with laboratory flames. These are more amenable to detailed study (hence, they are extremely valuable). Experiments with a variety of fuels have shown the same high electron densitites as rocket exhausts. Continued laboratory work is essential, but the basic problem of adequately simulating rocket exhaust conditions in the laboratory still remains.

Thus, knowledge of basic ionization processes is incomplete. Electromagnetic parameters of the flame plasma have been measured, but there is still insufficient information to fully understand the sources of ionization and to interpret the electromagnetic measurements.

Many theoretical studies have been published on the propagation of electromagnetic waves through flame plasmas. For the most part, these studies have consisted in the application of well-known principles of electromagnetic theory to assumed models of the flame plasma. Their limitations are two-fold: 1) inadequacy of the model as the result of lack of understanding of the basic flame ionization processes, and 2) lack of adequate data collection techniques on the actual missile flights to provide supporting data.

Theoretical analyses have included volume attenuation within the plasma, reflection at plasma boundaries, and refraction through varying plasma density, diffraction, and ray tracing.

Obviously, one of the best sources of flame effect data is the test range instrumentation. In fact, flame effects have been observed at the Atlantic Missile Range on virtually all ballistic missile tests and on all radiating instrumentation systems. Unfortunately, range instrumentation planners and operators were slow to accept the seriousness of flame effects, with the result that practically all recorded data has been "by-product" data, that is, data gathered for purposes other than the flame effect problem. Such byproduct data is usually recorded on low band-width recorders with inadequate attention to calibration and therefore is suitable only for identifying gross flame effects.

Liquid-propellant missiles in the past all have exhibited flame effects. These effects have been relatively small, so that most problems were solved by such simple measures as shifts to higher frequencies or use of larger antennas. With solid-propellant missiles, flame effects are much more severe, and it has become evident that these simple relief measures are inadequate both for present solid-propellant vehicles and for the larger liquid- and solid-propellant vehicles of the future.

In spite of flight test data deficiencies, some basic important characteristics of flame effects have been identified, such as dependence on aspect angle, dependence on frequency, and the effects of "flame noise" on precision tracking data. However, there are still numerous information gaps which can be resolved only by the use of improved data-gathering techniques. Such techniques would include wider selection of receiving site location, use of more accurate antenna patterns, monitoring of transmitted power, and recording of true signal strength rather than AGC voltages.

BIBLIOGRAPHY

S. Altshuler and R. V. Fredericks, "Radar Backscattering from an Inhomogenous Plasma," 9340. 2-49, September 1961.

G. Brown, "Microwave Attenuation During the Booster Phase of Atlas Flights," STL Report GM-TR-0156-00454, MT 58-17393, 22 August 1958. (S)

E. Burns, "Analysis of Minuteman Exhaust Products," prepared for Aerojet-General Corporation, 14 monthly reports, 15 August 1961 through 15 September 1962, plus Final Report 25 October 1962.

E. Burns, "Study of Microwave and Visible Lightwave Transmission Through Solid Propellant Exhaust Gases," prepared for NASA, 2 monthly reports and 1 quarterly report, 1 August through 1 November 1962.

E. Burns, "Spectrophotometric Study of Thallic Nitrate Complexion," J. Am. Chemical Society, Vol. 79, No. 866, 1957 .

E. Burns, "Analysis of Small Amounts of Metallic Aluminum Encompassed by Refractory Aluminum Oxide, " Anal. Chem., In Press.

E. Burns, "Evaluation of Some Analytical Procedures for Solid-Rocket Oxidants," Part 1, JPL Memo 20-158, 20 June 1958.

F. S. Coxe, "Operational Characteristics of the TLM-18 Automatic Tracking Antenna," IRE-PGSET, March - June 1959.

F. S. Coxe, "A Review of PCM at Cape Canaveral," STL Report, GM PTM-2936, MT 59-11741, August 1959. (C)

F. S. Coxe, "Capabilities and Limitations of the TLM-18 60-Foot Parabolic Telemetry Antenna," STL Report, December 1959.

J. R. Duffett, E. M. Simpson, and A. G. Smith, "Ionospheric Effects on UDOP Data," 12 May 1961.

J. R. Duffett, T. R. Morel, and A. W. Wren, "Summary Report, AZUSA Mark 1 Performance," 17 July 1959. (S)

J. R. Duffett, "Variate-Difference Noise in AZUSA Mark II Direction Cosines," FTM-403, 14 August 1961.

J. Gardner, "Preliminary Report on Diffraction of L-Band Radiation Around a Rocket Exhaust Above the Atmosphere," STL Memo, Report No. GM-TR-124, GM 42-139, MT 59-19290, 18 January 1957. (C)

BIBLIOGRAPHY (Continued)

A. A. Geiger, "Derivation of Formulas for the Prediction of Signal Scattering and Fluctuation Caused by Flame Plasma Turbulence," 9334. 13-5, 15 March 1963.

A. A. Geiger, "Minuteman Flame Noise History," 9334. 13-11, 11 July 1963.

A. A. Geiger, "Minuteman PCM Noise Spectral Density Caused by Flame Effects," 9334. 13-6, 18 March 1963.

A. A. Geiger, "Prediction of Noise Spectral Density of Signals Affected by Missile Exhaust Flame Plasma," 9334. 13-7, 20 March 1963.

A. A. Geiger, "Initial Study for the Performance of Statistical Analysis of Flame Noise and Attenuation," 9334. 7-46, 12 September 1962.

A. A. Geiger, "Calculation of Flame Plasma Transfer Function," 9334. 13-1, 23 November 1962.

G. Jeromson, "Atte[illegible]on of Radio Guidance Signals for the ICBM," GM-TR-19, 24 April 1956.

K. Kelso, "Doppler Shifts and Faraday Rotation of a Radio Signal in a Time-Varying Inhomogenous Ionosphere," TM-60-0000-09012, 23 February 1960.

J. Kliegel, and G. Nickerson, "Flow of Gas-Particle Mixtures Through Axially Symmetrical Nozzles," presented at the American Rocket Society Propellant, Combustion and Liquid Rockets Conference, Palm Beach, Florida, A. R. S. paper 1713-61, April 1961. (Also available as STL Report TM-7106-0023-MU000).

J. Kliegel, and G. Nickerson, "The Calculation of Supersonic Gas-Particle Flows in Axis-Symmetric Nozzles by the Method-of-Characteristics," STL Report 6120-8345-MU000, 1 May 1962.

H. P. Liepman, "Survey of Base Flow Recirculation and Base Heating Problems," presented at Symposium on Simulated Altitude Testing of Rockets and Missile Components at AEDC, Tullahoma, Tennessee, AEDC-TR-60-6, 1-2 December 1959.

H. P. Liepman, "An Analytic Design Method for a Two-Dimensional Asymmetric Curved Nozzle," University of Michigan, Project M-951 Report, June 1953.

H. P. Leipman, "Report on Supersonic Wind Tunnel Tests for Wright Air Development Center," University of Michigan, Report WTM 240, March 1953. (C)

BIBLIOGRAPHY (Continued)

H. P. Liepman, "Calibration Report on the University of Michigan 8 by 13 Inch Supersonic Wind Tunnel, Part V, Aerodynamic Calibration at Nominal Mach Number of 2. 43," University of Michigan, Report WTM 248, January 1954.

H. P. Liepman, "Development of a Variable Mach Number Sliding Block Nozzle and Evaluation in the Mach Number Range 1. 3 to 4. 0," University of Michigan, WADC TR 55-88, March 1955.

H. P. Leipman, "An Experimental Investigation of the Forces and Flow Fields Produced by a Jet Exhausting Laterally from a Cone Cylinder in a Mach 2. 84 Stream," University of Michigan Report WTM-255, November 1955.

F. Mastrup, "Ionization and Electron Capture," DA-36-039-SC-85269, GM-60-600-33667.

F. Mastrup, "Ionization Phenomena," DA-36-039-SC-88878, Contract Final Report, 8714-6008-SZ-000, September 1962.

W. C. Meecham, "Formulation of Electromagnetic Scattering and Reflection Problems in Plasmas with Magnetic Field," TR-60-0000-00436, December 1960.

M. Moe, "Method for Calculating Supersonic Flow in a Jet Containing Shocks," STL Report, 7 March 1957.

P. Molmud, "Flame Attenuation on the Thor Able-3 and -4 and Atlas Able-4," STL Report, MT 59-73807, 15 June 1959. (C)

P. Molmud, Flame Ion Conference at Princeton, 12-13 October 1959, STL Report, 1 December 1959.

P. Molmud, "A Theoretical Study on the Interference by the Reverse Thrust Plume on Communication (AZUSA) with the Minuteman Instrumentation Package," STL Report, MT 60-10968, 28 December 1959. (C)

P. Molmud, and J. H. Gardner, Proceedings of Conference on Ions in Flames, Stanford Research Institute, GM-60-0000-13732, MT 60-16759, 19-20 May 1960. (S)

P. Molmud, "Dependence of Polaris Exhaust Attenuation Frequency," STL Report, MT 61-91060, 1 August 1961. (C)

P. Molmud, S. Altschuler, and M. Moe, 'The Electromagnetics of the Rocket Exhaust," STL Report, GM-TR-0165-00397, 15 June 1958.

BIBLIOGRAPHY (Continued)

P. Molmud, "The Nonlinear Propagation of Strong Electromagnetic Waves Through Weakly-Ionized Media," TM-59-0000-00397, 14 October 1959.

P. Molmud, "Summary of Flame Effects," Working Conference held at JPL, 10-20 September 1957, STL Memo MT 58-16489, 24 October 1957. (S)

P. Molmud, "The Propellant Contaminant Program," STL Memo, MT 58 -321, 8 July 1958. (C)

P. Molmud and S. Altschuler, "Paraxial Transmission Through Expanding Rocket Exhausts," STL Memo, Report No. AP-51, 5 March 1956. (C)

P. Molmud, "The Altitude Dependence of Radar Signal Loss," STL Report, MT 58-16327, 3 June 1958. (C)

P. Molmud, "Flame Attenuation of Telemetry During Silo Launch of Minuteman," STL Memo, MT 60-13341, 22 January 1960. (C)

P. Molmud, "Comparison of Polaris 2nd Stages — Old Versus New Propellant," STL Memo, MT 60-19486, 7 December 1960. (C)

P. Molmud, "Expansion of a Rarefied Gas Cloud into a Vacuum," The Physics of Fluids, May-June 1960.

H. A. Poehler, "Plasma-Enhanced Breakdown of Telemetry Antennas During Ballistic Missile Powered Flight," IRE Transactions, Space Electronics and Telemetry, Vol. SET-7, December 1961.

H. A. Poehler, "The High Energy Solid Propellant Flame Attenuation Problem," The Classified Proceedings of the Second Symposium on the Plasma Sheath, AFCRL, 10-12 April 1962, Boston, Massachusetts (in publication). (C)

H. A. Poehler, "Staging, Its Effects on the Propagation of AZUSA Signals," American Rocket Society, Technical Paper 1660-61, 13 March 1961.

H. A. Poehler, "Exhaust Flame and Telemetry Breakdown Problems," IRE Transactions, Fifth Symposium on Space Electronics and Telemetry, Section 8-3, 19-21 September 1960.

H. A. Poehler and P. Molmud, "Prediction of Interference of Communication with Titan by Solid Propellant Vernier," STL Report, MT 61-91056, 14 July 1962. (C)

H. A. Poehler, "Breakdown of Rarefied Air Surrounding Nose Cone and Air Frame Telemetry Antennas on Series B," WS-107A-1 Flights, STL Report, MT 59-11700, 26 January 1959. (C)

BIBLIOGRAPHY (Continued)

H. A. Poehler, "Factors Affecting Telemetry Signals," presented to Flight Test Conference, American Rocket Society, Daytona Beach, Florida, 23-25 March 1959.

H. A. Poehler, "Plasmas, Their Importance in the Flight Testing of Ballistic Missiles," IEEE Conference Paper No. CP 63-111, Winter General Meeting, New York City, 27 January - 1 February 1963.

H. A. Poehler, "Analysis of Telemetry Dropouts in the Titan, Thor and Atlas," Proceedings of Conference on Ions in Flames, Stanford Research Institute, 19-20 May 1960. (S)

H. A. Poehler, "Telemetry Signal Strength Attenuation," Proceedings of the Study Group on Missile Launch Phenomenology," Stanford Research Institute, 18-19 October 1961. (S)

H. A. Poehler, "Observation of Radio Interference by Exhausts of Operating Missiles," STL Report, GM-PTM-5656, MT 62-82407, 7 September 1962. (C)

H. A. Poehler, "Flame Attenuation Problems on 401," STL Report, MT 61-75502, 16 May 1961. (S)

H. A. Poehler, "The A2X Attenuation Problem," STL Report, GM PTM-3262, MT 60-10836, 27 December 1960. (S)

F. L. Scarf, "Wave Propagation in a Moving Plasma," TR-60-0000-09245, August 1960.

A. G. Smith, "Derivation of Minimum Signal Strength Requirements for the Radar-Beacon System Used on AGS Launches," 30 August 1960.

A. G. Smith, "Minuteman Flame Effect at Downrange UDOP Sites – Addendum to Minuteman Flame Effect at Downrange UDOP Sites," 4 August and 3 October 1961.

A. G. Smith, "System Sensitivity Calculations – Planned UDOP System for Minuteman," 15 May 1961.

A. G. Smith, "Transfer of Electromagnetic Energy Between the FPS-16 Radar and an Elliptically Polarized Beacon Antenna," 17 January 1961.

C. Sonett, "Graphs of Electrical Properties of Ionized Media," STL Report, GM-TM-190, 30 August 1957.

J. R. Whinnery, "Ray Studies for Propagation Through the Flames," GMCC GM271.2064, Memorandum, MT 59-76794, 6 September 1955. (S)

Appendix XXIV. NUCLEAR RADIATION HAZARD TO ONBOARD TELEMETRY, TRACKING, AND COMMAND EQUIPMENT

1. INTRODUCTION

The effects of nuclear radiation on aerospace vehicle electronic equipment will present a critical problem in the very near future for spacecraft which carry nuclear reactors or which must operate for long periods in the radiation belts which surround the earth. To the extent to which successful execution of a mission depends on reliable operation of the vehicle-borne tracking, telemetry, and command equipment, this radiation problem has implications for the ground test environment.

The purpose of this appendix is to define the nuclear radiation problem for aerospace vehicle electronic equipment and to indicate the manner in which efforts devoted to its solution should proceed. While it is recognized that work toward this end is currently under way, this discussion will emphasize the need for additional work in this area.

2. THE NUCLEAR RADIATION ENVIRONMENT

a. Nuclear Reactors

Of importance today and with significance for future space exploration is the utilization of nuclear reactor propulsion systems and other similar exotic devices. Fortunately, a reactor is a controlled unit, and man to some extent can cope with the radiation hazards involved. However, total shielding from a nuclear reactor is impractical, especially with regard to nuclear powered vehicles, and this factor must be considered in the design of equipment so affected.

The appropriate spectrum of neutron energy within a carbon reactor is shown in Figure 1. (The peak would be slightly to the right for a water or beryllium moderated reactor.) The abscissa is the neutron energy E in million electronvolts (mev) and the ordinate n(E) is the number of neutrons at that energy. The curve indicates that the most probable neutron will have an energy of about 1 mev. The median neutron is at the 1.6 mev energy level, and the average neutron energy is 2 mev.

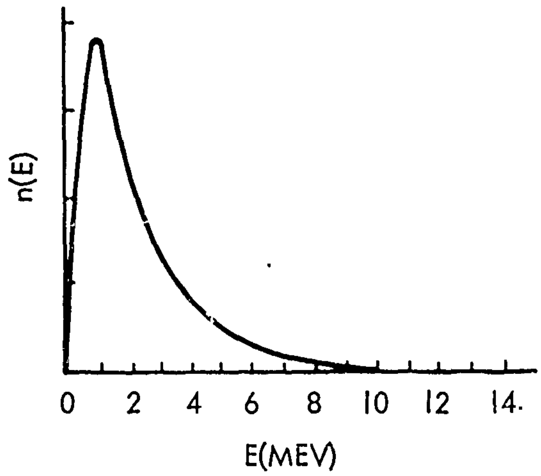


Figure 1. Spectrum of Neutron Energy within Nuclear Reactors

It is known that the damage caused by neutrons is proportional to

$$\int E \, n(E) \, \Sigma(E) \, dE$$

where the cross section $\Sigma(E)$ is a function of the probability that a neutron will collide with an atom to cause damage. $\Sigma(E)$ is inversely proportional to the square root of E except for energies below 1000 ev where it is approximately zero. The damage is, therefore, proportional to

$$\int_{0.001}^{\infty} \sqrt{E} \, n(E) \, dE$$

The function $\sqrt{E}\, n(E)$ is plotted in Figure 2 for the energy spectrum shown in Figure 1. From this graphical integration, plus knowledge of nuclear effects on transistors, it is determined that about half of the total damage to semiconductors is caused by neutrons having energies below 2 mev.

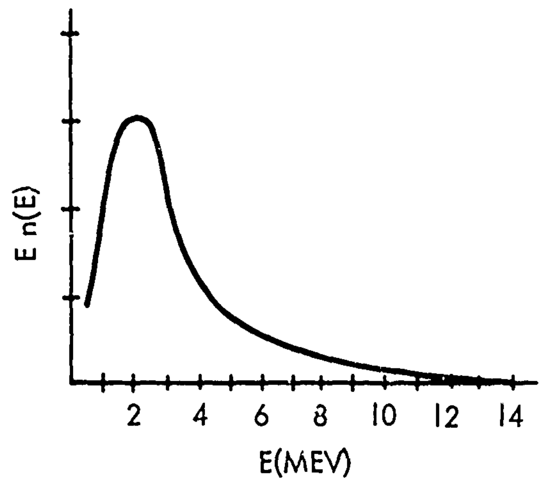


Figure 2. Function $\sqrt{E}$ n(E) for the Energy Spectrum Shown in Figure 1

Presently, particles that affect electronic equipment most seriously are photons (gamma ray) and fast neutrons. These particles are emanated from the nuclear reactor, but their intensity depends primarily on the type of shielding utilized and the power level of the reactor.

These factors are well understood and controllable through shielding techniques. (See Glasstone's Principles of Nuclear Reactor Engineering.) As a result, radiation intensity in the proximity of a reactor can be depicted with fair accuracy.

b. Space Radiation

Basic to the problem of nuclear radiation effects upon spacecraft electronic equipment is the environment to which a system will be exposed. Factors to be considered include:

Van Allen radiation belts
Galactic cosmic rays
Sporadic but intense solar proton beams
Artificial radiation belt
Electromagnetic radiation.

Pertinent data typically consist of principal radiant constituent, geographical location, and if applicable, its variation; and radiation intensity.

The trapped charged particles in the Van Allen belts constitute a particular hazard for spacecraft which must operate within these regions. Isointensity contours for these belts are shown in Figure 3, while Table I shows Van Allen's estimate of particle flux.

Table I. Van Allen's Estimate of Particle Flux

Particles	Energy	Intensity
	Heart of Inner Zone (3,600 km on geomagnetic equator)	
Electrons*	20 kev (max)	$2 \times 10^9/cm^2 sec/ster$
Electrons**	600 kev (max)	$1 \times 10^7/cm^2 sec/ster$
Protons**	40 mev (max)	$2 \times 10^4/cm^2/sec$
	Heart of Outer Zone (16,000 km on geomagnetic equator)	
Electrons***	20 kev	$1 \times 10^{11}/cm^2 sec$
Electrons	200 kev	$1 \times 10^8/cm^2/sec$
Protons	60 mev	$10^2/cm^2/sec$
Protons	30 mev	No information

*Accuracy within a factor of 10.
**Accuracy within a factor of 2.
***Theoretical considerations indicate this value may be too large.

3. EFFECTS OF NUCLEAR RADIATION

The damaging effect of high-energy electromagnetic or particle radiation, or both is determined by the energy, number of quanta encountered in some stated time period, and the physical properties of the materials. The materials most sensitive to radiation damage are the semiconductors. These materials compose the vital parts of the intelligence and power-plants on board a missile or spacecraft. They are used in the transistors and diodes which form the active components of solid-state telemetry units.

From Explorer VI Data During the Magnetically Quiet Period Before the Storm of August 16, 1959. The unlabelled points show locations at which fluxes were measured to be 2 x 10^7, 10^7, 2 x 10^6 and 2 x 10^5 electrons/cm^2/sec. Fluxes at points labelled 1 through 4 were measured by Pioneer III. At point 5 the energy spectrum and pitch angle distribution were determined in the Javelin flight. The outer zone boundaries were determined in region I by Sputnik III and in region II by Explorer XII.

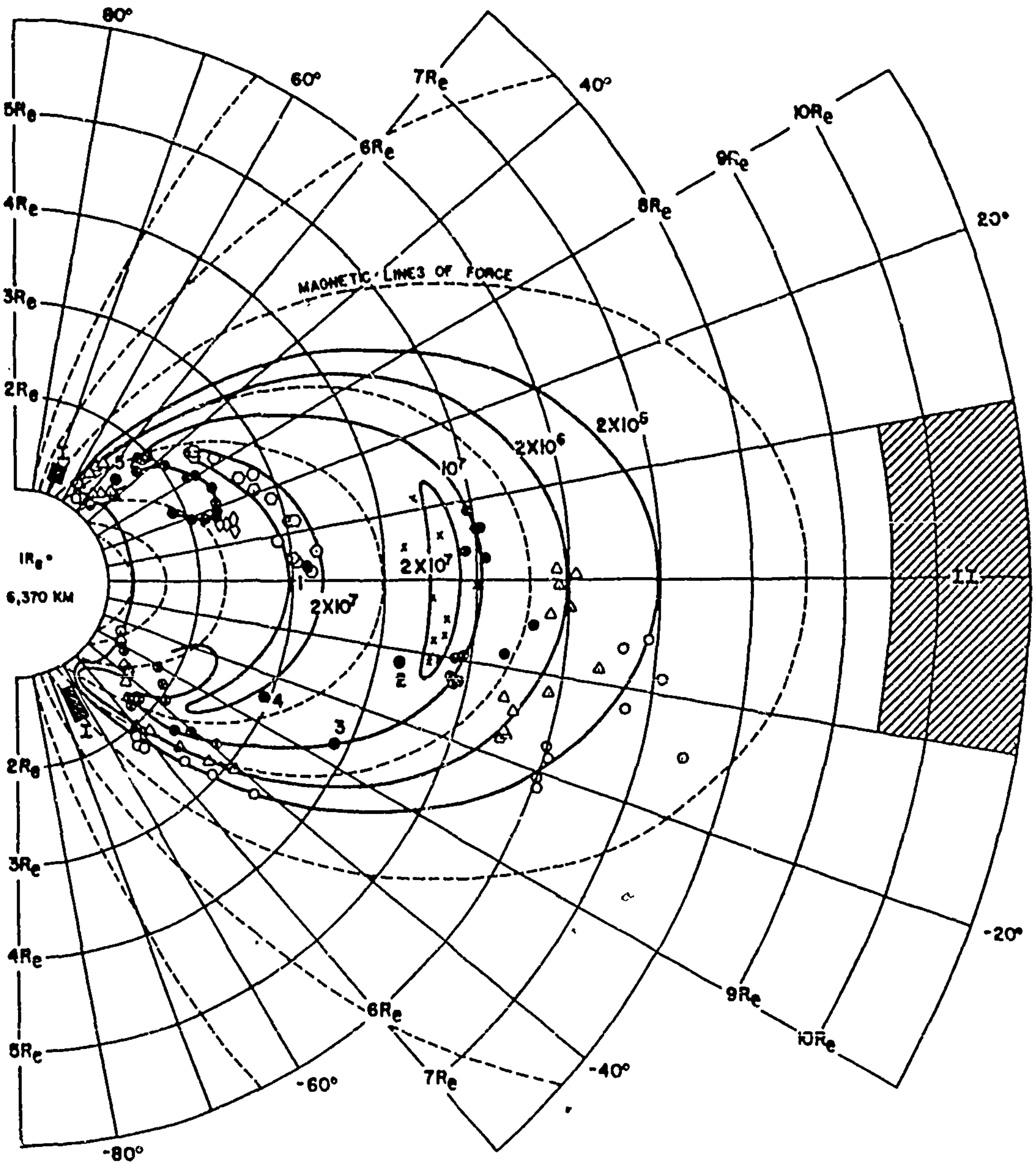


Figure 3. Isointensity Contours

Other sensitive materials include the plastics used as electrical insulators and/or dielectrics in interconnections and thin-film resistors, and capacitors forming the passive elements of a telemetry unit.

In semiconductor materials, much of the gamma-ray photoionization causes transient effects, raising the noise level in transistors and diodes, but decaying rapidly with removal of the gamma-ray source. However, some gamma-ray collisions cause displacement of atoms in the semiconductor crystal, resulting in permanent damage. Fast neutrons also cause such imperfections in the crystal structure and are, by far, the greater cause of damage. The effects of the crystal imperfections are to trap the majority current carriers. This affects the resistivity, the carrier lifetime, and the mobility of the majority carriers. The results are decrease in gain and increase in zero signal current (I_{co}) for transistors, and an increase in forward voltage drop for diodes and rectifiers.

Since any significant variations in transistor parameters will be reflected in the gain, many tables are formulated or focused to this parameter. Tables II and III give some typical results for neutron and proton irradiation of transistors.

The relative susceptibility of various electronic components to radiation damage is also of considerable interest. A partial list showing the levels at which radiation damage affects the operation of various components is shown in Table IV. The indicated orders of magnitude are only approximate, but do serve to indicate the trend.

4. IMPROVEMENT TO PRESENT ELECTRONIC SYSTEMS

Improving the operational capability of present-day electronic equipment in a nuclear radiation environment can be accomplished with the use of shielding techniques and radiation-hardened components.

a. Shielding Technique

The simplest protection against the presence of nuclear radiation is isolation through shielding. Unfortunately, the state of the art on shielding material does not justify this scheme, as weight and space are of paramount consideration in space vehicles.

Table II. Variations of Transistor Current Gain as a Function of Neutron Irradiation

Transistor	Type	Mfg	I_c (ma)	β/β_o at I_c and Indicated Netron Exposure (N cm^{-2})						
				$1(10^{12})$	$5(10^{12})$	$1(10^{13})$	$5(10^{13})$	$1(10^{14})$	$5(10^{14})$	$1(10^{15})$
L5405	PNP,GE	PH	50				0 86	0 64		
T1166	PNP,GE	PH	20			0 91	0 79	0 61		
T1075	PNP,GE	PH	5			1 00	1 00	1 00		
2N247	PNP,GE	RCA	30			0 43	0 196			
MACS	PNP,GE	PH	100				0 57			
2N139	PNP,GE	RCA	20			0 78	0 50	0 31		
T1041	PNP,GE	PH	500	0 50	0 21					
4JD1A17	PNP,GE	GE	100	0 73	0 44					
2N174	PNP,GE	DE	1000	0 46	0 15					
2N176	PNP,GE	MO	500	0 42	0 15					
H-7	PNP,GE	MH	1000	0 44	0 17					
2N94A	NPN,GE	SY	200		0 60	0 46	0 27			
HA5003	NPN,GE	HU	100		0 44	0 28				
4J4A5	NPN,Si	GE	10		0 13					
905	NPN,Si	TI	10	0 65		0 18				
952	NPN,Si	TI	10	0 21	0 17					
2N389	NPN,Si	TI	100	0 70						
ST-33	NPN,Si	TR	20			0 10				
ST-42	NPN,Si	TR	20			0 11				
T1257	PNP,Si	PH	50			0 78	0 42	0 34		
2N1385	PNP,GE	TI	10*							0 17
2N1141	PNP,GE	TI	1 0*						0 20	0 09
2N1195	PNP,GE	WE	2.0*						0 20	0 06
2N700	PNP,GE	MO								0 70
2N797	NPN,GE	TI							0 70	0 20
2N828	PNP,GE	MO							0 20	
2N1561	PNP,GE	MO							0 20	
2N834	NPN	TI							0 20	
2N753	NPN	TI							0 20	
2N1385	PNP,GE	TI							0 20	
2N706	NPN,Si	TI							0 20	
2N797	NPN,GE	TI						0 20		
2N736	NPN,Si	TI						0 20		
3N25	PNP,GE	TI	5			0 94	0 82	0 59		
2N384	PNP,GE	RCA	20		0 80	0 68	0 64			
2N167	NPN,GE	GE	5	0 63	0 45		0 07			
2N366	NPN,GE	TI	5	0 47	0 29	0 12	0.07			
2N336	NPN,Si	TI	10	0 56	0 32	0 09				
2N339	NPN,Si	TI	25	0 47	0 30					
2N335	NPN,Si	TI	10	0.64	0 32					
2N333	NPN,Si	TI	10	0 44	0 18					
2N332	NPN,Si	TI	10	0 80	0 45					
2N726	PNP,Si				0 98	0 93				
2N702	NPN,Si				0 99	0 86				
2N656	NPN,Si			0 83	0 44	0 30				
2N1154	NPN,SI			0 75	0 4					
2N710	PNP,GE				0 99	0 97				
2N697	NPN,Si					0 75	0 28			
2N797	NPN,GE					0 98	0 92	0 88		
2N705	PNP,GE		10*			0.35		0 35	0 10	

Index of Manufacturers

DE	-	Delco	RAY	-	Raytheon
GE	-	General Electric	RCA	-	Radio Corporation of America
HU	-	Hughes Products	SY	-	Sylvania
MH	-	Minneapolis Honeywell	TI	-	Texas Instruments
MO	-	Motorola	TR	-	Transitron
PH	-	Philco	WE	-	Western Electric

β/β_o = the ratio of the β after irradiation to the β before irradiation

*I_c in : ma

Table III. Variation of Transistor Current Gain as a Function of Proton Irradiation

Transistor	Type	I_B μa	Total Flux protons/cm^2	β/β_o	
2N128	PNP	20	$7.5(10^{12})$	0.86	22 mev protons
			$2.25(10^{13})$	0.35	
2N1303	PNP	30	$3(10^{12})$	0.65	
			$4.73(10^{12})$	0.27	
2N146	NPN	50	$7.5(10^{12})$	0.08	
			$1.5(10^{13})$	0.06	
			$2.25(10^{13})$	0.06	
2N169A	NPN	50	$7.5(10^{12})$	0.07	
			$1.5(10^{13})$	0.02	
			$2.25(10^{13})$	0.05	
2N1302	NPN	100	$7.5(10^{12})$	0.11	
			$1.5(10^{13})$	0.03	
			$2.25(10^{13})$	0.07	
2N128	PNP	30	$9.3(10^{11})$	0.94	240 mev protons
				0.97	
2N146	NPN	50	$9.3(10^{11})$	0.79	
				0.61	
				0.77	
2N224	PNP	75	$9.3(10^{11})$	0.34	
				0.39	

Table IV. Relative "Hardness" of Typical Electronic Components

Devices	Threshold Radiation: Fast Neutrons (nvt)	Threshold Radiation: Gamma Protons/cm^2
Transistors	10^{10} to 10^{13}	10^{14} to 10^{15}
Computer diodes	10^{13}	10^{14}
Rectifiers (thin-base silicon)	1.5 to 10^{15}	10^{15}
Electron tubes	5×10^{15} to 10^{18}	
Capacitors	10^{15} to 10^{18}	
Transformers	10^{18}	
Resistors	10^{18}	
Gyros (floated)	5×10^{14}	1×10^{16}

b. Radiation-Hardened Components

The use of radiation-resistant components in the electronic equipment is a step forward in the development of radiation-hardened equipment. Fortunately, extensive research programs have produced many components capable of withstanding high degrees of nuclear radiation. By integrating these components into the equipment design, a higher degree of confidence in operational integrity is achieved.

It will be necessary to incorporate the known radiation-resistant components into the present vehicle-borne tracking, telemetry, and command equipment. The difficulty associated with this technique is that radiation-hardened material, when exposed to radiation, characteristically exhibits certain dangerous phenomena; but when combined with other devices under power, its characteristics are highly modified.

Also, components which are known to be radiation-hardened may not be good substitutes for the original components (e.g., possess an undesirable temperature coefficient). Substitution of active elements poses an even greater problem because of the many varied parameters to be considered prior to substitution. In short, the development of radiation-hardened electronic equipment is not simply the direct substitution of components, but requires careful analysis.

5. DESIGN CRITERIA

In view of the factors discussed above, it is suggested that emphasis be given to design criteria, rather than to improvements in present electronic equipment.

One approach to the establishment of design criteria is as follows: detailed steps would be given for a designer to follow in order to satisfactorily meet the objective. This procedure would be similar to one he is presently following and would differ only to accommodate the new nuclear radiation stress factor. In this respect, the designer will be introduced to the use of the radiation profile curve for the selection of components. (This approach is analogous to a temperature profile in which a silicon transistor is selected over germanium for use in a high-temperature environment.) He will recognize the fact that many pertinent factors concerned with the design of the equipment will be contained in this radiation-profile curve. The limitation associated with the radiation-profile curve woul' ' also be discussed.

Use of magnetic cores would be clearly recommended in the radiation-profile curves because of their immunity to radiation, both in steady-state and transient conditions. Steps would be included to evaluate the pacing item with respect to radiation. Provision to consider alternate approaches should not be overlooked: e. g., a check point inserted to permit the designer to evaluate his status with respect to his end requirement.

Considerations of other environmental factors, such as temperature, altitude, shock, etc., should not be excluded from the procedure. It should be pointed out however, that data pertaining to component behavior in a radiation environment, coupled with other environmental factors, are not too plentiful.

During the optimization process, an important step (which is not usually found among radiation-free environment criteria) should be included. This step requires the designer to be cautious of any component changes that may be suggested during the optimization process. This is necessary because the range of values of the components is an important consideration in component selection. For example, the radiation effect on a certain type of resistor may differ markedly for different resistance values of the same resistor type. Subtle factors such as this can easily be overlooked, but are important in the design of radiation-resistant electronic equipment.

Also, the practice of using inverse feedback for compensating component degradation will continue to be a valuable design technique. In the normal environment, design engineers use this technique to account for transistor β variation due to aging, temperature, etc. The design criteria would emphasize that the technique can still be applied in the nuclear environment.

A final key point to establish in the design criteria is: when and how the designer determines he has reached the state of the art in the availability of a certain component with respect to radiation. This question is of prime concern to any designer of radiation-hardened electronic equipment.

1. A. Altshuler "Atmospheric Gamma-Ray Background," STL 7320.3-856, 5 December 1960.

2. A. Altshuler "Electromagnetic Propagation in Ionized Media," STL GM-00-4200-00379, 23 June 1955. (C)

3. S. Altshuler "Space Environmental Effects on Atlas," STL 6101-6689-MS-000, 12 March 1962. (S)

4. S. Altshuler "Topics for Theoretical Studies of the Interplanetary Plasma," STL 7320.3-500, 29 June 1960.

5. S. Altshuler "Upper Atmospheric Ultraviolet Detection," STL 7320.3-832, 8 December 1960.

6. Arnoldy, Hoffman and Winckler, "Solar Cosmic Rays and Soft Radiation Observed at 5,000,000 km from Earth" [inc].

"A Comprehensive Map of the Space Radiation Environment," STL 8644-6002-RU-000, 28 September 1962.

8. T.A. Farley "Soft Radiation Measurements on Explorer VI Earth Satellite," published as an Abstract in the Bulletin of the American Physical Society, Ser. II, Vol. 5, p 46, January 1960.

9. T. A. Farley and N. L. Sangers "Pitch Angle Distribution and Mirror Point Densities in the Outer Radiation Zone," Journal of Geophysical Research (in press).

10. G. W. Ford "Estimation of the Flux of Energetic Protons Above the Lower Radiation Belt," STL 5101-0040-RU-000, December 1961.

11. L. Glatt "Parameters Pertinent to the Detection of Ultraviolet Radiation in Space-to-Space Configurations," STL 6110-7244-RU-000, June 1962. (U)

12. C. Graves "Pyrolysis and Radiation Damage of Polytetrafluoroethylene (Teflon)," STL GM 47.3-83, 9 February 1959.

13. R. A. Howard "Possible Sources of Charged Particles in the Radiation Zones," STL 7320.2-97, November 1959.

14. J. Lindner "NASA Radiation Background Measurement Program," STL 7320.3-150, 14 December 1959.

15. J. Lindner "Attempted Measurements of Radiation Dosage During Missile Flights," STL TN-59-0000-00297, 31 July 1959.

16. J. Lindner "Composition of High Altitude Radiation Belt," STL GM-47-66, 2 September 1958.

17. J. Lindner "Cosmic Ray Experiments in Space," STL GM-TM-0165-00305, 8 October 1958.

18. J. Lindner "Rocket and Satellite Explorations of the Upper Atmosphere," STL GM-TM-0165-00306, 24 October 1958.

19. G. Munch "Contents of Interplanetary Space," STL GM-TM-0165-00301, 8 October 1958.

20. A. Rosen "Fluctuations in Intensity of High Altitude Radiation in the Trapped Radiation Zone," published as an Abstract in the Bulletin of the American Physical Society, Ser. II, Vol. 5, p. 46, January 1960.

21. A. Rosen "Dynamics of the Outer Van Allen Radiation Zones," published as an Abstract and presented at the American Geophysical Union Meeting of April 23 in Washington, D. C.

22. A. Rosen "Scientific Results from Magnetometer, Ionization Chambers and a Scintillation Counter on Pioneer I, II, and V Space Probes, and the Explorer VI Earth Satellite," published as an abstract and presented at the A. G. U. meeting in Berkeley, California.

23. A. Rosen "Results of Radiation Experiments on Earth Satellites," an invited paper presented at the American Physical Society meeting in December 1959.

24. A. Rosen and T. A. Farley "Observation of Soft Electrons from Explorer VI," presented at the Midwest Cosmic Ray Conference at the State University of Iowa, October 1959.

25. A. Rosen "Scintillation Counter Measurements," An Invited paper presented by T. A. Farley at the American Physical Society held November, 1959 in Cleveland, Ohio.

26. L. Scott "Average Number of Nucleons in Cosmic-Ray Induced Nucleon Cascades," STL 6120-7257-NU-000, February 1962.

27. E. Smith and A. Rosen " The Scientific Results of the Satellite Explorer VI," Proceedings of the 5th Ballistic Missile and Space Technology Symposium [inc] .

SPACE ENVIRONMENT DATA

1. T. A. Farley and A. Rosen "Charged Particle Variations in the Outer Van Allen Zone During a Geomagnetic Storm," Journal of Geophysical Research 65, 3494, October 1960.

2. A. Rosen, C. P. Sonett, P. J. Coleman, Jr., and C. E. McIlwain . "Ionizing Radiation at Altitudes of 3,500 to 36,000 Kilometers - Pioneer I," Journal of Geophysical Research, Vol. 64, No. 7, July 1959.

3. A. Rosen, P. J. Coleman, Jr., and C. P. Sonett "Ionizing Radiation Detected by Pioneer II," Planet Space. Sci. Pergamon Press 1959, Vol. 1, 343-346, also printed in Great Britain

4. A. Rosen, T. A. Farley, and C. P. Sonett "Soft Radiation Measurements of Explorer VI Earth Satellite," Proceedings of the 1st International Space Science Symposium, P. 938, North Holland Publishing Company, Amsterdam, 1960. Also presented at the Symposium at Nice, France, January 1960.

5. A. Rosen and T. A. Farley "Characteristics of the Van Allen Radiation Zones as Measured by the Scintillation Counter on Explorer VI," Journal of Geophysical Research (In press). July 1961.

6. A. Rosen "Geomagnetically Trapped Radiation and Interplanetary Cosmic Flux," Medical and Biological Aspects of the Energies of Space, Columbia University Press, 1961. Also presented at Energies of Space Symposium, San Antonio, Texas, September 1960.

Appendix XXV. LONG-HAUL COMMUNICATIONS

A. INTRODUCTION

This appendix is a survey of the capabilities of various media of long-haul communications for a global network of ground stations. The following areas are covered:

- Communication satellites
- HF radio
- Submarine cables
- Microwave radio relay
- Trans-horizon scatter propagation
- VLF/LF radio
- Lasers

The characteristics and the state of the art of each have been discussed as well as their anticipated capabilities in terms of bandwidth, reliability, and quality for the 1965 to 1970 time period. Some of the historical background has been given so as to indicate the amount of activity in development, especially recent development in a particular transmission medium. Reasons are given for not using a particular medium for the global range network, and advantages and disadvantages of other media are presented. Costs are developed so that comparisons between the several media can be made, since economic considerations are always of importance.

B. COMMUNICATION SATELLITES

1. BASIC CONSIDERATIONS

Communications by means of satellite radio relay has been shown to be technically feasible. Test data from the early experimental satellites such as Telstar, Relay, Syncom, and Echo indicate that future communication satellite systems can meet requirements for large quantities of long range telephone circuits, as well as high-speed digital data and television. Under NASA, Air Force, private industry, and foreign participation, the technology of communication satellites is advancing rapidly. With adequate priority and funding, some type of a global communication system using satellite relay could be realized in the 1965 to 1970 period.

Two distinct groups of orbiting earth satellites are considered: active and passive relays. For ease in comparison of the communication capability of active and passive relays, equations relating power received at one surface station to that transmitted at the other are given:

$$P_R \text{ (active)} = P_T G_T \frac{\lambda^2}{(4\pi R_1)^2} \left[g_r A g_t \right] G_R \frac{\lambda^2}{(4\pi R_2)^2}$$

$$P_R \text{ (passive)} = P_T G_T \frac{\lambda^2}{(4\pi R_1)^2} \left[\frac{4\pi\sigma}{\lambda^2} \right] G_R \frac{\lambda^2}{(4\pi R_2)^2}$$

where

P_R = power received at the surface receiver

P_T = surface-transmitter power

G_T, G_R = effective power gain of the surface transmitter and receiver antenna

R_1, R_2 = distances from transmitter and receiver to the relay

A = power gain through the relay

λ = wavelength at the signal

g_r, g_t = power gain of relay receiving and transmitting antennas

σ = scatter cross-section of the passive relay.

The terms in the brackets have the same overall significance in both cases. They give the effective gain of the relay as a receiving-reradiating element. The net gain at the antenna aboard the active relay is $g_r\ g_t$ The gain of the passive reflector is simply that of a receiving-reradiating antenna of effective aperture σ. Obviously the greatest difference between the two is the power gain provided in the active case. This differential can approach 50 to 100 db in practical cases.

There are other distinct capabilities that differ in the two cases which may actually negate the apparent power gain advantage of the active relay. For example, the passive relay can often accommodate a far broader frequency band, hence many more communication channels. Reliability and cost aspects are also involved.

a. Passive Relays

The only electrical difference among various types of passive relays is in the nature and magnitude of the scatter cross section. In Table I the several candidates for passive relays are compared. Weight

Table I. Scatter Cross Section

Item	Cross Section in Sq Meters	Relative Gain Over Large Spherical Shell (db)
Large spherical shell 41 meters diameter	1360	0
100 spherical shells 4.12 meters diameter	1360	0
1.868×10^4 spherical shells 3.0 cm diameter	1.585×10^4	+ 10.7
4.25×10^8 spherical shells 0.2 cm diameter	0.187	-38.6
1.29×10^{13} solid shot 0.001 cm diameter	6.15×10^{-10}	-124
2.1×10^9 half-wave dipoles 0.001 cm diameter random orientation	3.15×10^6	+ 34
Flat circular plate 82 meters diameter	3.5×10^{10} (max)	+ 74
Total mass: 350 lb		
Wavelength: 10 cm		

being one of the most significant parameters related to satellite launchings, this quantity has been held constant for each of the examples in the table. The reference scatter cross-section and weight is a 135-ft aluminum-coated mylar balloon. This has a large electromagnetic reflectivity and a scatter cross-section equal to its projected circular area (1360 sq m). At a frequency of 3000 mc this represents a reflection gain of 62.3 db over an isotropic receiving-radiating system. In order to establish a reference, a 10-kw transmitter would produce a satisfactory voice channel for a great circle range of 5000 miles using such a relay.

It must be noticed that the flat circular plate, with its apparent improvement of 74 db over the sphere, may encourage the drawing of erroneous conclusions. The plate is highly directive; for one degree off normal the scatter cross-section drops 62.4 db.

From Table I it may be determined that dispersed orbiting dipoles have the highest ratio of reflection cross-section to weight ratio. Figure 1 shows the scatter cross section per pound of material in two typical cases at a nominal frequency of 3000 mc. The research project West Ford is currently investigating the characteristics of a dipole belt as a satellite relay.

An example of relay dipole-belt performance for a station-to-station range of approximately 6500 nautical miles is shown in Table II. One thousand pounds of dipoles is assumed in orbit. The RF bandwidth is about 300 mc at a carrier of 3000 mc. Two effects peculiar to dispersed dipoles complicate proper utilization of this bandwidth. One is the significant difference in return-path lengths from dipoles in the antenna beams to the surface station. This causes multipath delays of fractional milliseconds. The other is the spread of doppler shifts through the beam which introduces radio frequency noise in the signal. To cope with these effects, the present state of the art requires frequency-hopping techniques which use more bandwidth than would be needed otherwise.

b. Active Relays

The active relay, as can be seen in the power equation, has substantial gain in the relay itself. Given the identical ground transmitters, antennas, and receiving equipment, the received signal as compared to the

Table II. Characteristics of Typical Orbiting Dipole System

Orbit altitude	4000 nmi
Frequency	3000 mc
Modulation	FDM*
Transmitted power	20 kw
Transmitted antenna gain	52 db
Receiving antenna gain	52 db
Receiver noise density	-209 dbw/cps
Receiver noise bandwidth	50 kc
Receiver noise power	-152 dbw
Scatter cross section	9.7×10^3
Equivalent noise power	-263 dbw
Capacity	50, 000 bits per second

* Special provision for frequency-hopping

passive relay may well be greater by 100 db. These systems have had initial experimental programs in the NASA Relay and Syncom and the AT and T Telstar. The Relay and Telstar programs have been extremely successful, not only in advancing satellite technology, but also in one of the major areas of effort, the establishment of ground stations located in and built by foreign countries. Stations have been or are being established in England, France, Germany, Italy, Brazil, Japan, and the U.S. During the first year of experimental operation, excellent cooperation with all participants was achieved.

Although major advancements are being made in communication satellite technology, a number of areas require advancement before an operational system can be realized. These are:

Satellite equipment reliability
Repeater configuration
Establishment of a global network of ground stations.

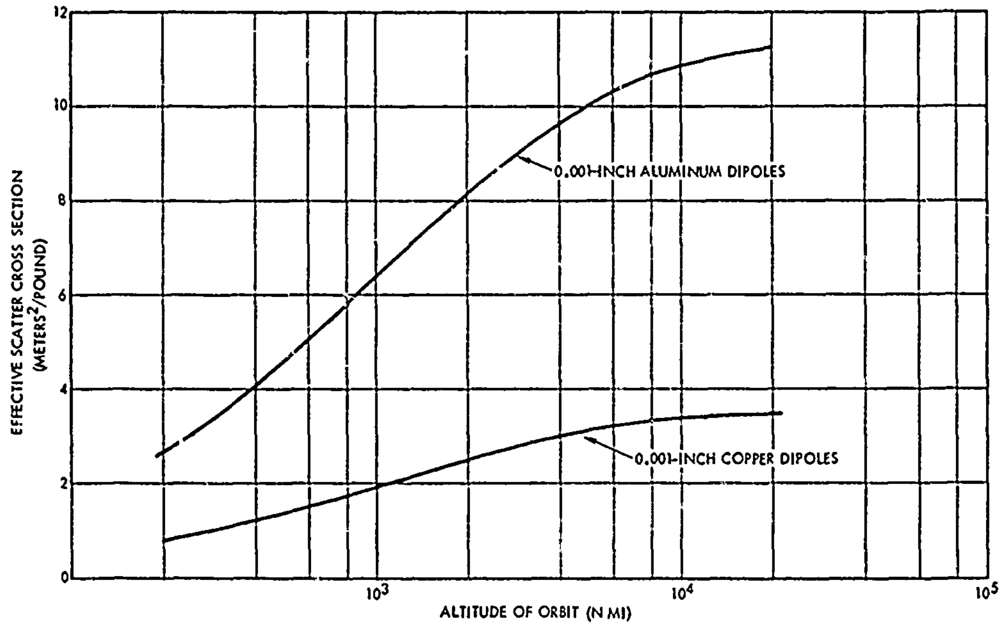


Figure 1. Scatter Cross Section of Orbiting Dipoles Per Pound of Material

1. Satellite Equipment Reliability

The active communication satellite contains several subsystems in addition to the repeater, all of which greatly affect its lifetime. The major items are: communications repeater; power supply; attitude control; and station keeping.

2. Communications Repeater

Repeater configurations are covered in Section B. 2 below. The reliability of the transponder itself, exclusive of the power amplification device, shows high promise of meeting long lifetime requirements. Solid-state devices can be employed for all active elements of the repeater except for the power stage. Traveling wave tubes for power levels of 1 to 20 watts show promise of meeting the minimum requirements for efficiency and long life.

3. Power Supply

The demand for energy sources for satellite applications over the past few years has led to studies of many methods and sources. Some of the combinations can be discarded immediately as not practical before 1965 or not suitable for long-term operation. The remaining methods are shown in Figure 2 as functions of specific weight to obtain power output for long periods of time. Oriented solar cells and thermoelectric generators using a radio isotope as the energy source offer substantial benefits in efficiency for these requirements. Since the satellite will be eclipsed from the sun during a portion of its life, a means for storing power must be available in solar-cell systems.

(a) Solar-Cell Battery Supply. Although solar cells can be made from a variety of semiconductors, at present only silicon cells are available commercially. Other semiconductors, such as gallium arsenide which has an ionization voltage near the theoretical optimum of 1.35 volts, are anticipated to give somewhat better efficiency than silicon, but preparation techniques for these materials are not far enough advanced, nor will be in the next few years, to enable fabrication of solar cells having efficiencies as high as those obtained with silicon.

Of the many techniques for mounting assemblies to a spacecraft, the best method from an efficiency standpoint is to mount the cells on a flat panel which is continuously oriented to obtain perpendicular

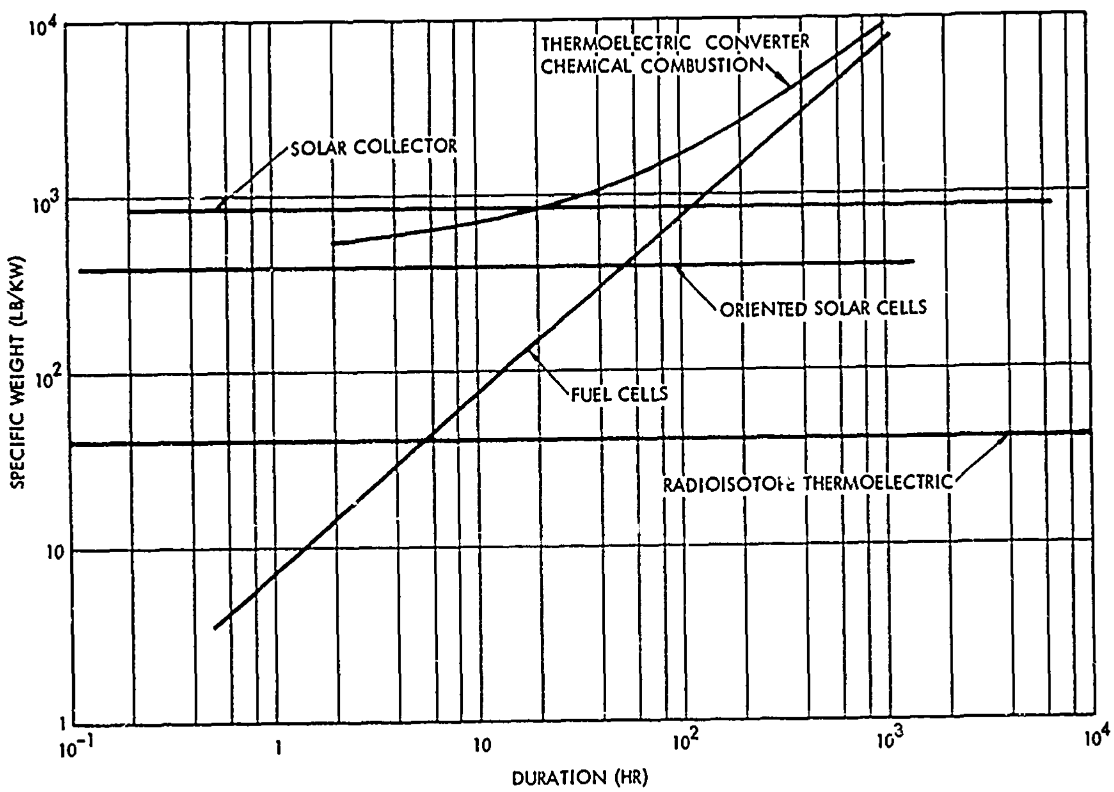


Figure 2. Specific Weights of Various General Types of Available Power Supplies

incidence of solar radiation. A power output of 8 to 10 watts per square foot of panel area can be obtained in near-earth orbits. Although this requires continuous attitude control of the spacecraft, it is likely that attitude control will be required for the directional antennas.

Sun-oriented solar arrays pose two problems to the satellite configuration design: an electromechanical drive is needed to keep the array properly oriented at all times, and a folding mechanism is required to fit the array within the booster fairing during launch. There are associated problems such as protection of the solar cells during unfolding and locking, and supporting the folded array to withstand boost loads and accelerations. The penalty in weight paid for the mounting and orienting structure for sun-oriented arrays is overcome substantially by the reduced requirements for solar cells.

When solar cells are used to supply primary power, storage batteries or radio-isotope thermoelectric generators are needed to supply power during eclipse and to supply short-duration peak power demands. Sealed nickel-cadmium batteries are at present the only tried and tested means of providing electric power aboard satellites during eclipse. They have been used on almost every earth satellite which has been launched to date for power storage and peak loads. The charge-discharge efficiency of sealed nickel-cadmium cells depends on the depth of discharge, charge and discharge rates, and cell temperatures. Efficiency will decrease for very long recharge periods or if charging continues after the cell is fully charged. One of the primary concerns in satellite applications is the cyclic life of the battery. At an altitude of 5600 nautical miles one can expect about 1500 eclipses per year. In addition to discharge during eclipse, partial battery discharge will occur during periods of high power drain whether the satellite is eclipsed or not. The design practice of using an oversized battery is frequently employed so that sufficient capacity will be available even if the battery capacity is decreased after repeated charge-discharge cycles.

The silver-cadmium battery cell is a relatively new development and is expected to yield several advantages in storage applications. It offers a compromise between the high cyclic life, low specific capacity of the nickel-cadmium cell and the high-capacity, low cyclic life of the silver-zinc battery. Sufficient test data is not available to confirm the design objective of the cells.

(b) Radio Isotope Power Systems. As shown in Figure 2 radio-isotope thermoelectric generators (RTG) will also meet the power requirements for communication satellites. However, RTG's are not expected to reach full operational capacity during the initial part of the time period specified. In addition, their use would pose a major problem in the configuration design because of the possible radiation damage to other components. For this reason, the RTG must be located as far away from the other components as possible and/or shielded from the other components with a consequent weight penalty.

4. Attitude Control

The simplest method of controlling the attitude of a satellite is to spin it; the momentum imparted by the spin will cause the spin axis to remain fixed in inertial space. This method has been successfully used on such first-generation satellites as Relay and Syncom. In the case of a stationary orbit, if the spin axis is to be kept oriented along the earth's local vertical, it is necessary to provide a constant impulse to cause the satellite to precess 360° per day. With presently possible impulse techniques, such precession would require a prohibitive weight. In addition, spin stabilization vertical to the earth would provide at best a solar-cell power supply efficiency of 25 percent of the sun-oriented array. Orienting the spin axis toward the sun would require achieving a precession rate of only 360° per year and would permit the same solar cell power supply efficiency as an oriented array, but the necessity in this instance to separately orient the satellite antennas if directivity is desired would again be costly in weight. In addition to the weight needed to maintain antenna orientation, approximately 2500 lb-sec of pneumatic system impulses would be required to maintain the spin axis orientation to the sun of a 500-pound satellite for a year. This would be the equivalent of 100 pounds of nitrogen in an optimum system. As the satellite orbits the earth, it is affected by several torque-producing influences: solar radiation, gravity gradients, magnetic fields, micrometeorites, internal rotating parts, and atmospheric drag. With the addition of mass expulsion, each of the torques, except that due to micrometeorites, could be used to provide control torques, or these can be treated as disturbance torques and their effects overcome by some other torque producing system.

A satellite whose principal moments of inertia are not equal will tend to orient one of its axes along the local vertical. This effect is caused by the gradient of the earth's gravitational field. However, the effect diminishes rapidly with altitude, and above 1000 miles the torques are more difficult to apply effectively. In the Tiros satellite, the earth's magnetic field was used to control system torques. This field, like the gravity-gradient field, decreases as $1/r^3$; therefore, satellites at higher altitudes would have to carry very large magnetic dipoles in order to obtain sufficient torques. There is a possibility that this type of control system could be used in high-altitude satellites.

At low altitudes, it is possible to use the torques created by air molecules for attitude control, but again, this source of torque is not available at high altitudes. In addition, such a method would introduce the need for variable orientation of solar-cell arrays with respect to the plane of the orbit.

Attitude control by mass expulsion (the release of cold or hot gas, ion propulsion, or plasma propulsion) is also a possibility. Of such methods, cold-gas systems are used most extensively, since they are understood most fully and appear to offer the highest reliability. Such systems, however, require more weight than a combination of cold gas and reaction wheels for spacecraft with long lifetime goals.

An inertial flywheel driven by a motor provides storage of momentum which can be used either for damping or storing cyclic disturbance torque variations. In combination with a mass expulsion system, the reaction wheel eliminates the need for releasing gas during cyclic torque variations and provides the capability for removing constant disturbance torques. Without a mass expulsion system, the size of an inertial flywheel and its motor drive to handle torques for long operation would be very large. By using a combination of the momentum storage capabilities of reaction wheels and the cold-gas nozzle thrust, one can achieve what appears to be the minimum weight for attitude control systems with lifetimes on the order of a year.

5. Orbit Control

To maintain control of orbital position requires first that the orbit error must be sensed, usually by ground tracking. Next an attitude sensing system is required on the satellite, and an attitude control system may be necessary to reorient the satellite before applying a corrective thrust. A propulsion system is then used to apply the corrective thrust, either by a cold-gas or hot-gas system. The weight tradeoff possible in the choice of the orbit control propulsion system is suggested by Figure 3, which plots various weight factors as a function of the velocity available from a typical cold-gas system (I_{sp} = 60) and a typical hot-gas system (I_{sp} = 240). This figure is also useful to indicate the weight of hot- and cold-gas attitude control systems using reaction jets, through the interrelation of the linear velocity ΔV and the angular momentum requirements of the attitude control system.

In arriving at a decision on the specific type of orbit control system, tradeoff studies as typified in the discussion of power supplies must also be made in such areas as propulsion system weight, space requirements, fuel or gas characteristics, reliability, etc. Propellant or pressurizing tanks usually take considerable volume from the satellite and require early consideration for efficient integration into the satellite design.

2. REPEATER CONFIGURATIONS

Perhaps the simplest way to classify satellite repeater configurations is with respect to the number of radio channels which the satellite provides. A radio channel is defined for our purposes as a separate and distinct path between the receiving antenna multiplexer (if any) and the point in the transmitter where the outputs(s) are first translated to radio frequency. That is, our definition of a radio channel allows for RF multiplexing of the channels in either receiver or transmitter or both, but not at IF, for example. A typical radio channel is depicted in Figure 4a.

With this definition in mind it is obvious that there are two extremes in the number of radio channels which a satellite can carry. The first results when a radio channel is supplied for each baseband channel (voice, teletype, or data) which passes through the satellite repeater. The second

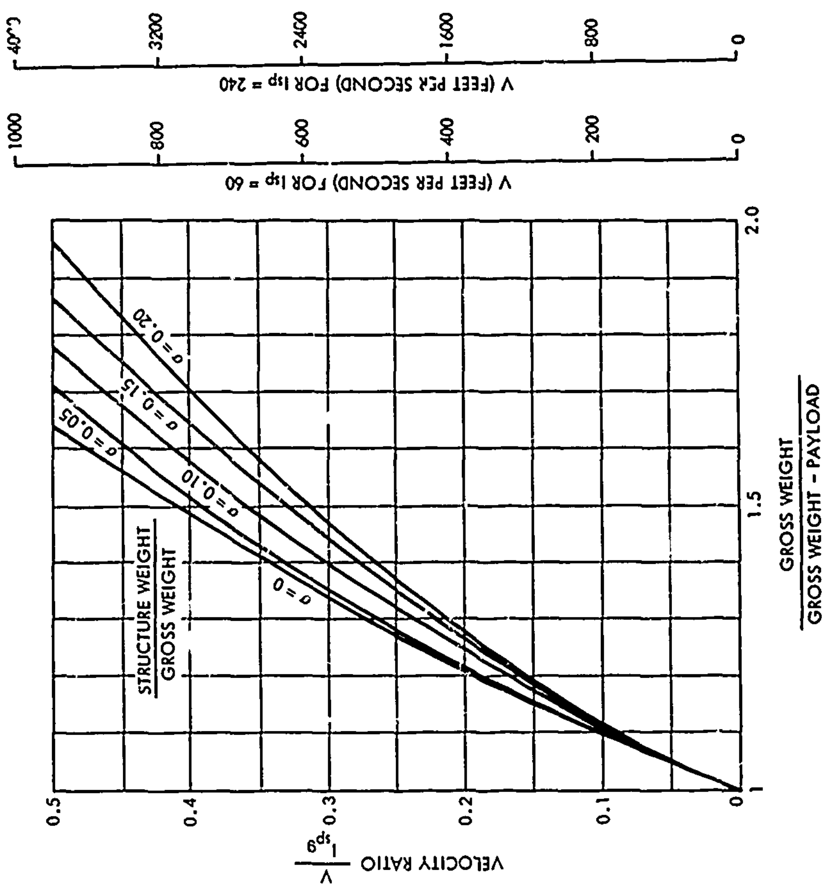


Figure 3. Hot- and Cold-Gas Control System Weights

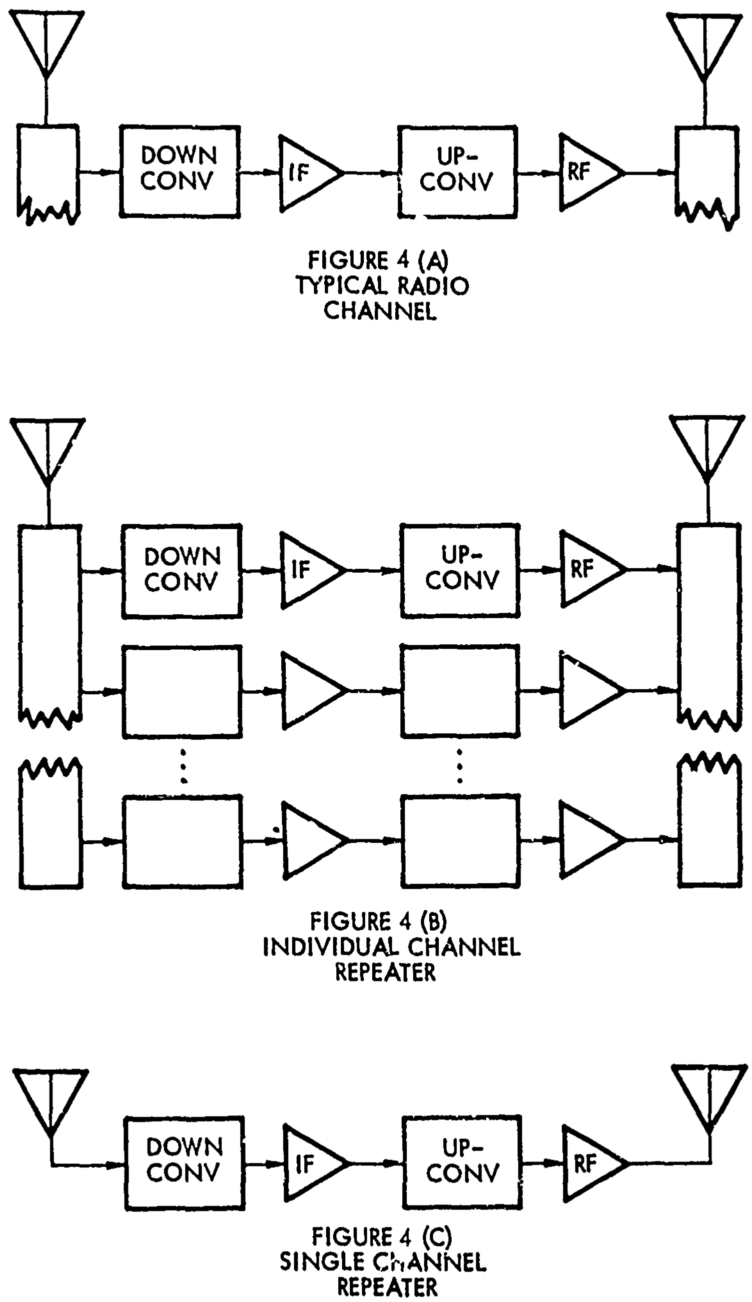


Figure 4. Radio Channel Arrangements

extreme occurs when a single radio channel is provided for all the traffic relayed by the satellite. The two extreme cases in the number of radio channels are depicted in the satellite repeater configurations of Figures 4b and 4c. In general, these will be impractical configurations but they provide a basis for comparison.

Another convenient division of satellite repeater configurations is with respect to how the baseband channels are grouped within a radio channel. The most obvious and straightforward grouping is done at the transmitter. That is, the baseband channels from each transmitting ground station are combined either by time or frequency division multiplex (TDM or FDM), and modulated on a radio frequency carrier. If these baseband channels are destined for a number of ground stations, then either the satellite or the ground station must separate the baseband channels. Alternatively, the transmitter could group channels according to their destination, thereby simplifying the selection process at the receivers. These two possiblities may be characterized as transmitter-grouped and receiver-grouped transmissions. Their effect on the satellite repeater will be investigated in Section B. 5 below.

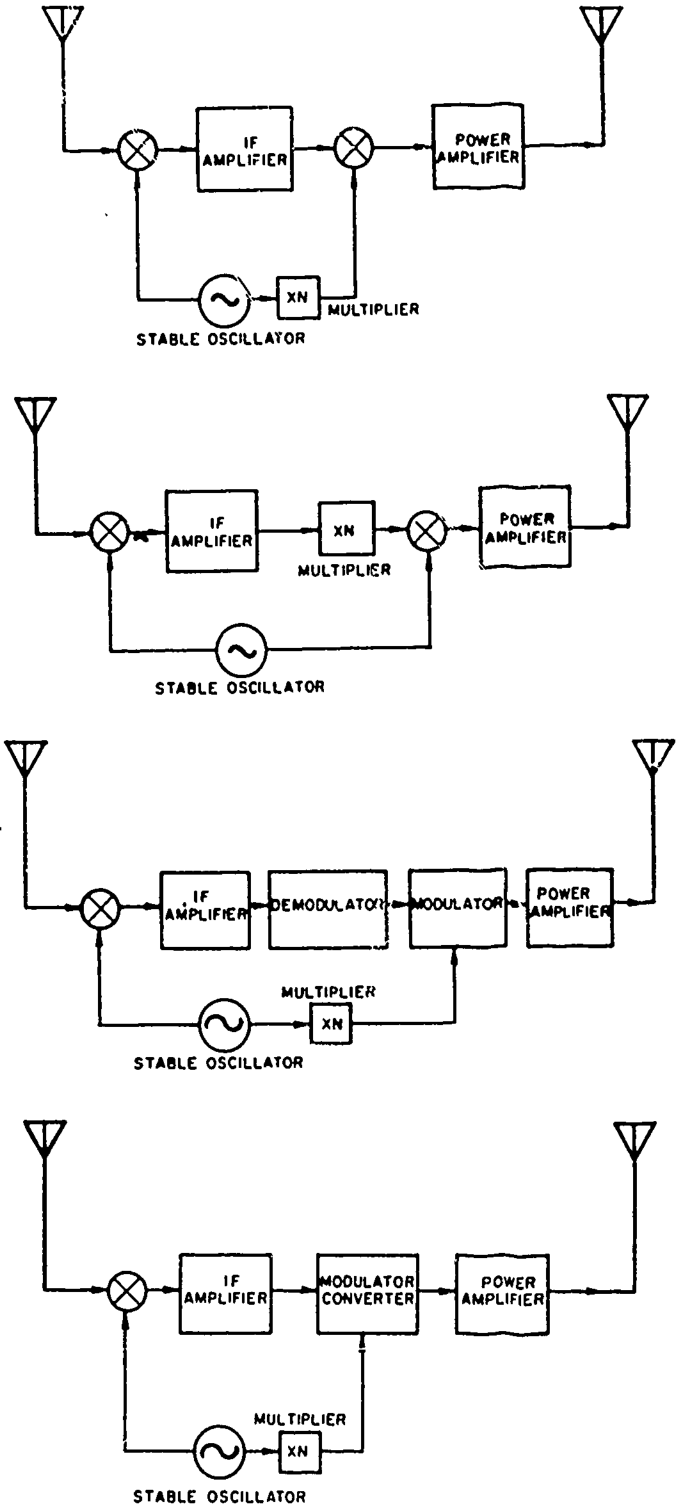
3. TYPES OF REPEATER CHANNELS

Before embarking on a discussion of various configurations of multiple-access satellite repeaters, it would be well to review the basic types of individual repeater channels. A thorough discussion of this subject has been given in Reference 1 and the intent here is to briefly review those findings. The basic repeater arrangements are the following:

The linear-translator repeater
The frequency-multiplier repeater
The demodulator-modulator repeater
The modulation converter repeater.

These four alternative configurations are depicted in Figure 5.

The linear-translator repeater amplifies the incoming signals linearly and translates them in frequency before re-radiating them to the ground. With this definition we have ruled out the translator utilizing hard-limiting from consideration in this category. Because of these characteristics the linear-translator is a very flexible configuration



(a) Linear Translator

(b) Frequency Multiplier

(c) Demodulator-Modulator

(d) Modulation Converter

Figure 5. Types of Repeater Channels

capable of transmitting almost any form of modulated signal. It thus provides growth potential in a system where one form of modulation must be accepted for an interim solution whereas another type of modulation may ultimately be used. One disadvantage of the linear repeater is that the final transmitter stage must also be linear and must therefore be operated well below saturation for the various forms of amplitude-modulated signal. For frequency and phase-modulated signals no such limitation exists; some thought, however, must be given to output levels to avoid intermodulation between signals and AM-to-PM conversion.

For a full-load test tone (simulating the actual communication signal) which frequency-modulates the carrier to a peak deviation f_d, the post detection signal-to-noise ratio at the baseband output of the ground receiver has been shown to be

$$\left.\frac{S}{N}\right)_{post} = 3M^2 \left.\frac{S}{N}\right)_{pre} \tag{1}$$

where $M = f_d/f_m$, and f_m is the maximum baseband frequency passed by the system. For the predetection S/N ratio the noise is measured in a band of width $2f_m$ centered at the carrier frequency and the satellite and ground contributions are indicated by

$$\left.\frac{N}{S}\right)_{pre} = 2f_m\left(\frac{\Phi_s}{S_s} + \frac{\Phi_g}{S_g}\right) \tag{2}$$

where Φ_s and Φ_g are the noise spectral density in the satellite and ground receivers, respectively, and where S_s and S_g are the signal powers measured at the same points in the satellite and ground receivers, respectively. The system is assumed to be operating above the threshold for the particular detection system employed.

The frequency-multiplier repeater produces a difference in input and output radio frequencies by frequency multiplication. It is therefore generally limited to FM and PM systems because limiting is usually employed. Since the frequency-multiplier transmits angle-modulated signals only, the amplitude linearity of the repeater is of little importance and the output amplifier stage in the transmitter can be used efficiently. On the other hand, it is clearly not as flexible as the linear repeater in handling various forms of modulated signals.

The frequency-multiplier repeater also has a disadvantage with regard to signal-to-noise ratio as compared with the linear repeater. It has been shown that although Equation (1) above holds also for the frequency-multiplier, the expression for the predetection N/S ratio is

$$\left(\frac{N}{S}\right)_{pre} = 2f_m \left(n^2 \frac{\Phi_s}{S_s} + \frac{\Phi_g}{S_g}\right). \tag{3}$$

That is, the noise contribution of the ground-satellite link is increased by the square of the frequency multiplication factor n. Thus to produce the same effective S/N ratio on this link as for the linear translator, the ground transmitter power must be increased by a factor n^2. The threshold requirements for the frequency-multiplier also turn out to be more severe than for the linear-translator repeater. Clearly, the frequency-multiplier repeater would be most desirable in a case where a high-power, narrow-bandwidth device such as a klystron was the ground transmitter, while a low-power, wideband device like a traveling wave tube was the satellite transmitter.

The demodulator-modulator, as its name implies, demodulates the incoming signal and then modulates the transmitted carrier. For our purposes we will assume that the incoming and outgoing types of modulation are the same; otherwise this repeater might equally well be called a modulation converter. Demodulation of the signal to baseband allows some processing of the information in the satellite. In a digital system, for instance, regeneration of the pulses could eliminate the noise added to the signal in the ground-satellite link. Also in a multiple-access satellite, frequency-division signals from several transmitters could be restacked in frequency, or time-division signals placed in a new format for the convenience of the ultimate receiver. However, any such processing schemes would suffer the penalty of increased weight and complexity which would have to be weighed against their beneficial results.

The noise analysis of the demodulator-modulator repeater for FM signals shows a behavior quite similar to that of the frequency multiplier. That is, if the modulation indices are M_1 and M_2 for the ground-satellite and satellite-ground links, respectively, the postdetection S/N ratio is obtained by replacing M by M_2 in Equation (1) above, and the predetection

S/N ratio is given by Equation (3) with n replaced by $a = M_2/M_1$. The threshold requirements are the same as for the frequency-multiplier repeater.

In the modulation-converter repeater an arbitrary input modulation produces both amplitude and phase modulation on a local carrier injected at the satellite receiver. The local carrier is large compared to the incoming signal so both AM and PM indices are small. The amplitude modulation is subsequently removed by limiting in the receiver. Following the limiter, the low index PM signal is multiplied in frequency so as to increase both the deviation and the center frequency of the signal.

The phase variations are given by

$$\theta(t) = \frac{E_s(t)}{E_L} \sin\left[(\omega_1 - \omega_2)t + M(t)\right] + \frac{n(t)}{E_L} \tag{4}$$

where $E_s(t)$ and $M(t)$ represent any amplitude and phase modulation on the original signal with carrier frequency ω_1, E_L is the amplitude of the injected local carrier, ω_2 its frequency, and n(t) is a noise voltage. Clearly, both AM and PM now appear on a carrier at baseband frequency of $\omega_1 - \omega_2$. Therefore, to prevent spectrum folding this frequency must be greater than half the original information bandwidth. That is for an FM signal, for example, the baseband must be twice as wide as for the other repeaters and 3 db more power will be required to achieve threshold. Single-sideband modulation on the up link would avoid this penalty. However, in any case Equation (4) indicates that a second demodulator will be required to recover the original information after the phase variations have been detected.

All of the four repeater types discussed above will have some features in common. For example, almost always the RF input signal will be linearly translated to an intermediate frequency where the bulk of the amplification will be accomplished. The demodulator-modulator only will further reduce the signal to baseband. For the linear repeater it would be possible to construct an all-microwave repeater using traveling wave tubes, for example. At the present state of the art, the number of TWT's required to supply the large gain required (100 to 110 db) would make the system noncompetitive with IF amplification. It is possible to make a re-entrant TWT repeater using fewer TWT's by feeding the signal, shifted in frequency each time, back through the same TWT a number of times. More experience would be required with such a system before it would replace conventional amplification at intermediate frequencies.

4. NUMBER OF RADIO CHANNELS

a. Individual-Channel Repeater

In this extreme form of channelization, each baseband channel becomes an individual radio channel (Figure 4b). The distinct radio paths may extend from individual input to individual output antennas, or may be multiplexed to a single receiving and a single transmitting antenna. Radio channels which are completely separated electrically can, of course, be completely separated physically as well. This physical separation can be used to provide even further electrical separation, such as isolation between antennas. A specific example of such a system (although 60 voice channels constitute a single radio channel in this case) is discussed in Reference 2.

Because of the relatively modest transmitter power required for a radio channel handling but a single voice channel, a solid-state transmitter suggests itself as a possibility. Such transmitters have been built with power outputs of 2.5 watts at 2.2 gc and a few hundred milliwatts at X-band frequencies. However, because of the difficulty of constructing solid-state amplifiers with high power outputs, a frequency-multiplying repeater or modulation converter would probably be used. Such a repeater would normally be limited to transmission (but not reception) of FM signals. However, even without demodulation the channels could easily be restacked in frequency in the satellite so that they become "receiver-grouped."

Other advantages of this type of multiple-access satellite are reliability through redundancy and simplicity. That is, loss of a single radio channel affects only a single voice channel rather than many. In a system employing many trunk channels, the loss of one degrades the overall quality of service very little. The possibility of using low-voltage, and low-dissipation semiconductor elements throughout the radio channel should also contribute to reliability. No special measures need be taken to equalize either differential signal levels (through programming of the ground transmitter, for example) or differential doppler shifts since each channel originates from a single source. Also since no intermodulation can arise in the radio channel, the transmitter output power capability can be used efficiently.

Among the disadvantages of the individual-channel repeater, the most obvious is the vast numbers of components required for a communication system of reasonable capacity. To provide n two-way voice channels would require 2n radio channels. Even though the components can be simpler in each channel, the total weight of each channel of a 2n-channel system will not be 1/2n that of two channels each carrying n voice channels. It is fairly obvious that such a system could never be a really lightweight system except for a very small number of baseband channels. A very-low-capacity space communication system is uneconomical for a commercial system but might be justified for some military purposes.

However, the impracticality of the extreme form of individual channel repeater does not destroy its usefulness as a means of demonstrating pros and cons of dividing a given traffic load between more and more radio channels. A very practical problem of this technique is the creation of more and more local oscillator frequencies. Unless this multitude of frequencies is generated from a common low-frequency oscillator followed by a multiplier chain, the number of spurious tones and the possibility of interference in the radio channels will grow faster than the number of channels. This tendency can be somewhat alleviated by using the principle of physical separation to reduce the coupling between antennas. Also if a common local oscillator supply is employed, the reliability of the system will be impaired unless a standby supply is included.

b. Single-Channel Repeater

The single-channel repeater is the opposite extreme from the individual-channel repeater. In this configuration (Figure 4c) all of the satellite repeater traffic is passed through a single wideband repeater channel. Clearly, for a high-capacity system all of the components (especially at IF) must be broadband and the final power output stage must be high-powered. On the other hand only a single receiving and/or transmitting antenna is required or can profitably be employed. That is, no attempt is made to isolate the baseband channels or combinations of baseband channels; rather they are lumped together by the repeater.

Conceptually, the single-channel repeater is very simple. In principle almost any of the four types of repeaters discussed in Section B.3 could be used. Practically, however, the linear-translator would usually be chosen. If the demodulator-modulator repeater were used, the basic modulation groups of baseband channels would have to be separated before demodulation, which defeats the purpose of the single channel. And, if the frequency-multiplier or modulation-converter were used, an already wideband signal would be increased in bandwidth by deviation multiplication, thereby taxing the capabilities of the multipliers and subsequent components. Indeed, the frequency multiplier is probably not applicable here because of the difficulty of filtering cross products which will be generated in the multiplier when several signals are present simultaneously.

The advantages of the single-channel repeater include modulation flexibility (by virtue of the linear-translator approach), reduction in numbers of components with a probable saving in weight, and minimal spurious tone interference. The latter advantage derives directly from the need for only one local oscillator chain. The claim for a net reduction in weight because of fewer components is somewhat shaky without a weight analysis of each specific satellite of this type. The wider bandwidths and the use of a linear repeater, and higher power required in the transmitter will usually preclude the use of solid-state devices in the transmitter portion of the repeater at least. Instead the tendency will be toward lightweight, high-efficiency TWT's for the final transmitter stage. And, because of the linear transmitter and receiver preceding it, any form of modulation can be passed by the repeater.

On the negative side, however, the linear single-channel repeater is wasteful of transmitter power because the transmitter must be worked well below its saturated output power to avoid intermodulation between the multiplicity of the signals. Furthermore, this same difficulty is present in earlier stages of the repeater. Amplitude nonlinearities in the receiver can equally well result in intermodulation products and even suppression of the weaker signals. In fact, precisely because the signals from different ground stations are not separated in the repeater, their levels must be equalized at the input to the satellite repeater by programming the output power of the ground stations. Also, there is no possibility of compensating differential doppler shifts so guard-bands must be assigned between signals from different ground stations.

One further drawback to the single-channel repeater is a poorer reliability, given the same quality of individual components. There is no inherent redundancy as in the individual-channel repeater because all signals pass through a single set of serial components. Thus redundancy must be added in the form of a second standby repeater. In addition, since high-voltage, high-dissipation devices are required in the transmitter, the lifetime and reliability of the transmitter are more subject to impairment.

5. BASEBAND CHANNEL GROUPING

In Section B.2 we classified repeater configurations according to whether the baseband channels were grouped for the convenience of the transmitter or the receiver. In a transmitter-grouped scheme the transmitter groups together either FDM or TDM signals destined for different receivers. Each receiver must then receive the entire transmitted signal, demodulate it, select the channels destined for it, and discard the rest. In a receiver-grouped system, either by the transmitters or in the satellite repeater, all the signals destined for one receiver are grouped together in frequency or time.

Transmitter-grouping of channels obviously makes the design of the transmitting equipment simple. A stack of FDM or TDM channels can amplitude- or angle-modulate a single carrier. This single modulated carrier becomes the natural unit for the satellite repeater to accept in one radio channel. Since the signal emanates from a single source, there are no problems of level control and guard bands for doppler shift. If the satellite repeater is of any type but the demodulator-modulator, the grouped transmitter signal must be transmitted as originally constituted to all the receiving terminals for which messages are intended. Each receiving terminal must have an essentially separate receiver for each transmitter with which it desires communication. This scheme then demands a multiplicity (n(n - 1) for n ground stations) of receivers and a single transmitter (a total of n) at each site.

Receiver grouping of baseband channels must be accomplished either by appropriate channel assignments at each individual transmitter or by rearrangement of the channels within the satellite. If the latter course is followed, then the system could have both transmitter- and receiver-grouping. This would simplify the ground station transmitter and receiver design at the expense of complexity in the satellite repeater. Except perhaps for single sideband modulation, a separate modulated carrier is required for each group of voice channels belonging to a transmitter-receiver pair. A partially separate radio channel is also a necessity in order for the rearrangement to be carried out. The rearrangement could also be done at baseband in which case a demodulator-modulator repeater is indicated. Clearly the additional complexity of the satellite repeater will generally preclude channel rearrangement within the satellite.

At this point it should be noted that the terminology receiver- and transmitter-grouping has meaning only if channels having different destinations are modulated on a single carrier, or if a number of separate carriers are passed through a single radio channel in the repeater. In terms of the number of radio channels required in the satellite, both types of signals seem to demand something between the two extremes considered in Section B.4. That is, except in the trivial cases of a single baseband channel per RF carrier, or one-way transmission requirement from a single transmitter, the individual-channel and single-channel repeater configurations are inappropriate. In fact, as we mentioned, the logical number of radio channels for transmitter-grouped signals is just twice the number of transmitter-receiver pairs to be served simultaneously. Thus an M-access satellite with M large will increasingly take on the characteristics of the individual-channel repeater. Conversely, as M becomes small and traffic per access increases, the satellite repeater will approach the single-channel repeater.

6. APPLICABILITY OF COMMUNICATION SATELLITES FOR RANGE COMMUNICATIONS

While experimental communication satellite programs (Relay, Telstar, and Syncom) have produced very encouraging results, a successful demonstration program is a far cry from an operational system possessing the capabilities required to provide the desired range support. There are positive indications that communication satellite systems will be developed and will become operational prior to 1970. The classes of communication satellite systems which may be contemplated here are commercial, military, and special purpose. While a great deal of planning is under way for a commercial satellite system, as of this moment it is too early to tell precisely what the characteristics of this system will be and when it will be operational. The question of random versus stationary orbits has not been resolved. For the random system, it is questionable if the required number of satellites can be placed into orbit within a reasonable period of time, due to insufficient launch facilities. The establishment of overseas ground stations would appear to be progressing at a remarkable rate from a cursory examination of the information below.

Location	Antenna Size
United States	
Andover, Maine	Horn, 60 ft
Nutley, New Jersey	Dish, 40 ft
Mojave, California	Dish, 40 ft
Europe	
Goonhilly Downs, England	Dish, 85 ft
Pleumeur Bodou, France	Horn, 60 ft
Raisting, West Germany	Dish, 85 ft
Fucino, Italy	Dish, 30 ft
South America	
Rio de Janeiro	Dish, 30 ft
Asia	
Kashima, Japan	Dish, 30 meters
Kamisu, Japan	Dish, 20 meters

As encouraging as this large number of stations might appear, the installations are primarily for testing. Also, it is to be noted that Western Europe, which already has excellent microwave links, has most of the large antennas, which only complicates the multiple-access problem. The rest of the world, especially the continents which lack good ground communications, also lack large antenna facilities.

The question of frequency also is not completely resolved. All three of the early experimental satellites operate on different sets of frequencies, ranging from 1.7 to 8 gc.

As the plans for and capabilities of the commercial system become more apparent, steps can be taken to incorporate this system into the overall communications system for a global range. On the other hand, pending the acquisition of successful operational experience with the communication satellite system, there is a clear and present danger in requiring the use of such a system as an integral portion of the backbone range communications network.

Next we consider the case for use of military communication satellites. The primary functions of the military communication satellites will be to provide reliable communications for the purposes of the United States Government. These purposes include priority military and diplomatic communications and will include, whenever the capacity is available, routine military and diplomatic administrative and lower-priority communications. In view of these primary purposes, it is not apparent that the communications requirements of a global range would receive sufficient priority so that the availability of wideband communication satellite channels could be guaranteed. In addition, as is the case for the commercial communication satellites discussed above, it is at the present time impossible to predict reliably the date when a military communication satellite system with a wideband communication capability will become operational. The history of the ill-fated Advent communication satellite program provides ample evidence that it would be ill-advised to place heavy reliance on such plans as are presently available for an operational military communication satellite system. The medium-altitude system presently in the design study phase is a multiple-launch system which may provide the means for launching a sufficient number of spacecraft for random systems. However, the communication capability is extremely limited, consisting of a few voice channels.

The fundamental problem with the military communication satellite as an integral part of the communications system for a global test environment is that such a satellite communications system would not be under the control of the range or the range planner. Accordingly, its characteristics and availability could not be guaranteed. For these reasons, until such systems are operational and the question of their priority use for various purposes can be resolved, the military communication satellite system (or systems) should not occupy a central position in planning for a global range during the 1965 to 1970 period. Rather, as plans become more definite, as bandwidths increase, and as it becomes possible to resolve questions of availability of suitable channels, such systems may be factored into the overall communications system for the global test environment where their use appears to offer significant advantages, and where availability is assured.

Finally, we have the possibility of developing a communication satellite system specifically for use in satisfying the communications requirements of the global test environment. Such a satellite system would overcome the major arguments, cited above, against planning the test environment communications network around the availability and use of a communication satellite system. This possibility deserves serious consideration. However, it must be appreciated that, when due account is taken of the ground station requirements, in addition to the satellite requirements proper, the development and operation of a special-purpose satellite communications system for range communications would be an extremely expensive matter. For this reason alone, it appears unwise to require a satellite communications system for use in the range communications net if these communications requirements can be met adequately by more conventional means. In addition, it should be recognized that funding to support still another satellite communications system, in addition to the commercial satellite system and the military communication satellite system, will prove difficult to obtain, especially in view of the fact that commercial communication satellites capable of performing long-haul communications will inevitably become available in a matter of time. The basic point here is that development of a special-purpose satellite communications system to support range requirements would assuredly be an uneconomical solution to the range communications problem if the functions of such satellites could be performed by conventional means or by means of other communication satellite systems which may safely be anticipated within a reasonable period of time. This point is sufficiently obvious that the priority for a range communication satellite system would undoubtedly be very low, possibly even to the point where funding support and program approval would not be forthcoming in time to permit useful employment of a special-purpose communication satellite system within the period of interest.

The preceding arguments indicate the inadvisability of planning for the use of communication satellites as a primary vehicle of the communications network for a global test environment. As plans for such satellite programs develop and as programs progress, range planning activities can take due account of them in assigning a portion of the communications traffic to such systems in those cases where specific advantages will result. Such assignments will likely be, for a number of years to come, on a case-by-case basis, while the primary traffic load for the global range communications network will have to be carried by conventional communications systems. If this is true, it implies, among other things, that only in special cases will it prove feasible, prior to 1970, to require television transmission in real time from an orbiting vehicle throughout its orbit to a central monitor and control point. Rather, when real-time television is an essential requirement for the range support function, reception and display directly at the receiving sites must be employed. Or, if these sites are provided with adequate point-to-point broadband communications facilities at centers located in the vicinity or, at least, on the same continent.

C. HF RADIO LONG-DISTANCE COMMUNICATIONS

1. INTRODUCTION

The primary method for achieving long-distance or global communications has been through the use of high frequency band radio (3 to 30 mc). The most significant feature of this frequency range is that transmission path loss is very low, and thus receiving system gain requirements are modest. Because of this feature the HF band has attracted widespread utilization and as a result has become extremely congested; and interference between stations has become a major problem.

2. TECHNICAL CONSIDERATIONS

a. Propagation: MUF-LUF

Early users of HF radio observed that for any fixed distance of transmission at any given time, there was an upper limit of frequency which could be received over the path. This frequency is referred to as the MUF or maximum usable frequency. At frequencies above the MUF, the wave skips past the receiving point. This skip results because the wave is not reflected back toward the earth at a sufficiently steep angle to reach the receiving point. As the frequency increases significantly past the MUF, i.e., towards and into the VHF region, the wave may not be reflected back to the earth at all and communications to any point on the earth is then possible only by the scatter mode.

As the frequency is decreased below the MUF, the wave becomes increasingly attenuated by absorption from electrons in the ionosphere. In addition, multiple reflections are propagated as steeper incidence reflections become possible, resulting in delayed and interfering transmissions arriving at the receiver. Higher levels of atmospheric noise are also encountered at lower frequencies. The LUF, or lowest usable frequency, is then defined as the frequency below which the S/N ratio at the receiver is less than that required for satisfactory reception. Therefore, a frequency as close to the MUF as possible, leaving some margin of safety against changes in the MUF with time, is the best choice for transmission.

Unfortunately, due to changing ionization in the ionosphere with time of day, season of the year, and sunspot activity, MUF's and LUF's are constantly changing with time for a given HF path. Consequently, to maintain satisfactory communications over the path, the circuit must periodically change frequencies.*

b. Operational Limitations of High-Speed Data Systems

The natural or operational limitations imposed upon the user of high-speed data systems which operate at high frequencies and make use of ionospheric reflection may be divided into three broad categories:

Path losses
Interference
Multipath phenomena.

Control of path loss depends to a large extent upon proper system engineering, economic factors, and the required reliability, and to a (statistically) minor extent upon unpredictable variations of the ionosphere. Path losses, however, may be controlled satisfactorily by known techniques, so that adequate S/N ratios can be maintained at receiving terminals by the proper choice of operating frequency. Depending on the required reliability, control of path loss is not usually a limiting condition for the transmission of digital data at HF, particularly if operation is primarily outside of the auroral zones.

On the other hand, interference to HF transmission is a problem which can be controlled only by frequency-allocation action and proper system engineering so that adequate received fields are provided. At the time that such interference becomes intentional, additional system considerations are required. (A discussion of antijamming techniques is beyond the scope of this discussion.)

As opposed to path loss and interference, the theoretical limitation of signaling speed at HF has been due to multipath propagation phenomena, where signals from a given transmitter are received over

*A good example of this phenomenon is in the commercial broadcast band. At night the LUF is generally below this band so that radio stations can be heard great distances. During the day, the LUF is above the band and only local stations can be heard via the ground wave.

several possible paths, causing pulse smearing at the receiver. Such a situation is illustrated in simplified form in Figure 6, where the ionosphere supports transmissions from a minimum of four paths, the high and low of both a single- and two-hop path. The corresponding situation existing at the detector of a hypothetical receiver is illustrated in Figure 7, showing both wave cancellation and pulse stretching due to the arrival of multipath components. For successful digital operation at HF, means must be provided in the transmission system to circumvent these natural phenomena. As indicated, multipath accounts for variations in signal strength and consequently S/N ratio changes. Under low signal strength conditions, circuits are susceptible to radio interference as well as noise. Double-sideband amplitude modulation is singularly sensitive to selective fading* since carrier fading results in harmonic and intermodulation distortion. Fortunately, exalted or local carrier use has

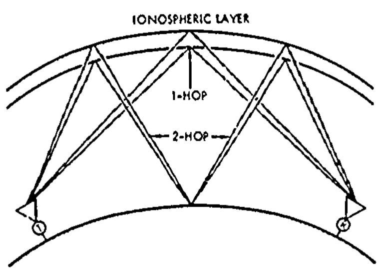


Figure 6. Multipath Transmission

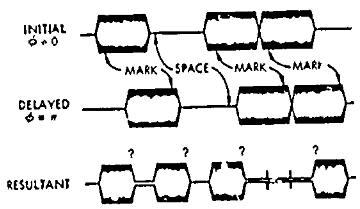


Figure 7. Mark/Space Uncertainty from Two-Component Multipath Delay Smearing at High Speed

*Selective and flat fading result from phase relationships between received signal components, as these add vectorially. Flat fading results when the components undergo little group delay, whereas selective fading arises from relatively long delays. For a two-path condition, frequency fading occurs with separations given by f = 1/t, where t is the group or envelope (AM) delay between the paths.

greatly alleviated this difficulty. Diversity reception, particularly space diversity, provides, together with single-sideband reception,* the means to arrest fading significantly. Thus, from the S/N ratio viewpoint, the performance of a HF radio circuit is primarily determined by the transmitted power and the absence of direct interference.

3. LONG-HAUL HF TRANSMISSION

The signal distortion mentioned earlier refers to pulse stretching or jitter; this effect is a direct result of differences in group delay between multipath components. Estimates of potential delays, based on a realistic model of the ionosphere have been derived and are presented in Figure 8. From the data presented therein, it is seen that delays in excess of 3 or 4 msec are possible as a result of multiple geometric mode propagation components. However, since these delayed components are greatly attenuated,** their signal disruptive effects are minor. To this effect Figure 9 illustrates the multipath delay spread as a function of distance and operating-to-MUF ratio. These curves were derived by NBS (Report No. 7206) from experimental data as well as geometrical calculations. On the basis of the information presented, it may be concluded that pulse lengthening (the cause of adjacent pulse interference) does not generally exceed 2 to 3 msec for appropriately designed and operated long-haul HF radio links.***

Automatic transmission of discrete information (data) over HF point-to-point radio links is generally accomplished on the basis of FSK, diversity reception (space or frequency), reliable propagation analysis, and operational procedures.

*SSB reception on the basis of detection with exalted or local carrier.

**Operation close to the MUF enhances this condition.

***This conclusion agrees with W. Lyons (Reference 3): "Under ordinary circumstances and availability of operating frequencies, we can say that the maximum multipath distortion we must contend with is approximately 2.5 msec. This figure establishes the minimum duration of a pulse to be 5.0 msec if we are to assume that we can identify a signal element which is 50 percent of normal length. Thus we can recognize that the highest practical speed for FSK working is roughly of the order of 200 bauds which is equivalent to 300 wpm when using the five-unit code."

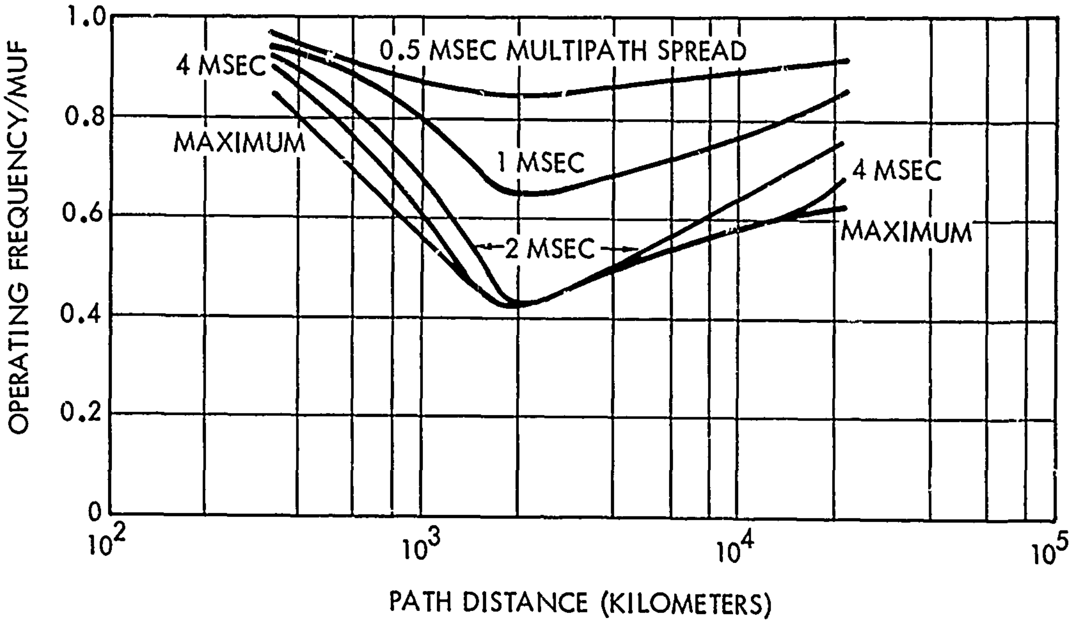


Figure 8. Multipath Delay Times

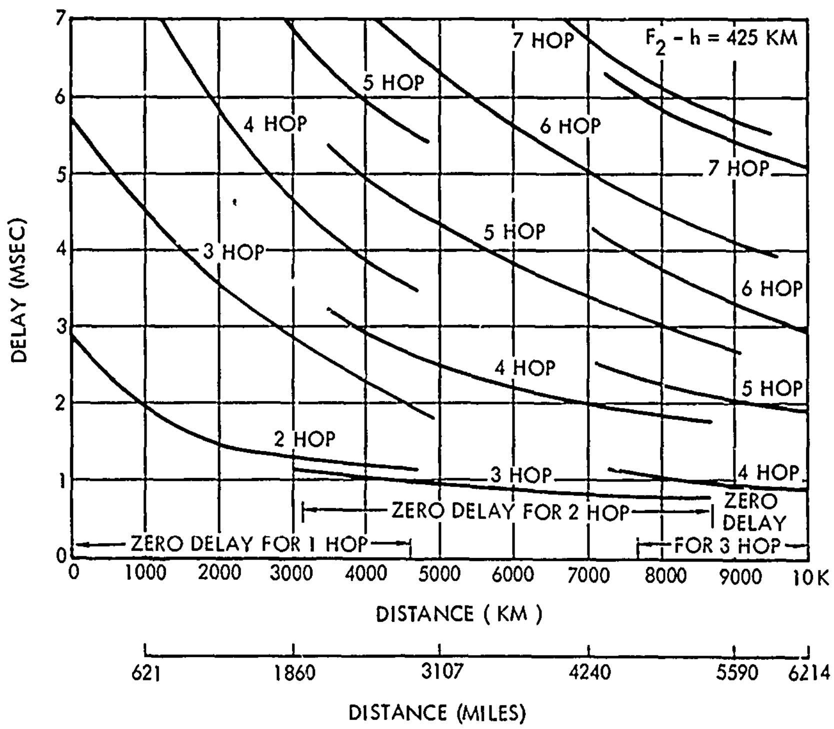


Figure 9. Multipath Delay Spread

The facts briefly summarized above explain past success of the 132-baud FSK transmissions referred to earlier (Reference 4). Similarly, good results have been reported by Jacobi (Reference 5) with respect to frequency division, subcarrier FM transmission at 120 bauds (8.3-msec signal elements). To quote Jacobi, "In field tests of a number of frequency division multichannel radiotelegraph systems for use in the high frequency range (3 to 30 mc), it was found that FM subcarriers, in conjunction with SSB and space diversity, provided the best method of transmission of those tried. Error rates of 0.02 to 0.14 percent were obtained on a transcontinental circuit."

A thorough statistical analysis of more recent experiments conducted by the Radio Corporation of America (Riverhead) under the sponsorship of the Department of Defense, provides further basis upon which to judge the performance of long-haul HF data systems. Figure 10 drawn from this analysis shows the percentage error rate as a function of signal element length (subcarrier frequency modulation) for a Hawaii to Riverhead, New York, radio circuit. It is seen, for example, that the error rate with a 5-msec pulse is less than five times that corresponding to a 10-msec signal element, 90 percent of the time. Admittedly, the error rate for both cases is much higher (two orders of magnitude) than the results reported by Jacobi; however, the 90-percent coverage and use of frequencies much below the MUF account for these differences.

4. SHORT HAUL HF TRANSMISSION

HF transmission over relatively short distances accentuates certain problems associated with this general mode of propagation. With reference to Figure 8, it may be seen that transmission delays associated with multipath components of ionospheric propagation increase as the range decreases (assuming an ionization density sufficient to support the mode). Thus, for 100-mile transmission, the second-hop F signal exhibits a delay of approximately 3 msec. Taking into consideration the possible ground wave component present at shorter distances, it may be concluded that fading and pulse stretching (or jitter) create serious transmission problems.

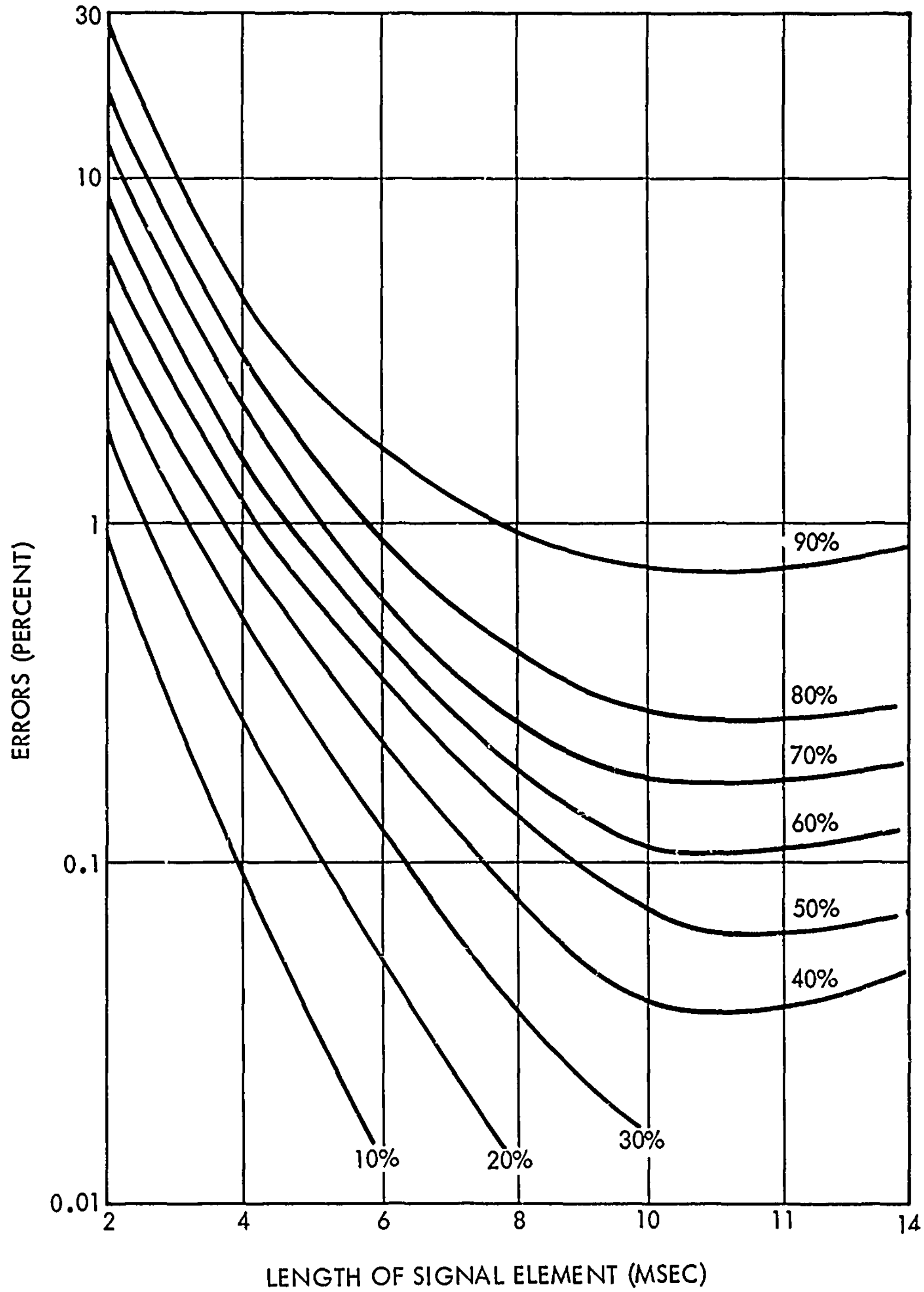


Figure 10. Error Rate as a Function of Signal Element Length

5. HF ANTENNAS AND TRANSMITTERS

Antenna gains are limited by both physical size and by the fact that radio waves arriving at two points even a short distance apart will vary in phase relationship and, hence, will not add properly in phase. Therefore, an electrical limitation of gain is reached in the vicinity of 20 db for ionospherically propagated signals. Antennas having gains of this value cover immense acreage and can only apply to ground stations.

The spatial decorrelation mentioned above is used to advantage in practice in space diversity receiving systems. Dual and triple diversity are commonly used as a means of selecting the optimum signal at any specific time. More advanced systems combine rather than merely select the best outputs.

Not to be confused with antenna gain for the land station is the concept of antenna pattern control. Although actual gain of the received signal is limited by noncoherence of the received wave, this does not hold true for the transmitted signal, nor does it limit the receiving antenna's ability to discriminate against unwanted signals arriving from other directions. This is most important, as the vast majority of long haul HF circuits employ rhombic antennas, which have very poor sidelobe suppression. In fact, most antennas in use today on HF circuits have poor sidelobe control to the detriment of their ability to discriminate against unwanted signals.

In order to increase the system gain at one end of the propagation link, a broadside array of vertical radiators with a Dolph-Tchebyscheff current distribution could be utilized to provide the higher directivity. The Dolph-Tchebyscheff distribution enables the selection of a desired sidelobe level and computation of the element excitation values in terms of a parameter related to the sidelobe level value. Sidelobe levels of the order of -30 db are attainable with arrays that are more than five wavelengths long. For this case, the half-power beamwidth in the horizontal plane is approximately equal to

$$\phi = \frac{60}{\frac{e}{\lambda}}, \text{ degrees}$$

where e is the length of the antenna array and λ is the transmitted wavelength. In a typical case,

$$\frac{e}{\lambda} = 5$$

so that

$$\phi = 12 \text{ degrees}$$

In the case of shipborne antennas, in which antenna size is severely restricted, gains are probably limited to 7 to 10 db. The antenna problem is further complicated by the necessity for rapid changes in frequency, thus virtually eliminating super gain antennas like the Yagi.

The transmitter power used is usually based more on the cost-power relationship of radio transmitters than on a requirement for a specific signal strength. That is, the control console, frequency synthesizer, and tuning mechanisms are to a large extent the same, whether the transmitter power is 100 or 100,000 watts. At the present state of the art, high-powered amplifiers in the 10 to 50-kw bracket are available from several manufacturers and represent a bargain in system gain. The fact that many services on the air simply cannot afford this kind of power and will move off the frequency, thus minimizing interference, may be a significant factor under some circumstances. Unintentional jamming will exist, therefore, particularly from the so-called emerging nations with their limited budgets.

6. MODULATION TECHNIQUES

Modulation techniques will be discussed in two categories; analog signal and digital signal techniques. Because of multipath, wideband FM is not practical on HF circuits. Therefore, most analog information is transmitted using amplitude modulation. Present practice is toward almost complete use of single sideband to conserve bandwidth in the crowded HF band. In SSB-AM the principal problems are noise, interference from other stations, and both flat and, to a minor degree, selective fading. Noise and interference are best attacked by methods which result in more signal power at the receiver. This means using the minimum bandwidth necessary in the receiver, properly oriented directional antennas, and adequate transmitter power. Fading is more difficult to combat but diversity reception can provide some relief. Selective fading generally limits the useful bandwidth of an analog signal to about 3 kc over long distances as this is approximately the bandwidth over which the correlation coefficient of the signal amplitudes is usually greater than 0.5.

Digital signals include TTY and data transmission. The principal problems here are fading and multipath, regardless of whether the system is on-off AM, FSK, or PSK. Significant improvement in error rates can be obtained by the use of diversity reception. However, multipath limits the data rate that can be used on any one channel due to intersymbol influence from delayed signal paths. Multipath delays usually lie in the range from 2 to 8 msec, which allows maximum data transmission rates of 100 to 500 bits per second. To obtain higher data rates, multiple channels are used and the composite frequency multiplexed signal (in a 3 kc bandwidth) is SSB modulated on a HF carrier. Thus, HF radio users today generally employ single-channel FSK or multiple channel FSK/SSB and differentially coherent PSK/SSB, all with space diversity reception. Error rates of about one in 10^3 or 10^4 bits during periods of good propagation conditions are to be expected with these systems. Improved error rates can be achieved with error detection and correction techniques at the price of reduced data rates.

Other techniques have been developed for overcoming the multipath limitations on data rate. Two such techniques (see References 6 and 7) will be mentioned here briefly. The first technique is a frequency-stepping scheme which consists of sending the short pulses needed for higher data rates on different frequencies in a sequence. In this way, the system is always operating on a frequency that is not contaminated by multipath from the preceding pulse or pulses. The frequency-stepping is continued until the multipath has subsided on the original frequency. Then the sequence of frequencies is repeated. Another more complex scheme is the RAKE system. Here the digital or mark-space sequence is coded into a wideband signal; e.g., 5-msec pulses after coding will occupy about 10-kc bandwidth. Portions of the signal possessing different delay times are isolated by correlation techniques and then added together after inserting the proper delays to make them coincide. The

RAKE system has demonstrated excellent results in the presence of strong multipath signals where ordinary systems would not be able to function. Development work under way indicates the desirability of using RAKE for high bit rate systems up to 10,000 bits per second. The main drawback to RAKE is its uneconomical use of the radio spectrum.

Figure 11 shows the general block diagram of the transmitter and receiver of a four-phase system. The transmitter contains little that is novel; it simply groups the incoming binary information stream into quaternary numbers which advance the phase of the transmitted signal by 0°, 90°, 180°, or 270° according to whether the quaternary number to be transmitted is a 00, 01, 11, or 10, respectively. The carrier frequency (1667 cps) and the timing are stable to 1 part in 10^5, being derived from a crystal oscillator at approximately 100 kc. Actually, four carrier signals with quadratic phase relationships are supplied to the balanced modulator, and the coding mechanism selects one of these four phases by means of the modulator. It is to be remembered that the phase selected is advanced from its preceding value of 0°, 90°, 180°, or 270° depending on whether the coding mechanism recognizes a 00, 01, 11, or 10 from the information source. That is, the modulation involves a running reference so that an absolute phase reference is not required.

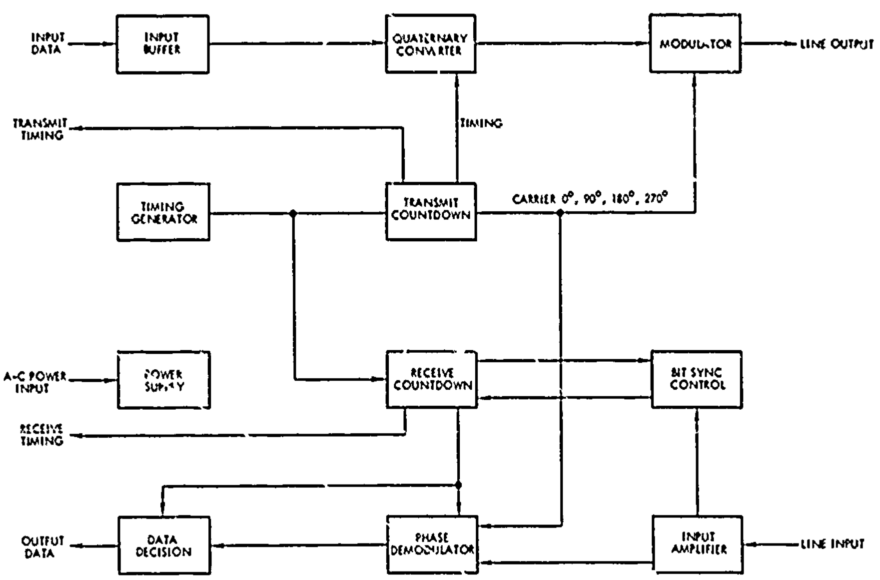


Figure 11. HC-270 Digital Data Transceiver, Block Diagram

The symbol-to-symbol reference is easily obtained in the two-stage binary counter. A 00 signal adds nothing to the state of this counter; a 01 adds a count to the 2^0 end of the counter; a 11 adds a count to the 2^1 counter, and a 10 adds a count both places. There is, of course, a carry connection between the 2^0 and 2^1 stage. Each of the four possible states of the counter selects a unique carrier phase to drive the balanced modulator.

The receiving system involves somewhat more complex procedures and new concepts. The detection process requires bit synchronization in order to cross-correlate each incoming "quaternit" with a pair of quadratic references. The memory and decision circuits also require bit synchronization. Bit synchronization is accomplished by slaving a digital AFC loop (driven from an internal clock) to the output of a narrowband filter tuned to the signaling frequency. This filter is fed the detection envelope product of the line signal. Since the transmitted signal is bandwidth-limited, it has an envelope containing bit-synchronization information. Provision is included also for external clock use to meet unusual stability requirements.

The detection scheme involves integrating the outputs of two phase detectors which have quadrature carrier references. Since the detection involves comparison of a quaternit with the preceding quaternit, the reference carrier need not be exactly the same frequency as the transmitter. Twelve-cps (or even 30-cps) frequency translation is permissable; this tolerance will be more than adequate safequard against carrier shifts.

The outputs of the phase detectors are integrated by means of RC filters which also act as bit storage elements. This last function is accomplished by a unique circuit designed to provide simplicity and significant insensitivity to internal variants. In short, the receiver decomposes the received signal into orthogonal components. The phase of the signal during each quaternit is detected and remembered in terms of orthogonal coordinates established by the carrier frequency references of the phase detectors.

The phase detection method centers about the orthogonal transformation matrix. This circuit generates a set of four vectors which are the original detected vector shifted by 0°, 90°, 180°, and 270° respectively. For example, 180° shift is accomplished by inverting both the I (inphase) and Q (quadrature) channels. As each quaternit is detected by the I and

Q channel phase detectors and integrators as a two dimensional vector, it is immediately converted into a set of four orthogonal vectors. The preceding quaternit is stored as a vector in the hold (or memory) circuit. Next the remembered vector must be compared with each of the vectors of the orthogonal transformation matrix, and it is evident that if the remembered vector is most like the incident vector (transformed by 0°) than 00 has been sent. If it is like the incident vector transformed by 180° phase shift, a 11 is detected. Similar arguments show that 01 and 10 are also uniquely determined by which output terminals hold the like signal. The essence of the detection is in this determination of the phase of the quaternit by noting which orthogonal transformation matrix terminals have the vector most like the vector in the hold circuit.

In order to use a binary threshold circuit, the signal from the subtracting network is fullwave-rectified so that it is always positive and approaches zero only when the two vectors are like each other. The threshold is placed between zero and the lowest value of unlike vectors.

7. EFFECT OF SOLAR ACTIVITY ON HF PROPAGATION COMMUNICATIONS

Since solar activity influences the earth's ionosphere to a considerable degree, it plays an important role in long distance communications via ionospheric radio propagation. Knowledge of the sun's influence on the ionosphere and predictions of solar activity are essential to forecasting propagation conditions for long distance radio communications between points throughout the globe. The following material presents a brief discussion of some of the effects of solar activity on the ionosphere and propagation. Finally, some predictions of the sun's activity for the next 11-year solar cycle will be given.

a. Solar Effects in the Ionosphere

Solar activity influences both the upper and lower useful frequencies for HF radio by control of the ionization density in the upper and lower regions of the ionosphere. The well known 11-year cycle in solar activity is quite evident in the degree of ionization of the upper atmosphere. While solar activity probably influences other factors in the ionosphere, e.g., mass motions and temperature, there is little doubt that large changes in the ionization density produced in the ionosphere occur with changes in solar activity. As an index of solar activity, the conventional Zurich sunspot number R is used. Table III illustrates the variation of the MUF and the FOT (frequency of optimum transmission) as a function of month, time of day, and sunspot number (SSN). The sunspot number has ranged from zero up to as high as 200 (latest solar cycle) over a solar cycle. The period of minimum solar activity is referred to as sunspot minimum and the period of maximum solar activity as sunspot maximum. The period between maxima covers a solar cycle and has been determined by observations over the last two hundred years to be about 11 years in length.

Generally, reliability is increased as the sunspot number or FOT is increased. Higher ionization results in higher MUF's with correspondingly lower attenuation and atmospheric noise levels.

The solar flare, a short duration bright emission from an active region on the sun, is the phenomenon apparently associated with solar effects on the ionosphere. The flare emits large amounts of both photon and corpuscular radiation, some of which reach the earth's upper atmosphere causing perturbation in this region. Few large flares occur during the two or three years of sunspot minimum, but during sunspot maximum years it is common to observe several large flares and many smaller flares each month. The solar flare has three important effects on the ionosphere for the radio communicator. These effects are (1) the short wave fadeout (SWF), (2) the polar cap absorption event (PCA), and (3) most important of all, ionospheric and magnetic storms.

The SWF is caused by greatly enhanced D region absorption due to increased electron densities in this region. Since the D region is the lowest layer of the ionosphere, radio waves travel through it before being reflected from the higher regions. As a result, radio waves fade out at the receiver from the much higher path losses occurring in the D region.

The SWF phenomenon occurs simultaneously with the observance of a solar flare and is short-lived. However, tens of minutes to several hours later, protons arrive from the larger flares and are channeled into the upper atmosphere of the high latitudes by the earth's magnetic field. Consequently, the D region absorption increases over the entire polar

Table III. Typical Variation of MUF and FOT with Month, Time of Day, and Sunspot Number for a Range of 2671 Nautical Miles*

	SSN 10				SSN 150			
	June (mc)		Dec (mc)		June (mc)		Dec (mc)	
(Greenwich Mean-time) GMT	MUF	FOT	MUF	FOT	MUF	FOT	MUF	FOT
1	17.9	15.2	10.6	9.0	20.8	17.7	28.5	24.3
2	18.7	15.9	10.1	8.6	21.4	18.2	23.9	20.9
3	17.8	15.1	9.7	8.3	20.8	17.7	19.7	16.8
4	15.2	12.9	9.3	7.9	19.3	16.4	16.8	14.2
5	12.5	10.6	9.7	8.2	17.6	15.0	15.3	13.0
6	10.4	8.8	10.3	8.7	16.0	13.6	14.4	12.3
7	9.0	7.7	10.7	9.1	14.5	12.4	14.0	11.9
8	8.6	7.3	11.4	9.7	13.6	11.6	13.8	11.8
9	8.7	7.4	11.6	9.9	13.0	11.0	13.2	11.2
10	8.9	7.6	11.5	9.7	12.5	10.6	12.7	10.8
11	10.0	8.5	10.7	9.1	13.1	11.1	12.2	10.3
12	12.5	10.6	9.3	7.9	14.9	12.7	11.1	9.4
13	14.6	12.4	9.6	8.2	16.6	14.1	12.4	10.5
14	15.6	13.3	13.6	11.5	17.0	14.6	20.2	17.2
15	16.1	14.3	18.4	15.6	17.1	16.5	31.3	26.6
16	16.1	15.5	20.6	17.5	18.4	18.4	38.5	32.7
17	16.5	16.5	21.1	17.9	19.9	19.9	40.4	34.3
18	17.2	17.2	21.2	18.0	20.7	20.7	39.5	33.6
19	17.0	17.0	20 9	17.8	20.5	20.5	38.1	32.4
20	16.4	16.4	20.9	17.8	19.8	19.8	37.3	31.7
21	16.2	15.5	20.7	17.6	18.6	18.3	37.2	31.6
22	16.0	14.3	19.4	16.5	18.5	16.6	36.9	31.4
23	16.1	13.0	16.7	14.2	18.7	15.9	36.0	30 6
24	16.8	14.3	13.0	11.1	19.5	16.6	33.0	28.1

*National Bureau of Standards Calculations. MUF's and FOT's shown are monthly medians of hourly medians.

areas and results in the phenomenon of the polar cap absorption event, interrupting communications in these areas.

The third effect, magnetic and ionospheric storms, results from low energy corpuscular radiation that arrives at the earth 20 to 40 hours or more after a flare. These particles are responsible for magnetic and ionospheric storms of 2 to 5 days or more and for auroras and other related effects. For long distance radio communications, the ionospheric storm is probably the most important consequence of the solar flare. The SWF is short-lived and, as such, can probably be tolerated; the PCA event is usually restricted to high latitudes, but the ionospheric storm is a world-wide phenomenon that can last days and in severe cases affect almost all ionosperic propagation.

Ionospheric storms occur in association with magnetic storms. During sunspot maximum they tend to be shorter (2 to 3 days) and more severe, and during sunspot minimum, longer (4 to 5 days) and less severe. The chief manifestation of the ionospheric storm is the lowering of the MUF and, hence, narrowing (or elimination) of the range of frequencies between the LUF and the MUF. In addition, under storm conditions the ionosphere frequently ceases to reflect as a smooth layer, returning instead, diffuse and scattered signals--a phenomenon called "spread F," after the F layer of the ionosphere.

The forecasting of ionospheric storms is an obvious need since communications can be rerouted or cable transmission arranged for important traffic during this period. During the sunspot maximum period, the majority of storms occur in association with solar flares as discussed above. However, during the lower portion of the solar activity cycle this relationship no longer works. In its place, a pattern of storms tending to repeat every 27 days or so begins to form and this tendency strengthens as the sunspot minimum period approaches.

b. Consideration of the Next Solar Cycle

During the winter of 1957-58, smoothed sunspot numbers (SSN) reached the highest values ever recorded and are now unquestionably declining. A minimum SSN of less than 10 clearly may be expected in 1964-65, because nearly all previous minima have reached this low figure. During this predicted imminent period of low sunspot activity, MUF's will be drastically reduced. Instead of exceeding 40 mc, the MUF's may rarely reach 20 mc and often below 10 mc. Grote Reber, writing on the history of radio astronomy (Reference 8) states "It is interesting to see how the mystifying peculiarities of short-wave communications of 1930 gradually have been resolved into an orderly whole. The solar activity minimum of the early thirties must have brought with it abnormally low-critical frequencies. Many a winter night was spent fishing for DX at 7 mc when nothing could be heard between midnight and dawn. It is now clear that the MUF over all of North America was well below 7 mc for several hours." Based on experience from past sunspot minima, the average traffic capability of the ionospheric spectrum will be less than half the capacity experienced during the last few years.

Successive sunspot maxima seem to build up gradually and then collapse. Dr. S. G. Lutz has pointed out that there may be a long-term sawtooth envelope to sunspot maxima as shown by the dashed curve in Figure 12 (Reference 9). The past peak, being the highest ever recorded, could be the last in this century to reach these values. HF radio communication over paths of approximately 250 miles will also be vulnerable to the influence of declining sunspot activity.

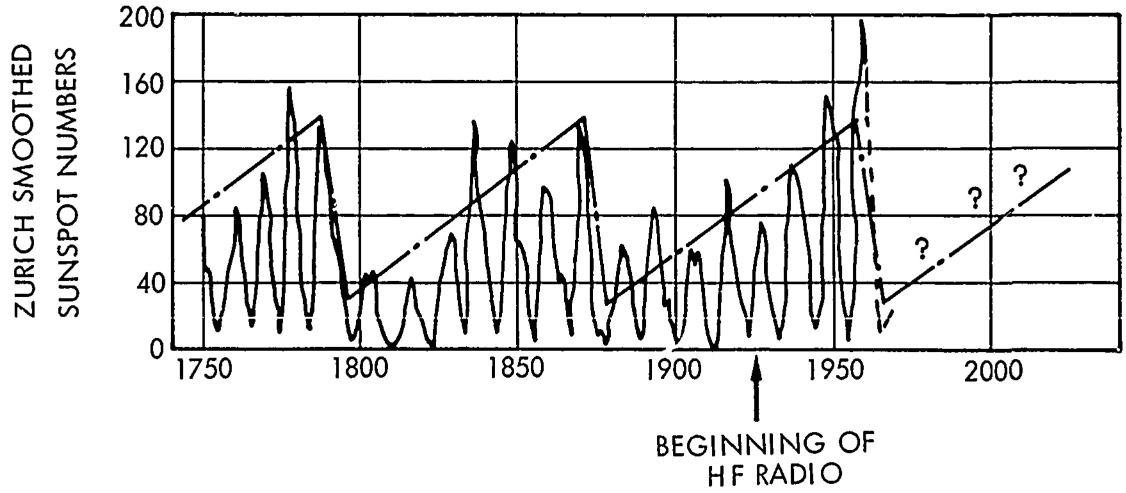


Figure 12. Sunspot Cycles Since 1749

Figure 13 shows predictions of sunspot activity during the period April 1962 through April 1965. These predictions were prepared by the Central Radio Research Laboratory of the National Bureau of Standards (Reference 10). As is apparent from this figure, the smoothed sunspot numbers are predictable with reasonable accuracy. On the basis of semi-theoretical and statistical analyses of previous sunspot numbers (References 11-12), it appears likely that the next cycle will exhibit much less activity than the last one. The sunspot minimum period will apparently occur in the 1964 - 1965 period and the sunspot maximum in the 1968 to 1969 period. The most probable maximum sunspot number, based on statistical analysis, will be in the 110 to 160 range.

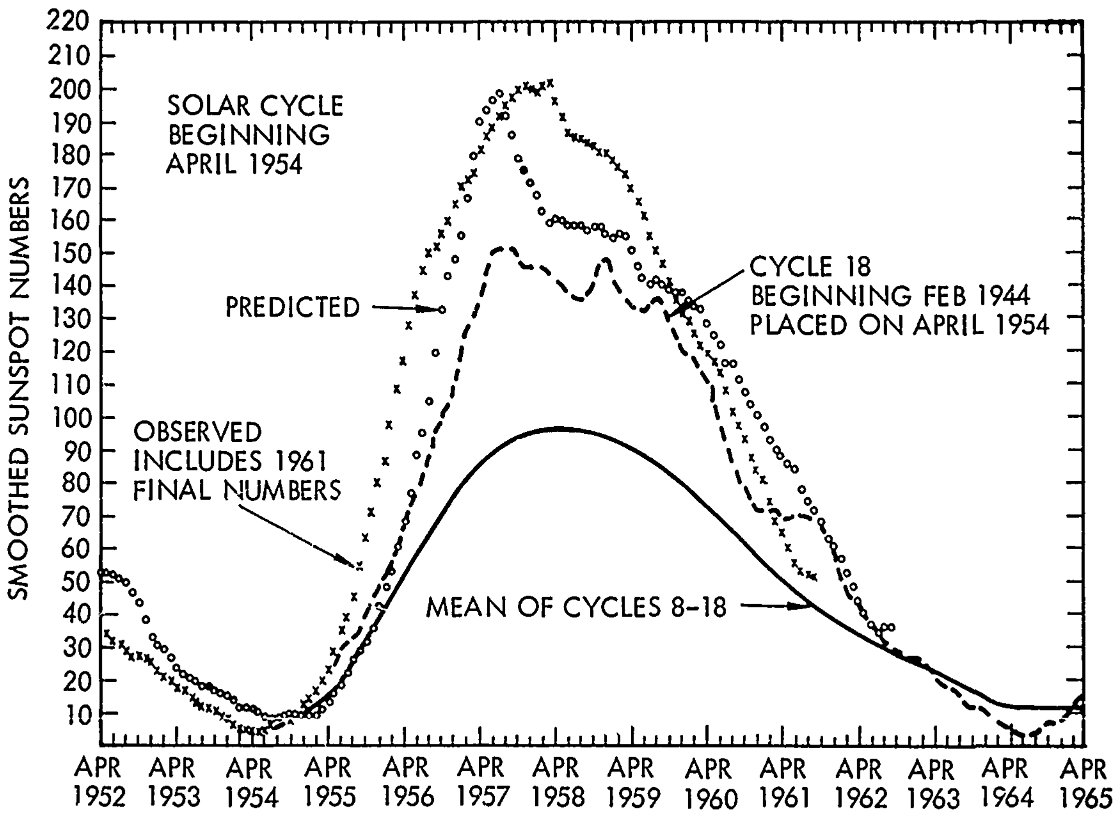


Figure 13. Predicted and Observed Sunspot Numbers

The Central Radio Propagation Laboratory predicts sunspot numbers three months in advance in its Ionospheric Predictions publication issued monthly. A 12-month moving average of sunspot numbers is used. In about 1965-66, a much more reliable estimate of the sunspot maximum number can be made.

8. HF COMMUNICATION SYSTEM PARAMETERS

Neglecting interference from other stations, transmitting power is determined primarily by the required reliability. Reliability, in turn, is determined by a number of factors, of which the most significant are range, atmospheric noise, location of receiver and transmitter, man-made noise, month, time of day, sunspot number, and transmitter power.

If one computes the power required to communicate between any two distant points on the earth's surface with normal conditions assuming no interference, the answer is in milliwatts. If conversely we assume the necessity to retain communications during a sudden ionospheric disturbance, the required power will exceed the present state of the art of high power transmitters and is in the order of tens of megawatts. Obviously, either of these extremes must be tempered with practicality. Transmitter power is one of the major factors in high frequency communications in that it is readily variable from zero to about 1 megawatt at present.

Choice of optimum power is, of course, a compromise. As a signal increase of 3 db requires twice the power, it is obvious that system gain achieved by power is costly indeed. Another important factor in output power is the corona or arc-over problem of the antenna system. With powers of 10 kw or greater, care must be taken such as insuring adequate radius on all conductors and large arc-over paths on insulators. This problem would further be complicated on ship board by the salt spray environment. At the present state of the art, the antenna problem is the limiting factor in transmitter power.

The choice of an optimum frequency is the greatest variable available to the HF system planner. That is, if a very astute choice of the maximum usable frequency can be made and retained, not only will the signal be of maximum strength, but contaminating factors such as multipath will be at a minimum. If the precise maximum usable frequency for 1-hop F layer were precisely followed, the following characteristics would be present:

1) No multipath
2) Interference from sources other than those within a few hundred miles of the desired source of signal would not exist, as there is no mechanism of propagation. At the MUF, only one hop can exist and signals a mere few hundred miles further or closer are not reflected.

An easy conclusion would be to recommend that one should always use the MUF. Unfortunately, as the MUF varies on an hourly, daily, monthly, and annual basis, accurate or precise following of the MUF is virtually impossible. It is, however, reasonable to follow the MUF within a few percent.

A relatively new method for overcoming the effects of constantly changing MUF's and LUF's is the use of "ionospheric sounders." The use of these sounders has been found to be an effective means for improving the reliability of long range HF communications, according to the results of a study by Granger Associates (Reference 13). The ionospheric sounder is a transmitter which sends out a series of pulses in a short burst which are spaced across the HF band. The signal is received at the receiver site and also at the transmitter by back scatter from the ionosphere. The signal can then be displayed on a storage-tube oscilloscope and from the display the MUF's and LUF's are readily discernible. The signal received at the transmitter site from a given back-scatter angle with the ionosphere will be identical to the signal received at a point on the earth along the identical forward scatter angle with the ionosphere. Therefore, the ionospheric sounder should indicate the same propagation conditions for the ionosphere at the midpoint of the path between transmitter and receiver at both sites. By periodically sounding the ionosphere, the HF circuit should be able to operate on the optimum frequencies at all times, barring the complications of frequency allocations and priorities.

Another possibility for determining the frequency of optimum transmission (FOT) is to employ a two-frequency technique. For over 25 years, long-distance HF point-to-point communications circuits have followed the MUF by transmitting the same information on two channels, separated in frequency by a few megacycles. If the error rate increases on, say, the higher frequency channel while holding on the lower channel, a shift to a lower frequency is effected. Conversely, should the lower channel depreciate, a shift to a higher channel is effected. This is a man-controlled action and works quite well. It would seem quite reasonable to consider developing an automated system to accomplish the same goal. Such a system would employ two additional carriers, placed on either side of the data carrier, separated by a megacycle or less and transmitting a series of approximately 10-msec pulses. A comparison would be made of the pulse train received on the lower carrier with that received on the upper carrier, amplitude comparison being one factor. Multipath would show up on the lower channel first, as a second or third pulse in the train. This would be an indication that the frequency is too low. Should the multipath signal arrive with greater amplitude than the main signal, we have an indication of back scatter. The possibilities of such a system appear to justify an investigation.

9. COSTS

HF radio is easily the most economical method for obtaining long distance communications. The main costs are the antenna installations, particularly with diversity reception, and transmitter power generation. As distances of 6000 to 7000 miles can be obtained with a single link, relay stations are not necessary.* Costs are thus reduced significantly. In addition, antenna installations are not nearly as costly as for scatter communication systems. Transmitter powers for HF installations typically run from 1 to 50 kw dependent on distances and grade of service desired.

10. CONCLUSIONS

Though HF communications suffer from many ill effects, such as interference, low usable bandwidth, and limited reliability, HF radio is

*Except in cases where relay around disrupted paths is desired to maintain communications during periods of abnormal ionospheric conditions such as storms.

by far the cheapest and in many cases the only method for obtaining long distance communications. For instance, for instrumentation ships at sea, HF radio would be the only means of communication, with the possible exception of satellite or aircraft relay.

The performance of HF circuits can be optimized by choosing the correct frequency (near the MUF) and a modulation method to suit the requirements of the data to be transmitted, such as SSB for voice, FSK for low-speed teletype, and complex techniques such as the RAKE system, four-phase modulation, or for the most difficult cases, quantized frequency modulation or possibly spread spectrum techniques. Some question remains as to the benefits derived from the more complex systems when due consideration is given to the maintenance problems posed. Clearly, any modulation-demodulation system must employ sufficient self-check capability to facilitate maintenance by personnel of limited experience. Error rates of 0.1 percent or less are completely feasible.

Principal problems to be overcome are:

An accurate solution to operation at or near the MUF at all times

Accurate prediction of MUF and FOT

Accurate prediction of ionospheric disturbances

The total problem of interference in the high frequency spectrum. This would include the eventual moving of many services now on HF to other parts of the radio spectrum

D. SUBMARINE CABLE FOR LONG-HAUL COMMUNICATIONS

1. INTRODUCTION

The investigation into global range instrumentation requirements has indicated a need for very-long-haul, point-to-point surface communication circuits of high quality and reliability. Based on reliability, performance, and relatively high capacity, submarine cables, especially of advanced design, offer interesting possibilities. Cable systems which have a bandwidth of approximately 400 kc for each direction of transmission will be available for service in the Caribbean and Northern Atlantic areas in 1963 and in the far flung Pacific areas (Japan, Philippines, etc.) in 1964 and 1965. The maximum single link (due chiefly to power restrictions) is approximately 4000 nautical miles. However, a 10,000 nautical mile section can be made up of several links composed of combinations of submarine cable, land lines, microwave links, etc. Even though this arrangement does not fulfill the prediction of some articles (2-mc cable system for a 10,000-nautical mile section) (References 14 and 15), it can be assumed that a submarine cable system capable of even wider bandwidths is not only feasible, but should be available early in the 1970 to 1975 era.

2. HISTORICAL BACKGROUND OF SUBMARINE CABLES

The first successful trans-Atlantic submarine cable was put in service in 1866 with the first messages sent by DC telegraph at a rate of two words a minute (representing a bandwidth of less than 1 cps).

The cable structure has progressed from stranded copper wires, gutta percha insulation, and soft armoring wire, through continual loading of the conductor and steel armoring wire, to a coaxial arrangement with low density polyethylene for electrical insulation and high-density polyethylene for mechanical protection (omitting the overall steel armor wire and resultant twisting while laying).

The first submerged repeaters were put in service in 1943 (see Table IV). Flexible repeaters, able to pass through the laying gear without being adversely affected by the twisting and untwisting of the cable as it was laid, were then used for the first deep-sea installations. This

Table IV. History of Submarine Cables

	Date	Length nmi	Use*	No. of Chan-nels	Sub-merged Re-peaters	Remarks
Newfoundland-British Isles	1866	1900	A	1	No	First successful deep-sea submarine cable
Trans-Atlantic			A	1	No	21 additional cables from 1869 to 1928
California-Hawaii	1902	2277	A	1	No	First cable across Pacific
Hawaii-Midway	1903	1336	A	1	No	
Midway-Guam	1903	3656	A	1	No	
Guam-Manila	1906	1647	A	1	No	California to Manila 7716 nmi
Key West-Havana	1921	104	B, C	5	No	First voice cable - 1 voice chan and 4 tele chan
California-Santa Catalina	1923	40	B, C		No	First use of carrier telephony
Wales-Isle of Man	1943		B, C	5	Yes	First use of submerged repeaters - shallow water
England-Germany	1946	196	B, C	5	Yes	Shallow water installation
Key West-Havana	1950	104	B, C	24	Yes	First use of submerged repeaters - deep water
Scotland-Newfoundland	1955	1940	B, C	36	Yes	First long haul use of submerged repeaters
Washington-Alaska	1956	550	B, C	36	Yes	Long haul use of submerged repeaters
Newfoundland-France	1959	1950	B, C	36	Yes	Long haul use of submerged repeaters
Florida-Puerto Rico	1960	950	B, C	48	Yes	3-kc channel bandwidth
New Jersey-Scotland	1963	3400	B, C	128**	Yes	Will be first long-haul single cable, deep sea
Hawaii-Guam	1964	4000	B, C	128**	Yes	
Guam-Tokyo	1964	1320	B, C	128**	Yes	
Guam-Okinawa	1965	1200	B, C	128**	Yes	
Guam-Philippines	1965	1150	B, C	128**	Yes	
Vancouver-Hawaii	1965-70	2400	B, C	60	Yes	British Commonwealth - cable compatible with U. S.
Hawaii-New Zealand	1965-70	4350	B, C	60	Yes	British Commonwealth - via Fanning and Fiji Isles
New Zealand-Australia	1965-70	1320	B, C	60	Yes	British Commonwealth

* A - DC Telegraph
B - VF Telegraph
C - Telephony

**Slow scan television possible

type of repeater only amplified in one direction; therefore, it took two cables to make the system. The art has progressed to where a single cable and repeater are now available.

The submarine cables listed in Table IV are just a small portion of the submarine cables installed throughout the world (Reference 16). The cable capacity has progressed from two words a minute on one channel of DC telegraph in 1866, through voice frequency telegraph and telephone capable of 400 words per minute by 1950, to a bandwidth of 400 kc in 1963 with extensive capability. The 400-kc cables are capable of 128 two-way channels of 3-kc bandwidth which combine into groups of 16 channels each (48 kc), and supergroups of five groups each (240 kc). This means that there can be combinations of telephone conversations, medium speed data (2400 bits per second), and/or teletype circuits simultaneously.

3. CABLE PROPERTIES

The armorless, deep-sea coaxial cable is approximately 1-1/4 inches in diameter and is distinctly superior to previous designs for underwater telephone transmission systems. The makeup is as follows: (1) a core of 41 wires of high carbon steel (for strength), (2) an inner copper conductor (surrounding the steel core) with a DC resistance of 1.64 ohms per nautical mile at 3 degrees C, (3) a low-density polyethylene layer over the inner copper conductor, (4) a return conductor of 10 mil copper tube around the polyethylene, and (5) a high-density black polytheylene outer cover approximately 1/8 inch thick. The cost of the cable is approximately $5,000 per mile.

The use of the steel spine allows a thicker insulation and a larger inner conductor with a subsequent reduction of transmission losses. This is accomplished without any increase in the diameter of the cable over the steel armored cable used previously. Also, the steel spine, in addition to adding strength, allows cable laying tensions without cable twisting (less than one-half turn every 100 feet). Extensive twisting of the cable would damage an armorless cable.

The cable properties are such that a plus or minus power plant of 5500 volts maximum can be used to provide power on the center conductor to the submerged repeaters. By using a double-ended, series-aiding power feed (positive 5500-volt supply at one end and negative 5500-volt supply at the opposite end), there can be a maximum route of 4000 nautical miles where the total voltage drop (repeaters and cable) is 11,000 volts.

4. REPEATER PROPERTIES

The newly developed armorless cable has made it possible to use a rigid repeater of new design (Reference 17). Formerly, when laying submarine cable, it was necessary to stop or slow the ship in order to place the rigid repeater. Through the use of a unique cable-repeater handling system, uninterrupted laying of cable with rigid repeaters is now possible. The major feature of the system is a linear cable engine used to pay out cable and the bulky rigid repeater in a straight line at a steady speed of approximately 800 ft/min. It is thought that this arrangement will be able to lay future cables with their repeaters without major modifications.

The rigid repeater (as contrasted to the previously used flexible repeater) can handle transmission in two directions and is designed to operate reliably for 20 years. It is shaped like a miniature torpedo, tapered at both ends. It weighs 500 pounds, is 50 inches long, and has a maximum diameter of 13 inches. The outer case is beryllium copper and is able to withstand a pressure of 11,000 psi, which is a value over that required for the anticipated deepest laying depth of 4000 fathoms.

The repeaters are made of broadband, fixed-gain amplifiers using specially designed electron tubes similar to the ones used with such success in the first trans-Atlantic repeater cable. Low mutual conductance electron tubes are used instead of high mutual conductance tubes or transistors because of the proven reliability, lower mechanical failure probability, and long life. Amplifiers are connected in parallel so that if one fails, the second one carries the load. Other features of the repeater include (1) an arrangement of a fusible alloy device which operates if a tube cathode heater opens, bypasses the heater, and inserts a compensating resistance, and (2) gas tubes connected across the input and output of the repeater which protect the components from high voltage surges.

A spacing of approximately 20 nautical miles will be used, the exact spacing depending on ocean temperature and pressure. Two bands of frequencies are used; a low band (116 to 512 kc) in one direction, and a high band (652 to 1052 kc) in the opposite. Directional filters in conjunction with power separation filters are used to separate the frequencies and direction of transmission. Prudent choice of multiplexing equipment allows 128 channels of 3-kc bandwidth. Interfering signals from unwanted transmission paths are at least 90 db below the power of the desired signals.

As a matter of interest, these repeaters will be used initially in a new submarine cable linking Florida and Jamaica (1963), and later Jamaica and Panama. They will also be used in a new trans-Atlantic cable scheduled for completion in 1963, and in Pacific cable routes to be completed in 1964 and 1965. It takes about 15 months to assemble and test a repeater, and it is expected that when full production is realized, 400 repeaters can be produced annually at an approximate cost of $60,000 per repeater.

5. CONCLUSIONS

Submarine cable offers a very fast, reliable means for the transmission of information which does not require great bandwidths. In other words, real-time video transmission (4 to 5 mc) is not possible with a single submarine cable system today. However, a frequency spectrum of approximately 400 kc is available for the simultaneous transmission of slow scan video (approximately 150-kc bandwidth), in addition to voice and/or teletype. This configuration would be possible by a skillful combination of the 48-kc groups. High-speed data (32 kilobits per second requiring approximately 50-kc bandwidth), in connection with other means of communications (voice, teletype, etc), would be possible in the same way.

Terminal arrangements are currently in operation for increasing the available number of channels on submarine cables. One arrangement (TASI), uses high speed switching for allocating channels during transmission of information. TASI provides a two to one increase in voice channels. It would be well to insure that TASI or similar equipment is not associated with circuits (channels) used for data transmission.

The submarine cable facilities are reasonably jam-proof, although not possessing a very great degree of immunity to spoofing without additional precautions. Secure transmission would not be available without additional terminal equipment. This additional equipment would tend to reduce the system reliability, which is one of the more attractive features of submarine cable systems. A degree of privacy of transmission is inherent in a submarine cable system.

Preliminary figures indicate that the 128-channel submarine cable system will cost approximately $10,000 per mile (installed), based on a 3000-nautical mile link. The pacing delivery item appears to be the repeater with a lead time of 15 months.

The progress which has been made since 1950 and the demand for wider bandwidths indicate that a submarine cable system capable of transmitting video should be available by the 1970 to 1975 era. Mr. Dingman, Executive Vice President of ATandT, has stated that ATandT expects to have a new transistorized submarine cable system available by 1966, capable of 720 voice channels or one TV channel. This means that when television is transmitted, there will be no space for voice and data and that alternate media would have to be found for the transmission of this information.

Such cables would be considered as "backbone" routes with spur or trunk lines connecting into this "backbone" at the most favorable locations. Any of several available communication modes could be used for these trunk lines (land lines, microwave, smaller submarine cables).

The impending availability of satellites for long haul communication does not lessen the importance of the submarine cable system. The requirement always exists for back-up and/or alternate routes, especially for man-in-space programs. There will be a continuing need for trunk lines to interconnect satellite ground stations with control and processing centers.

E. MICROWAVE RADIO RELAY SYSTEMS FOR LONG-HAUL COMMUNICATIONS

1. INTRODUCTION

Microwave radio relay systems have assumed an increasingly important role in ground communications throughout the world. About 27 percent of the long distance telephone circuit miles and more than 78 percent of the inter-city television circuit miles of the Bell System are provided by microwave radio relay systems. A trend toward relatively greater use of this type of radio transmission is expected to continue.

The line-of-sight characteristic of microwave transmission restricts somewhat its employment in global communication systems due to large areas of water to be traversed. The average spacing of 30 miles for repeater stations with associated equipment will have a political impact in countries which are unfriendly or where there may be a question as to future relationship.

The wide bandwidth available and relatively low transmitted power, along with a high degree of reliability, provide distinctive features which make microwave radio relay attractive for long-haul communications. These same qualities also are attractive for short-haul use, such as connecting an outlying instrumentation station into a long-haul "backbone" communication system.

The simplest as well as the most important fact about electromagnetic wave propagation in space is that a wave tends to travel along a straight line in the absence of obstructions. If the transmitted beam either wholly or partially hits the earth's surface or any physical obstacle, it will be reflected, or more correctly, it will be scattered. When a microwave signal has a "sky" component, it is little affected by the atmosphere in the troposphere and stratosphere regions but is strongly influenced by the ionosphere. For microwaves started toward the horizon, reflection by the ionosphere becomes possible.

The microwave region is normally thought of as being between the upper end of the radio waves and the lower end of the infrared waves, or in terms of frequency from 1000 to 100,000 mc (Reference 18). For the purpose of this report, the term microwave radio relay refers to the frequency region from approximately 1000 to 15,000 mc, using specialized directional antennas with narrow transmitted beams and line-of-sight tower locations. There are several good microwave radio relay systems such as the MM600 (2000 mc), TD-2 (4000 mc), TH (6000 mc), and TJ (11,000 mc) in service today.

2. CHARACTERISTICS

a. General

Due to the line-of-sight restriction, repeaters must be located on an average of 30 miles apart (depending on type of terrain, frequency used, and other factors). These repeaters compensate for the propagation loss. The repeater stations (as well as the terminal stations) usually have self-supporting towers so that the antenna systems can be placed above the surrounding terrain and oriented very carefully. With very narrow transmitted beams, and towers located on the highest promontory, interference due to obstacles and ground reflection is reduced. Some microwave systems use different antennas for transmitting and receiving with waveguide (circular or rectangular) leading from the antenna to the microwave transmitter and receiver. For repeaters, the equipment is usually mounted in huts built close to the tower. In some cases passive reflectors at the top of a guyed tower are used to deflect the microwaves to antennas located at the bottom of the tower. This simplifies the waveguide "run" to the transmitter and receiver and also cuts down on the cost of the installation, although more real estate is required since guy wires are needed for the towers.

The repeater stations as well as terminal stations require a prime source of power. In the continental United States commercial power is supplied and diesel alternators used for standby. In certain localities in the world the diesel alternator may have to be used as the prime source with a second diesel alternator as backup.

Microwave radio relay systems are able to achieve wide bandwidths with relatively low transmitted power. In the case of the TD-2 (Reference 19) system there is 1/2 watt of transmitter power output while in the MM600 (Reference 20) system the transmitter power output is 15 watts. To accomplish this, highly directional and very efficient antennas are used to transmit narrow microwave beams.

b. Transmission Range

The average repeater spacing is approximately 30 miles, and depending on type of microwave system used, systems can be built up to 4000 miles (References 19, 20, and 21). There are other systems good for 200 to 300 miles. The so-called long-haul systems are arranged so that groups of 4-kc channels can be dropped or added at certain repeater points making this type of communication system a good candidate for a "backbone" system.

The TD-2, TH (References 19 and 21), and MM600 microwave radio systems are representative of types of long-haul systems (3000 to 4000 miles). The TJ (Reference 22) microwave radio system is representative of a short-haul system (200 to 300 miles). However, due to the frequency of the TJ system (11,000 mc), raindrops and other forms of precipitation produce attenuation in the transmission path. The attenuation depends on rate of rainfall over the path rather than on total amounts, and when used in areas having periods of high rates of precipitation, a reduced repeater spacing and subsequent reduced system range must be accepted. The systems using the lower frequencies are not affected by this phenomenon.

c. Capacity and Reliability

The communication capacity of long-haul microwave radio relay systems is approximately 600 4-kc message channels or one television channel for each RF channel. These systems have 12 to 16 RF channels. However, all the RF channels are not available for separate information channels as frequency diversity techniques are utilized to provide automatic protection against either equipment failure or selective fading in the propagation medium. The use of one protection channel for each working channel gives the greatest reliability and best performance.

The short-haul systems normally do not have the capacity per RF channel as compared to the more expensive long-haul systems. However, the short haul TJ radio system has the capability of transmitting television a maximum distance of 100 miles.

The reliability of transmission when using the one-for-one frequency diversity exceeds 99 percent (Reference 20).

3. COSTS

There are definite economic reasons for selecting one form of communication system over another. These reasons may be based on the channel requirements, installation costs, maintenance and operating costs, and some intangibles. Construction costs differ from one area to another, and there is a wide variation between equipment costs, construction costs, and maintenance costs at overseas locations.

Microwave radio relay costs are proportional to the length of the circuit and somewhat dependent on accessibility to the sites. It is estimated that a 3000-milesystem would have 100 repeaters with two end terminals. Using this as a basis with 600 4-kc voice channels, a repeater installation costing $275, 000 each (including switching equipment), and a terminal installation costing $860, 000 (including multiplex equipment), the total would be $29, 220, 000. On a per-mile basis this would be $9740 (Reference 23), and on a per-channel per-mile this would be approximately $17. Annual costs which include amortization, operating personnel, spare parts, transportation, test equipment, utilities, etc., must be considered when it becomes a choice between two or more systems on a dollar basis.

4. CONCLUSIONS

The foregoing discussion has pointed out that a microwave radio relay system possesses many of the qualities required for long-haul communications where reliability and high quality are quite important. The microwave systems have wide bandwidth capability and can transmit video information in real time or great amounts of data (megabits per second).

Since it is a line-of-sight system it is not subject to the severe fading associated with ionospheric transmission. However, there is some fading as atmospheric conditions change and protection in the form of frequency diversity is used to offset this.

The microwave systems, with their advantages of wide bands of useful frequencies and high directivity and resolving power of the narrow microwave beams obtainable with antennas of physically convenient size, have some disadvantages: (1) The average spacing of 30 miles for repeaters places a severe restriction on localities where a system can be used. One of the big problems in a global network is the large bodies of water to be traversed. Unless land masses can be found within the line-of-sight characteristic, the use of microwave relay will have to be restricted to the different large areas of land such as continents, or for short-haul spur routes. Due to the number of stations required, the acquiring of real estate and the assurance of future friendly relations (in case of foreign countries), presents knotty problems. (2) Another disadvantage is that although attenuation of a microwave signal by the earth's atmosphere is usually small, it may be selectively large for certain bands of microwaves. For instance, the absorption by water vapor is selectively high for waves in the K-band (λ = 1 cm) and by oxygen is high for waves in the so-called K/2 band (λ = 1/2 cm). Therefore, it is important to consider weather conditions all along the route before choosing a frequency.

Although the useful frequency at the high end of the microwave range is approximately 15,000 mc, it can be expected that an increase in the spectrum of 5 to 10 times can be achieved by further improvement of power generation at the high end of the microwave range (Reference 18).

F. TRANS-HORIZON SCATTER COMMUNICATIONS

1. INTRODUCTION

A relative newcomer to long-distance communication techniques is forward scatter propagation of radio signals past the radio horizon. Scatter propagation is generally broken into two categories, "ionospheric scatter" and "tropospheric scatter," named after the region in the upper atmosphere where the scattering of the radio waves takes place. The two scatter techniques are generally different in their applications, reliability, operating frequencies, etc.

Ionospheric scatter communications can support multiplex teletype service or at most one voice channel between two points spaced approximately 600 to 1200 statute miles apart. Such circuits operate in the VHF region between 30 and 50 mc and achieve a performance or reliability of communications exceeding an HF skywave circuit but falling short of a cable circuit. Tropospheric scatter communications are applicable at shorter ranges, from 100 to 700 statute miles. Much greater communication capacities can be realized with tropospheric scatter circuits. Bandwidths up to 2 to 4 mc can be achieved at the shorter ranges with proper design of equipment. These circuits operate in the VHF, UHF, and SHF frequency bands and reliability of communications has approached that of cable, land-line, or microwave line-of-sight systems. In fact, tropo-scatter circuits have been used in tandem with cable, land-line, and microwave circuits. A prime example is the White Alice Network in Alaska.

Applications of both iono- and tropo-scatter circuits have been many in the past 4 or 5 years. Most of the applications have been military ones. Two examples are: (1) the Army Pacific scatter network consisting of seven major links (six of which use iono-scatter and one tropo-scatter) stretching from the Philippines to Palau (1095 mi) to Guam (829 mi) to Ponape (1017 mi) to Wake (1025 mi) to Midway (1183 mi) and to Hawaii (1180 mi); and (2) the White Alice Network which provides toll quality telephone circuits for both military and civilian service throughout the state of Alaska and consists of many tropo-scatter links. The reliability of both these systems has been quite high, exceeding 99 percent even for the iono-scatter links.

2. CHARACTERISTICS

In this section, a brief discussion is presented of those characteristics having a bearing on the design of VHF ionospheric scatter and UHF-SHF tropospheric scatter communication circuits. The equipment involved in typical scatter circuits is presented here. The transmitter site consists of a high-power transmitter (5 to 50 kw), operating at a fixed frequency, and a directional antenna (20 to 40 db gain), pointed near the horizon and in the direction of the receiving site. The receiving site consists of two directional antennas pointed at the scattering volume illuminated by the transmitting antenna, and two receivers which combine the outputs of the two antennas in some established manner. The two receiver-antenna systems are for space diversity operation and provide some measure of protection against the fading of the received signal which is characteristic of scatter communications. Since both receiving and transmitting antennas must point toward regions low in the sky, the antennas must be located above the general terrain in front of each antenna.

a. Transmission Range

Ionospheric scatter propagation is best used when the range is about 600 to 1200 miles. At this distance, the transmission loss varies only a few decibels if the antennas are pointed at the point in the sky appropriate for the path length. The small variation in loss is due to the fact that forward scattering becomes more efficient as the scattering angle becomes smaller or the antenna approaches the horizon for the longer paths. Beyond the 1200-mile region the losses increase rapidly as the curvature of the earth begins to mask the common scattering volume in the D region of the ionosphere (85 km high), where the scattering effect takes place. At ranges less than 600 miles, the larger scattering angles cause large path losses.

At ranges less than 600 miles, tropospheric scatter systems are more advantageous. The path lengths for tropo-scatter systems have ranged from about 100 miles up to as high as 800 miles with path losses increasing in a roughly linear fashion with distance. However, the total transmission loss for a tropo-scatter path is much less than for an iono-scatter path.

b. Operating Frequencies

The most useful frequency range for ionospheric scatter is from 30 to 60 mc. Several considerations affect the choice of operating frequency. Path loss increases with increasing frequency so that use of the lower part of the band minimizes power requirements. At these lower frequencies, however, long range interference effects become a problem due to propagation of strong signals to great distances by the F2 layer during maximum solar activity periods. Thus, it becomes necessary to plan carefully for frequency allocations at the lower end of the band, as for HF, or to change to frequencies above the F2 MUF with consequent increased power requirements or reduced channel capacity. In arctic latitudes, polar disturbances sometimes cause intense absorption at frequencies in the lower portion of the band. As a consequence, arctic circuits requiring the highest order of reliability must have the capability of operating above 45 mc. At temperate latitudes, absorption has not proven to be a serious factor down to 30 mc. In conclusion, therefore, the conflicting considerations to be weighed are the higher power requirements at the higher frequencies vs the interference and attendant frequency allocation problems and the arctic absorption effects of the lower frequencies.

Tropospheric scatter circuits operate in the VHF, UHF, and SHF bands. Many factors, some conflicting, enter into the choice of an operating frequency. These factors may be grouped into three areas: propagation characteristics, performance and costs of RF equipment, and circuit design specifications. In general, the path loss increases with increasing frequencies but increasing antenna gains at the higher frequencies tend to compensate for the increased losses. Available RF power outputs are smaller at the higher frequencies and therefore considerably influence the choice of frequency. The trend has been to favor the higher frequencies for short and intermediate length paths, especially where large bandwidths are required, and to favor the lower frequencies for relatively narrowband long distance paths.

c. Capacity and Reliability

The communications capacity of ionospheric scatter circuits is limited to perhaps a dozed 60-wpm multiplex teletype channels or a single voice channel. Reliability of communications with these capacities can equal or exceed 99 percent with sufficient transmitter power. For example at 50 mc, approximately 5 kw per 60-wpm channel is required for 99 percent reliability with an error rate of 1 in 10,000 characters. Error detection techniques can reduce this power somewhat. Intelligible voice transmission requires about 10 kw at 30 mc for 99 percent reliability, and with many repeater circuits between sending and receiving locations, power requirements could increase to 40 to 50 kw.

The communications capacity of tropospheric scatter circuits is much greater, due mainly to the much lower path losses. Here, capacity is a function of range. It appears feasible to obtain satisfactory transmission of standard black and white television signals over distances up to 200 miles. At the longer distances over 500 miles, 30 to 50 voice channels are possible. Transmission of digital information rates of several thousand bits per second with very low error rates is also possible. With 50-kw transmitter powers, communication reliability up to 99.9 percent has been obtained.

3. COSTS

With scatter communications the main costs lie in the large gain antenna systems and the large RF powers of the transmitters. For ionospheric scatter the costs are identical over any range from 700 to 1200 miles so that costs of the transmitting and receiving sites are the governing factor. The main variable on costs will be the communications capacity and quality of service desired. Other factors such as site acquisition and development will influence the costs to varying degrees depending on location. To extend the range beyond 1200 miles will increase costs since repeater stations would then be necessary.

For tropospheric scatter the costs are considerably dependent on tradeoff studies between distance, number of repeaters, and communications capacity. Unlike iono-scatter, for a given communications capacity and quality of service, the distance or length of hop will have a strong influence on costs since it will affect the total number of repeaters necessary.

4. SUMMARY AND CONCLUSIONS

In summary, the applications of trans-horizon scatter communications appear to be as follows. Ionospheric scatter is the most reliable radio technique available for fixed communications service in the range from 700 to 1200 miles. In addition, although the lower VHF band is crowded, some relief is offered from the congestion of the HF band. The major drawback in iono-scatter is the low communications capacities available, e. g., at most 12 to 15 teletype channel for a 50 to 60 kw transmitter. The reliability of communications or grade of service, while not as good as line of sight or cable systems, is still in the neighborhood of 99 percent. Tropospheric scatter offers a grade of service approaching or exceeding 99. 9 percent with, in addition, much greater capacity, e. g., up to 200 to 300 voice channels at the shorter ranges. The main drawback here, however, is the limited range for a single hop -- 100 to 700 miles. Consequently for long distance paths, land masses must be available for repeater sites. Of course, each additional repeater decreases the realibility of the overall communications path.

Improvements in grade of service and capacity for trans-horizon scatter communications lie in the development of advanced modulation-detection techniques and increased RF power generation. A significant improvement in capacity and range, however, is not likely since the propagation losses are an invariant product of nature.

G. VLF/LF COMMUNICATIONS

1. INTRODUCTION

Because of the reliability of low-frequency radio signals, they are extremely important to the military for long-range communications and worldwide standards of frequency and time. The very-low-frequency band (VLF) extends from 3 to 30 kc, while the low-frequency band (LF) extends from 30 to 300 kc. Radio communications began with the use of low frequencies, but when ionospheric reflection with HF waves was discovered, the VLF/LF bands were neglected until fairly recently.

Renewed interest in the low frequencies, especially the VLF portion, has occurred in the last few years because of the relatively low attenuation and stable phase characteristics of signals propagated in this band. Applications of low frequencies include communication, transmission of timing and frequency standards, navigational aids, and as an experimental tool for studies of the conductivity of the earth, characteristics of the lower ionosphere, etc. As will be shown, however, there are some large disadvantages in the use of low frequencies which greatly offset some of the advantages. As a result, their use is very limited when compared to the uses made of the higher frequency bands.

2. PROPAGATION CHARACTERISTICS

Radio waves follow three paths to reach the receiving antenna. The first of these is the direct wave which follows a straight line through space. Due to refraction from the atmosphere, the direct wave follows, to a certain degree, the earth's surface with a curvature of about four-thirds the earth's radius. The direct wave diminishes rapidly as the distance between transmitter and receiver falls into the shadow region and is insignificant for long distances.

The ground wave is propagated by the conducting and dielectric properties of the earth's surface. The attenuation of the ground wave is dependent on the conductivity of the ground between the transmitter and receiver and is also frequency-dependent in a complex manner with, in general, the losses increasing with rising frequency. At VLF, the ground wave diminishes inversely as the distance up to about 300 miles over sea water and ground of good conductivity; for ground of poor conductivity, the inverse distance holds true for only about 120 miles. At LF the same rules apply, but the distances are reduced to about 200 miles over sea water and ground of good conductivity and to only about 15 miles for ground of poor conductivity. At distances well beyond these, the ground wave diminishes at a much faster rate and becomes negligible.

The sky wave results from reflections by the ionosphere. At VLF the ionosphere is an excellent reflector because of the fact that penetration into it becomes a small fraction of the long wavelength. Because of convergence effects which become quite appreciable at great distances, the small ionospheric absorption and the inverse distance loss are counteracted to an important extent. This results in a sky wave which is strong enough to support very long distance VLF communications both day and night. The same general effects apply to LF at night also, but during the day the ionospheric absorption increases to a serious extent, particularly in the summer. The absorption rises rapidly with frequency.

To the long waves of VLF and LF signals, the ionosphere looks like a rather smooth reflector as contrasted to the roughness it exhibits to short radio waves. The smooth reflector causes the sky wave signal strength to be free of rapid fading for VLF and nearly free for LF. The fastest fades for LF are about one per minute. Slow changes in ionospheric absorption produce correspondingly slow changes in the sky wave signal strength. These changes usually occur when the sun is rising or setting on the ionosphere at the midpoint of the path.

Deep fades occur at points where the ground wave and the sky wave can substantially cancel each other. These cancellation points are not stationary but move as the ionospheric height and absorption changes. The region of cancellation is usually between about 200 and 400 miles. Much past this range, cancellation cannot occur since the ground wave is attenuated sufficiently to preclude it.

The same general discussion on fading or amplitude stability applies also to phase and frequency stability of VLF and LF signals. The smooth ionosphere and its slow variations produce no serious short-term perturbations in phase and frequency. Small diurnal variations occur due primarily to sunrise and sunset. However, the variations are similar from day to day. Standard frequency transmissions at VLF have been measured and precisions of several parts in 10^{10} over one hour have been attained at very long ranges, 3000 miles or more. By averaging over a 24-hour interval to minimize diurnal effects, trans-Atlantic frequency measurements of 2 parts in 10^{11} are achievable.

3. NOISE AT LOW FREQUENCIES

One of the serious drawbacks to the use of VLF and LF is the noise environment encountered at the low frequencies. Atmospheric noise is by far the major contributor and is that noise radiated from thunderstorms in the atmosphere. This radiated energy is rich in the low frequencies and the lower in frequency one goes, the stronger the noise field at any point. Due to the concentration of thunderstorms in the lower latitudes, atmospheric noise is higher in these regions. In the polar regions, noise is at a much lower level but still exists due to propagation of noise fields from the lower latitudes.

The atmospheric noise levels at any given point also have diurnal and seasonal variations. Atmospheric noise levels are lowest in the morning and begin rising rapidly in the late afternoon. This is caused by propagation factors and the rise and fall of thunderstorm activity. The noisiest period is generally at night. In addition, the noise is highest from March to September in the Northern Hemisphere and from September to March in the Southern Hemisphere because thunderstorm activity follows the summertime from hemisphere to hemisphere. Predictions of the atmospheric noise levels throughout the world in four-hour time blocks for the four seasons are published by the International Radio Consultative Committee (CCIR). In addition, the National Bureau of Standards operates a group of worldwide noise recording stations which attempts to up-date the CCIR predictions. Since atmospheric noise is nongaussian and tends to be very impulsive during high noise periods, amplitude probability distributions of the noise envelope are given.

Man-made noise also contributes to the overall noise environment and like atmospheric noise is strongest at the lowest frequencies. Sources of such noise are: ignition systems, corona from high-voltage power lines, commutation in motors and generators, electric railways, fluorescent lighting, etc.

4. DESIGN CONSIDERATIONS

a. Transmitter

As discussed above, the low radio frequencies suffer the least from attenuation and fading and therefore offer good reliability for long-range radio communications. However, because atmospheric and man-made noise have their strongest components in the low frequencies and are likewise propagated with the same ease as low-frequency signals, radio transmitters in the VLF and LF bands have to be very powerful to over-ride the noise at the receiver.

Also, because of the large wavelengths involved, it is not feasible to build transmitting antennas with directional characteristics. Thus, power cannot be beamed in preferred directions. In addition, the radiation efficiency of the antennas is often 25 percent or less and in the very best designed systems does not exceed 50 percent. All these factors combined require the transmitter and antennas to be capable of radiating large powers. VLF transmitters, for instance, are usually of 100 kilowatts or higher rating; and in the case of the Navy's VLF transmitter in Maine, the power rating of the transmitter is two megawatts.

b. Modulation

A major limiting factor to modulation methods is the extremely limited bandwidth of the transmitting antennas. This limited bandwidth is due to the high Q of the antenna which in turn results from its very short electrical length. Naturally narrow bandwidth channels cannot transmit high information rates. Thus, transmissions at VLF are limited to low information teletype or Morse code service. Some VLF antennas cannot accommodate 60-wpm teletype and therefore transmit 15-wpm code. LF antennas suffer less from high Q's and voice transmissions are practical in the upper part of the band.

Because VLF and LF transmissions suffer serious disturbances from atmospheric noise, it is not practical to amplitude modulate the radio carrier by small and varying amounts such as is done in commercial broadcasting. Throughout most of the world and during most of the time, the noise would obliterate the small variations in carrier level on long range circuits. The only AM that becomes practical is on-off keying of the carrier. Distinguishing between full on and zero carrier is possible except during periods of worst noise. On-off keying lends itself nicely to International Morse Code and coded signals for navigational aids. The usual keying rate is about 15 wpm for Morse code. On-off keying for teletype service has been generally unsuccessful because of the vulnerability to atmospheric noise.

Angle modulation is also used at VLF and LF even though higher order sidebands and consequently larger bandwidth than for AM result. Angle modulation has a distinct advantage over AM in that the system is relatively insensitive to atmospheric noise when operated above threshold. At present, most angle modulation is in the form of FSK. Single-channel 60-wpm start-stop FSK teletype is the most common form used at VLF. In the LF band where more bandwidth is available, time multiplex teletype is used extensively. A common system is one which carries four 60-wpm channels.

c. Receiver

In contrast to the large transmitting antennas required for VLF and LF, the receiving antennas can be quite small. The main requirement is that the antenna be sufficiently sensitive that the noise power it intercepts from space be at least 10 db above the noise of the antenna itself and the noise of the receiver front end.

In the receiver itself, which is of either the tuned radio frequency type or the superheterodyne type, a main design consideration is the noise reduction technique used in the receiver. Two well-known and worthwhile impulsive noise reduction techniques exist. The simpler of the two is known as clipping and the other as blanking or "hole punching." Both systems are employed in an early stage of the receiver where the bandwidth is largest. The impulsive atmospheric noise will be of high amplitude and short duration and can be easily clipped or blanked out, reducing the noise into the later stages of the receiver. In this way signal-to-noise improvements of about 10 to 20 db can be achieved, thus reducing the degrading effects of atmospheric noise on the signal.

5. COSTS

VLF and LF transmitter installations are very costly. The costs are due mainly to the tremendous power requirements mentioned above and to the huge transmitting antenna installations. These installations usually cost in the millions of dollars and generally consist of steel towers 500 to 1000 feet in height. At the base of the towers is an elaborate ground plane consisting of buried conducting cables radiating like the spokes of a wheel for hundreds of feet. The ground plane and the tall towers are necessary to achieve reasonable antenna efficiencies.

6. SUMMARY AND CONCLUSIONS

VLF and LF systems provide extremely reliable but low capacity communication service over long ranges. The main drawbacks to VLF and LF communications are the costly transmitter installations with their large antennas and high power requirements, the low usable bandwidth on the transmitting antenna, and the high noise environment from atmospheric and man-made sources. The main feature of VLF and LF communication is the extremely reliable signal available over long distances. Such signals are relatively free of fading, except for well-behaved diurnal effects, and are extremely stable in phase and frequency.

For the global test environment, the VLF and LF radio bands may be best utilized for transmission of range timing and frequency standards to take advantage of the stable signal. Reliable FSK teletype channels may also be used which will operate practically error free except during periods of very high noise level (perhaps 1 percent of the time or less, especially in the high latitudes where noise levels are lower). Considering the communications capacity obtained, and the cost of the transmitter and antenna installation, this is an extremely costly system from the standpoint of the amount of communication service bought. Therefore, except possibly for transmission of range timing and frequency standards which can be accomplished with much less transmitter power, VLF and LF radio does not appear to be a practical solution to global communications problems.

H. OPTICAL COMMUNICATIONS BY LASERS

1. INTRODUCTION

The most critical requirement in a global communications system is a reliable long-distance communication link of large capacity and reasonable power requirements. Laser techniques, at first glance, appear to be ideally suited to meet this requirement. Unfortunately, some inherent major drawbacks lead to the conclusion that lasers will not provide a practical solution to the communication requirements for the testing range in the 1965 to 1970 time period.

The first and most obvious drawback is the line-of-sight restriction for a light beam. This is similar to the restrictions on microwave systems in that point-to-point surface communications are limited by the earth's curvature and surrounding topography. The other major drawback is the instability of the atmosphere for propagation of laser frequencies. Weather conditions cause large variations in attenuation and refraction of laser beams traveling through the atmosphere and a loss of signal transmission results due to fading or beam bending, or both between transmitter and receiver antennas.

A technique that shows some promise is the use of so-called "light pipes" to provide a communication path for the information-carrying optical beam. The principal role of the pipe is to maintain a clear and homogeneous atmosphere for propagation of the laser beam from transmitter to receiver. The pipe must be composed of straight sections. joined by bends with reflecting mirrors. They must be evacuated or filled with a good transmitting medium such as dry nitrogen gas. Bends must be kept at a minimum to prevent large reflection losses. The installation and maintenance of such pipes may prove to be difficult since the pipes must be optically aligned to prevent losses from reflections off the walls. The use of a dielectric guide such as fiber optics for the propagation medium suggests itself as a means for replacing the optically aligned pipes but the attenuation of several decibels per meter is too high to permit communication over any practical distance.

The following text discusses some of the features of laser communications as compared to present day communications by microwave and radio frequencies. Some possible laser systems are discussed briefly and are summarized in Table V.

2. LASER FEATURES APPLICABLE TO COMMUNICATIONS

The fact that a coherent light beam (lasers) and a radio frequency or microwave beam are all electromagnetic waves, suggests applying radio principles to light beams. However, important differences exist in the properties of the two media, mainly due to the large difference between the frequency band of lasers and that of radio frequencies or microwaves. The following discussion points out the differences and their effects on communications.

a. Directivity or Beamwidth

One of the significant features of laser frequencies is the ease with which extremely narrow beamwidths can be generated. Attendant with these narrow beams are extreme directivity and high gain, resulting in very efficient transmission of energy over long distances. For example, a laser operating at one micron wavelength (near infrared) would require an aperture of four inches for a 10-microradian coherent beam with a resulting gain of 112 db. At X-band (10 gc), a 10 microradian beam would require a reflector diameter of 9800 feet with a surface accuracy to within 0.1 inch. At present, the largest gains that appear feasible at X-band frequencies are about 60 to 65 db.

b. Modulation Bandwidths

Another significant feature at laser frequencies is the very large increase in spectrum space or available bandwidth for transmission of information. To take full advantage of this large available bandwidth, the modulation of the light beam must be very wide band. Recent research (References 24 and 25) has led to modulation techniques where phase and intensity modulation of light beams by microwave energy has been accomplished. Bandwidths in the gigacycle region have been achieved with low modulation power. It therefore appears that a bandwidth as large as the entire radio spectrum today will be available on a single coherent light carrier.

Table V. Summary of Sample Laser Systems

Circuit	Configuration	Advantages	Problems	Conclusion
Long-Haul (trans-ocean)	Laser beams traveling in evacuated pipes analagous to sub-cable communications	Millions of times larger bandwidth than sub-cables Low radiated power levels Few repeaters necessary along path	Pipes must be pressurized or air-tight Pipes must be in optically straight sections Misalignment of pipes due to ground instability	If the laying of optical pipes is feasible and economically justifiable along ocean bottoms or buried in earth, then laser communications has definite advantages over sub-cables
	Comsat Relay Earth-Satellite-Earth	Low radiated power High gain beams Wide bandwidth	Alignment of beams due to extremely narrow bandwidths Absence of a good laser band window through the atmosphere under bad weather conditions (clouds, rain, fog, etc.)	Due to high propagation losses under bad weather conditions at all laser frequencies, Comsat application is extremely doubtful
Short-Haul (line of sight)	Laser frequency relay similar to TD-2 microwave links across U.S.	Low radiated power Wide bandwidth	Atmospheric losses for same reason as above Alignment of beams in link	Application is very doubtful due to high losses in bad weather. If sufficient laser power becomes available, short hops of 5-10 miles may prove feasible. Present TD-2 link across U.S. has 30 mile hops between points

c. Propagation

The atmosphere is remarkably transparent for electromagnetic waves in the visible, near-ultra violet, and far infrared regions of the spectrum. Unfortunately, however, bad weather conditions such as fog, rain, and clouds absorb and scatter the above bands, giving rise to very large attenuations. Even in clear weather, layers of air at different temperatures can curve the path of a light beam due to differences in refractive index, making it difficult to align the transmitter-receiver path and resulting in broadening of directivity. Thus, the atmosphere is not very reliable as a propagation medium for laser frequencies.

Even though clear water or sea water is partially transparent to certain frequencies such as blue-green light, the losses are still too large to permit the use of water as a propagation medium except for very short ranges. For instance, at the most optimum wavelength (0.47 micron), the attenuation in very clear sea water is on the order of 130 db per kilometer.

Therefore, the beam must travel in a well-chosen path such as indoors or within a pipe where the atmosphere can be controlled, or in a region where there is no atmosphere such as in space.

d. Noise Characteristics

The fundamental limitations on receiver sensitivity at laser frequencies and microwave frequencies are very different due to the different nature of the limiting noise in each band. In the microwave region, the minimum detectable signal is limited by the thermal noise level at the antenna, given by the well-known value kTB. In the laser band the thermal noise becomes negligible compared with the quantum or self-noise due to the particle nature of light. The minimum level now, instead of being equal to kTB, is equal to hFB where h is Planck's constant and F is the frequency.

At a wavelength of one micron, this minimum noise level equals 2 x 10^{-15} watt per cycle per second, while at microwave frequencies it equals 4 x 10^{-21} watt per cycle per second for a temperature of 300°K. Of course, quantum noise also exists at microwave frequencies, but it is negligible compared to thermal noise (e.g., at 10 gc the quantum noise level is more than 30 db below the thermal noise level at $300^{\circ}K$). Thermal noise also exists to a significant degree at laser frequencies if the thermal temperature of the noise source is high enough. For example, the sun or any incandescent body radiates significant energy at laser frequencies. To take full advantage of the minimum quantum noise level, a laser receiver must be protected as much as possible against extraneous light from such objects. This can be accomplished by either providing filters with bandwidths of only a few Angstrom units or by use of the pipes mentioned above where the pipes exclude any outside light sources.

From the numbers above one can see that the quantum noise limit is significantly above the thermal noise limit (23 db above), so that much more power is necessary to detect light signals. This disadvantage is more than overcome, however, by the high gain possible with the extremely directive light beams.

3. POSSIBLE OPTICAL COMMUNICATION CIRCUITS

The requirements for global communications fall into two broad categories: long-haul circuits such as might span the Pacific or Atlantic, and trunk line or short-haul circuits that would tie remote sites into the main long-haul circuits. The use of laser communications for both of these requirements will be briefly investigated here.

a. Long-Haul Circuits

A long-haul circuit could conceivably take the form of a beam of light traveling down a pipe made up of straight sections joined by bends where necessary. Each bend would have to contain a reflecting mirror to pass the light around the bend. The entire length of pipe would have to be either filled under pressure with a transparent medium such as dry nitrogen gas, or evacuated and sealed from the atmosphere. In addition, the number of bends would have to be kept at a minimum to cut down the reflection losses. When necessary, a repeater may be used analogous to sub-cable repeater amplifiers. Since the pipe must be optically aligned , keep the beam in the center of the pipe, the medium in which the pipe is placed, e.g. the ground or ocean floor, must be very stable. Since earth tremors and ocean currents are common, this alignment could prove to be a serious problem.

Consider a laser circuit with a 10-cmdiameter pipe 4000-km long. Assume a carrier of one micron wavelength (3×10^{14} cps) modulated with a 10,000-mc bandwidth signal. Previously it was stated that a 4-inch (about 10-cm) aperture could generate a 10-microradian beam with 112-db gain at 1-micron wavelength. The received signal power at the other end of pipe (neglecting reflection losses) is given by

$$P_o = \frac{P_t G_t D_r^2}{16 R^2}$$

where

P_t = transmitter power

G_t = transmitting gain (112 db)

D_r = receiving aperture diameter (10 cm)

R = distance (4000 km)

The noise power at the receiver is given by the quantum noise divided by the quantum efficiency of the detector.

$$N_o = \frac{hFB}{q}$$

where

h = 6.6×10^{-34} joule-seconds

F = 3×10^{14} cps

B = 10,000 mc

q = detector efficiency (realistic value 0.001)

Using these numbers and a desired signal-to-noiseratio of 20 db, the transmitted power necessary is 20 watts. This power has not been attained yet by a CW gas laser. However, if this power can be achieved, 10^{10} bits per second could be transmitted on this circuit at a virtually negligible error rate. This represents a fantastic improvement over present sub-cable networks. The use of a dielectric guide, such as optical fibers, appears as a possibility for use in place of the evacuated or pressurized pipes. However, due to the extremely high losses in fibers (up to several decibels per meter), communication over any practical distance becomes impossible.

The use of communication satellite systems (COMSAT) for long-haul is now being given serious attention. Use of laser circuits for the links from ground to satellite has been considered but this application is virtually excluded by the extremely poor propagation through the atmosphere in bad weather or clouds, as stated earlier.

b. Short-Haul Circuits

Again, the use of pipe circuits for laser short-haul circuits appears as a possibility here. Due to the short distances involved, short-haul may be the most practical use of "piped light beams." The optical alignment and reflection losses would be a lesser problem. Nevertheless, only if extremely large bandwidths are required does the use of lasers appear desirable over present communication techniques.

The use of short-haul "piped light beams" in tandem links analogous to the microwave links across the continental United States appears to be definitely feasible. Repeaters quite possibly can be built into the pipe, thus resulting in a laser version of submarine cable communications with millions of times more bandwidth or capacity.

REFERENCES

1. J. A. Develet, Jr., "Selected Topics on Modulation Systems for an Active Wideband Communication Satellite," STL, 8949-0006-NU-000, 23 April 1961.

2. D. P. Sullivan, "SSB-FM Communication System Design," STL, 8616-6008-RU-000, 30 March 1963.

3. W. Lyons, "Design Consideration for FSK Circuits," 1954 IRE Convention Record.

4. H. O. Peterson, J. B. Atwood, H. F. Goldstine, G. E. Hansell, and R. E. Schock, "Observations and Comparisons on Radio Telegraph Signalling by Frequency Shift and On-Off Keying," RCA Review, March 1946.

5. T. E. Jacobi, "Multichannel Radio Telegraph System for HF Circuits," RCA Review, December 1948, pp. 704-720.

6. R. Price and P. E. Green, Jr., "A Communication Technique for Multipath Channels," Proc. IRE, Vol 46, March 1958, pp. 555-570.

7. A. R. Schmidt, "A Frequency Stepping Scheme for Overcoming the Disastrous Effects of Multipath Distortion on HF FSK Communication Circuits," 1959 IRE National Convention Record, Part 8, pp. 176-186.

8. Grote Reber, Proc. IRE, January 1958.

9. S. G. Lutz, "An Eventual Communications System," IRE Transactions of the National Symposium on Space Electronics and Telemetry, 1960.

10. "Predicted and Observed Sunspot Numbers," CRPL-F212, Part B, Solar-Geophysical Data, April 1962.

11. K. A. Anderson, "Preliminary Aspects of Prediction Aspects of Solar Cosmic Ray Event," NASA Technical Note TN D-700, April 1961.

12. A. I. Ol, "The Possible Periodicity of Solar Activity," translated from Astronomicheskii Zhurnal, Vol 37 No. 2, March-April 1960, pp. 222-226.

13. "Ionospheric Sounder Tested," Electronic News, 29 April 1963, p. 22.

14. C. A. Parry, "Planning Aspects of a Globe-Circling Communication System of High Capacity, Reliability, and Performance Quality," IRE Transactions on Communication Systems, September 1961.

15. E. I. Green, "Progress and Problems in World Communications," Electrical Engineering, Vol 79, June 1960, pp. 494-498.

16. "List of Cables in World's Submarine Cable Network," Secretariat General De L'Union Internationale Des Telecommunications, 15th Edition.

17. "New Two-Way Repeater Amplifies 128 Telephone Conversations," Bell Laboratories Record, September 1962.

18. Encyclopedia of Science and Technology, McGraw Hill. [inc]

19. "6000 Mc Radio Relay System for Broadband Long Haul Service in the Bell System," AIEE, Vol 76, 1957.

20. "A New High Capacity Microwave Relay System," AIEE, Vol 78, 1959.

21. "TH Radio System," Bell System Technical Journal, Vol XL, November 1961.

22. "The TJ Radio System," Bell Laboratories Record, April 1959.

23. "Economic Analysis of Communication Systems," 7th National Communication Symposium, 1961.

24. C. J. Peters, "Gigacycle Bandwidth Coherent Light Traveling-Wave Phase Modulator," Proc. IEEE, Vol 51 No. 1, January 1963.

25. W. W. Rigrod, and I. P. Kaminow, "Wide-Band Microwave Light Modulation," Proc. IEEE, Vol 51 No. 1, January 1963.

BIBLIOGRAPHY

K. A. Anderson, "Preliminary Aspects of Prediction Aspects of Solar Cosmic Ray Event," NASA Technical Note TN D-700, April 1961.

H. Bosc, "Principal Uses of Coherent Light," Electrical Communication, Vol 37 No. 4, 1962.

"Communications by Piped Light," Mechanical Engineering, Vol 82 No. 12, December 1960.

G. C. Dacey, "Optical Masers in Science and Technology," Science, Vol 135 No. 3498, 12 January 1962.

J. A. Develet, Jr., " Selected Topics on Modulation Systems for an Active Wideband Communication Satellite," STL, 8949-0006-NU-000, 23 April 1961.

A. L. Durkee, D. Metcalfe, W. H. Tidd, "White Alice System-Design and Performance," IRE Trans. on Communication Systems, CS-7, December 1959.

Economic Analysis of Communication Systems, 7th National Communication Symposium, 1961.

Encyclopedia of Science and Technology, McGraw Hill.

E. I. Green, "Progress and Problems in World Communications," Electrical Engineering, Vol 79, June 1960, pp. 494-498.

"Ionospheric Sounder Tested," Electronic News, 29 April 1963, p. 22.

T. E. Jacobi, "Multichannel Radio Telegraph System for HF Circuits," RCA Review, December 1948, pp. 704-720.

G. Jacobs, "Optical Communications," Proceedings of National Electronics Conference, Vol 17, 1961.

"List of Cables in World's Submarine Cable Network," Secretariat General De L'Union Internationale Des Telecommunications, 15th Edition.

D. G. C. Luck, "Some Factors Affecting Applicability of Optical Band Radio to Communication," RCA Review, Vol 22 No. 3, September 1961.

S. G. Lutz, "An Eventual Communications System," IRE Transactions of the National Symposium on Space Electronics and Telemetry, 1960.

W. Lyons, "Design Consideration for FSK Circuits," 1954 IRE Convention Record.

"A New High Capacity Microwave Relay System," AIEE, Vol 78, 1959.

"New Two-Way Repeater Amplifies 128 Telephone Conversations," Bell Laboratories Record, September 1962.

A. I. Ol, "The Possible Periodicity of Solar Activity," translated from Astronomicheskii Zhurnal, Vol 37 No. 2, March-April 1960, pp. 222-226.

C. A. Parry, "Planning Aspects of a Globe-Circling Communication System of High Capacity, Reliability, and Performance Quality," IRE Transactions on Communication Systems, September 1961.

C. J. Peters, "Gigacycle Bandwidth Coherent Light Traveling-Wave Phase Modulator," Proc. IEEE, Vol 51 No. 1, January 1963.

H. O. Peterson, J. B. Atwood, H. F. Goldstine, G. E. Hansell, and R. E. Schock, "Observations and Comparisons on Radio Telegraph Signalling by Frequency Shift and On-Off Keying," RCA Review, March 1946.

"Predicted and Observed Sunspot Numbers," CRPL-F212, Part B, Solar-Geophysical Data, April 1962.

R. Price and P. E. Green, Jr., "A Communication Technique for Multipath Channels," Proc. IRE, Vol 46, March 1958, pp. 555-570.

"Radio Transmission by Ionospheric and Tropospheric Scatter," Proc. IRE, Vol 48 No. 1, January 1960.

A. R. Schmidt, "A Frequency Stepping Scheme for Overcoming the Disastrous Effects of Multipath Distortion on HF FSK Communication Circuits," 1959 IRE National Convention Record, Part 8, pp. 176-186.

"6000 Mc Radio Relay System for Broadband Long Haul Service in the Bell System," AIEE, Vol 76, 1957.

Grote Reber, Proc. IRE, January 1958.

W. W. Rigrod, and I. P. Kaminow, "Wide-Band Microwave Light Modulation," Proc. IEEE, Vol 51 No. 1, January 1963.

D. P. Sullivan, "SSB-FM Communication System Design," STL, 8616-6008-RU-000, 30 March 1963.

"TH Radio System," Bell System Technical Journal, Vol XL, November 1961.

"The TJ Radio System," Bell Laboratories Record, April 1959.

J. R. Wait, "Diffraction Theory of LF Sky-Wave Propagation," Journal of Geophysical Research, Vol 66 No. 6, June 1961.

J. R. Wait and H. H. Howe, "The Waveguide Mode Theory of VLH Ionospheric Propagation," Proc. IRE, Vol 45 No. 1, January 1957.

A. D. Watt and E. L. Maxwell, "Characteristics of Atmospheric Noise from 1 to 100 kc," Proc. IRE, Vol 45 No. 6, June 1957.