FINAL REPORT

SOLAR-PUMPED LASER

CONTRACT NAS 9-3671

Prepared for

INSTRUMENTATION AND ELECTRONICS SYSTEMS DIVISION
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
MANNED SPACECRAFT CENTER
HOUSTON, TEXAS

Prepared by

DEFENSE ELECTRONIC PRODUCTS
RADIO CORPORATION OF AMERICA
CAMDEN, NEW JERSEY

NOVEMBER 1965

FINAL REPORT

SOLAR-PUMPED LASER

Contract NAS 9-3671

Prepared for

Instrumentation and Electronics Systems Division
National Aeronautics and Space Administration
Manned Spacecraft Center
Houston, Texas 77058

Prepared by

Applied Research
Defense Electronic Products
Radio Corporation of America
Camden, New Jersey 08102

(Authors: J. Bordogna, W. Hannan, C. Reno, R. Tarzaiski)

November 1965

ABSTRACT

The objective of this program was to fabricate a laboratory model of a solar-pumped modulated laser with which technical assessment of an optimum system could be made. The ultimate goal is the transmission of real time television data over ranges of 50 million miles.

Results of this program indicate that the capability of a solar-pumped modulated laser to generate high power coherent radiation at optical wavelengths and its inherent advantages of small size and weight recommend it for spacecraft application. Experimental apparatus used in testing solar-pumped modulated lasers, experimental results of successful transmission of real time television pictures, and tradeoffs for space communication systems are given. It is shown that for the desired information bandwidths, a communication range of 50 million miles requires narrow beamwidths, necessitating extensions in the state of the art in beam collimation and tracking.

TABLE OF CONTENTS

Section		Page
I	INTRODUCTION	1
II	ANALYSIS OF SYSTEM PERFORMANCE	3
	A. Major System Components	3
	B. System Parameters and System Performance	4
III	COMPARISON OF POSSIBLE DEEP-SPACE LASER PUMP POWER SOURCES	31
	A. High-Intensity Lamp Pumping	31
	B. Solar Pumping	34
	C. Comparison of Solar and Lamp Pumping	34
IV	TECHNICAL DETAILS OF A SOLAR-PUMPED LASER COMMUNICATION SYSTEM	37
	A. Laser Crystals	37
	1. Physical Characteristics of Laser Crystals	37
	2. Optimization of Crystal Size	44
	B. Solar Collector	46
	C. Thermal Analysis	58
	1. Introduction	58
	2. Laser Crystal Temperature Distribution and Cooling	59
	3. Optimum Design of Uniform, Rectangular Radiating Fins	64
	D. Modulation	73
	1. Magnetic Modulation	73
	2. Electro-Optic Modulation	77
	3. Wideband Electro-Optic Modulator	95
	E. Beam Extraction, Collimating, Directing, and Aiming	98
	1. Beam Extraction	98
	2. Beam Collimation	100
	3. Directing and Aiming the Beam	105
V	EXPERIMENTAL RESULTS	114
	A. Laboratory Pumping	114
	1. $CaF_2:Dy^{2+}$ Laser	115
	2. YAG:Nd^{3+} and YAG:Nd^{3+} - Cr^{3+} Laser	117

TABLE OF CONTENTS (Continued)

Section		Page
	B. Operation of the Solar-Pumped Laser Model	119
	1. The $CaF_2:Dy^{2+}$ Laser	119
	2. YAG:Nd^{3+} and YAG:Nd^{3+}-Cr^{3+} Lasers	121
	C. Flow Tube Construction	124
	D. Electro-Optic Crystal Modulator	128
VI	CONCLUSIONS AND RECOMMENDATIONS	130
Appendix		
A	Relationship between Photodetector Output Current and Incident Light Power and Signal-to-Noise Ratio of an Optical Receiver	132
B	Relationship between Transmitted Power and Received Power in an Optical Communication System	136
C	Background Noise	138
D	Threshold Detection Level	140
E	Atmospheric Transmission	144
F	Solar Cell Array Weight vs Useful Output Power	152
G	Optimization of Crystal Size for a Cylindrical Laser in a Solar-Pumped Laser System	156
H	Drive Current Required for Magnetic Modulation	172
I	Drive Power for Magnetic Modulation	180
J	Comparison of Modulation Methods	185
K	Range of an Optical Communication System as a Function of Modulation Bandwidth	199
L	Characteristics of Electro-Optic Crystals	202
M	Optimum Electro-Optic Crystal Length	211
N	Relationship between Signal-to-Noise Ratio and Modulation Bandwidth	215

Section I

INTRODUCTION

For communications purposes, the electromagnetic wave spectrum may be divided into two broad regions with a vague boundary in the vicinity of 1 THz. In terms of wavelength, this boundary corresponds to a few tenths of a millimeter and is thought of as lying between the ultramicrowave and infrared regions. The so-called "radio frequencies" below this boundary can be generated by conventional electronic techniques, as high-efficiency, high-intensity, coherent electromagnetic waves. But the generation of high-intensity, coherent waves above 1 THz has, until recently, been limited by an inability to manufacture electronic cavity resonators with dimensions compatible with the short wavelengths involved. The solution to this problem has been achieved through the generation of high frequencies by natural atomic resonances, the outstanding example of which is the laser — available at this time in several forms: gas, crystal, and semiconductor diode. Thus, possibilities now exist for communication at frequencies above 1 THz.

With its ability to generate both coherent and high-intensity radiation at optical wavelengths, the laser provides good reasons for using an optical carrier wave instead of a conventional radio wave for certain applications. For example, more communication channels are possible at optical frequencies, and the wider channel bandwidths available allow wideband modulation methods to be easily employed. Also, the inherent narrow beamwidth of the emitted laser radiation makes longer ranges possible with less power than required by conventional systems. This is especially true for communication from vehicles exploring deep space. Of course, this narrow laser beamwidth, which establishes the longer range property, also places severe tracking requirements on the system. Be that as it may, the

properties of the laser suit it well for use as a communication device for manned deep-space missions.

The implementation of an optical communication system requires the solution of new problems. Some design procedures may follow closely those of a radio communication system; others may diverge radically. Modulation and detection techniques must be investigated along with sources of noise and properties of the transmission medium. The characteristics of laser materials must be understood, and these materials must be continually studied and improved. New devices must be created to modulate and detect the optical carrier.

The approach to this investigation depends, of course, upon the operational requirements of the ultimate system desired. At present, and for the next decade at least, the minimum communication requirements for manned deep-space missions are: a video channel of 5-MHz bandwidth (the equivalent of one commercial quality television channel), a voice channel of 4-kHz bandwidth, a telemetry channel of 1-kHz bandwidth, and a mean transmission distance of 50 million miles. The many aspects of the analysis to follow below are concerned with developing a solar-pumped laser communication system to fulfill these requirements. The analysis begins with the definition of system parameters and a conceptual spacecraft design for the system. Relationships among these parameters are studied and correlated with results obtained through experimentation with a laboratory model of a solar-pumped laser. In addition, a technical assessment is made of the capabilities of an optimum system of this type to meet the stated minimum requirements for manned deep-space missions.

Section II

ANALYSIS OF SYSTEM PERFORMANCE

A. MAJOR SYSTEM COMPONENTS

The major components of a solar-pumped laser communication system are shown in Fig. 1. The laser crystal is mounted at the focal point of a parabolic dish (solar energy collector). In operation, the collector is automatically pointed at the Sun, thus solar-pumping the laser. The laser beam is then suitably modulated, focused into a narrower beam, and transmitted. A parabolic dish at the receiving station focuses the received energy on a photodetector, the output of which is amplified and demodulated.

The design of each of the components making up the system of Fig. 1 must take into consideration the requirements of the system as a whole. Also, because several of these components are extensions of the state of the art, they will force limitations on the performance of the system. Thus, overall design will, of necessity, be based on a consideration of tradeoffs among the many parameters involved. For example, design of the system optics involves the establishment of tradeoffs among collected solar power, collector dish diameter, circle of confusion, laser output power, laser crystal size, tracking accuracy, type of pumping (end or side), and so on. Selection of the laser crystal involves consideration of absorption and radiation spectra, pumping threshold, spatial and temporal distribution of laser output, energy levels and efficiency, and thermal properties. Design of the receiver must be based on good noise qualities, sufficient photodetector responsivity at the wavelength of the laser selected, antenna gain, and detector temperature. Modulation and demodulation methods appropriate for the bandwidth required and the power available must be considered carefully. Devices for accomplishing the type modulation desired must be created and optimized.

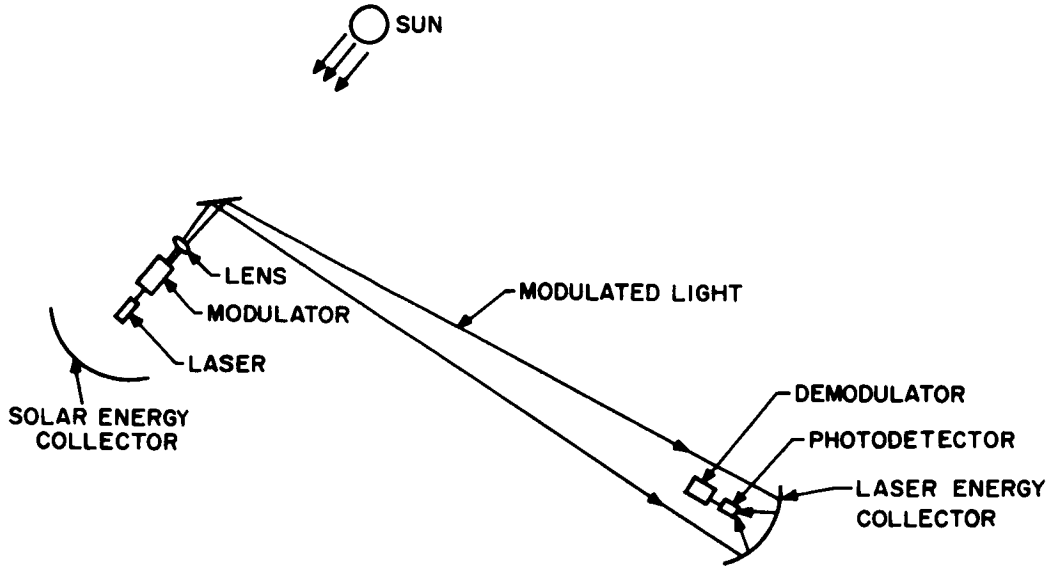


Fig. 1. Major components of a solar-pumped laser communication system.

All of these problems are interrelated in many ways; it is the object of this report to first clarify what the problems are and second, to provide sensible and practical solutions. The first step toward this end is a discussion of the many parameters involved and the generation of tradeoff curves among these parameters. From a study of these curves in relation to the minimum communication requirements set forth in Sec. I above, a possible system will be developed.

B. SYSTEM PARAMETERS AND SYSTEM PERFORMANCE

The maximum range which can be achieved with a given laser transmitter power (or, conversely, the transmitter power needed for a given communication range) depends upon the beamwidth of the transmitter, the type of modulation used, the area of the receiver dish, and the sensitivity of the receiver. In addition, the properties of the transmission medium and, since we are communicating with light,

background noise caused by reflected sunlight (if any) must be considered. The following analysis provides definitions of these variables and develops mathematical relationships among them. The end result is a system of tradeoff curves from which solar-pumped laser communication systems can be designed. Using these curves, optimum systems for the years 1966 and 1970 are presented.

As shown in Appendices A, B, and C, respectively, the relationships among signal-to-noise power ratio, required transmitter power, received signal power, background noise power, and range through the Earth's atmosphere and through space (or both) can be expressed in the form of the following three equations (corresponding to Eqs. A-14, B-4, and C-3):

$$\frac{S}{N} = \frac{\rho^2 P_s G^2 R_\ell}{2eB\left(\rho P_s + \rho P_b + I_d\right) G^2 R_\ell + FkTB} \tag{1}$$

$$P_t = \frac{\alpha_t^2 R^2}{T_a T_o D_r^2} P_s = \frac{\pi \alpha_t^2 R^2}{4 T_a T_o A_r} P_s \tag{2}$$

$$P_b = T_a T_o \frac{\xi M B_{opt} A_r \alpha_r^2}{4} \tag{3}$$

where

S/N = signal-to-noise power ratio

P_t = required transmitter power (W)

P_s = received signal power (W)

P_b = received background power (W)

G = gain of post-photodetector amplifier

ρ = responsivity of photodetector (A/W)

R_ℓ = load resistance (Ω)

e = charge on an electron (1.6×10^{-19} C)

B = bandwidth of pre-demodulation filter (Hz)

F = noise factor of post-photodetector amplifier

k = Boltzmann's constant (1.38×10^{-23} J/°K)

T = receiver front-end temperature (°K)

I_d = photodetector dark current (A)

α_t = transmitter beamwidth (rad)

α_r = receiver beamwidth (rad)

R = range (m)

D_r = diameter of receiver collecting optics (m)

A_r = area of receiver collecting optics (m^2)

T_o = transmission of receiver optics

T_a = transmission of atmosphere

M = solar irradiance ($W/m^2/\mathring{A}$)

B_{opt} = bandwidth of receiver optical filter ($\mathring{A}$)

ξ = reflectivity of background

As described in subsequent sections of this report, the YAG laser is the best CW crystal laser available for solar-pumped use at the present state of the art for several reasons: it is easy to handle, it operates at elevated temperatures, its emitted radiation has a wavelength which can be detected with a good photodetector, and its output can be wideband modulated with an available electro-optic crystal.

The wavelength of the radiation emitted by the YAG laser is 1.06 μm. As shown in Fig. A-1, the S-1 photosensitive surface (the best available at this wavelength) responds to this wavelength with a responsivity of 0.3×10^{-3} A/W. A good photodetector with an S-1 surface is the RCA 7102 multiplier phototube. The salient feature of this tube is the high, essentially noise-free gain of its electron multiplier. This gain is high enough (for the application at hand) to make amplifier noise N_a (see Eq. A-12) negligible compared to the amplified shot noise N_s (see Eq. A-10) from the photocathode. Thus, with the multiplier phototube detector, Eq. 1 becomes

$$\frac{S}{N} \approx \frac{\rho^2 P_s^2}{2eB(\rho P_s + \rho P_b + I_d)} \tag{4}$$

This is an important equation, for it tells us what the received signal power must be for a given signal-to-noise ratio. In particular, Eq. 4 allows us to calculate the lowest received signal power tolerable for acceptable reception of the transmitted information. This minimum acceptable received signal power is dependent on the lowest signal-to-noise ratio at which the receiver will function. According to the analysis given in Appendix D, this receiver "threshold detection level" is 9 dB, i.e., a signal-to-noise ratio of approximately 10:1. Substituting this value of S/N into Eq. 4, we get the following quadratic equation on relating P_s^*, the received signal power at threshold, to the system parameters:

$$\left(P_s^*\right)^2 - \frac{20eB}{\rho} P_s^* - \frac{20eB}{\rho}\left(P_b + \frac{I_d}{\rho}\right) = 0 \tag{5}$$

With Eqs. 2, 3, and 5 we can now analyze the performance of an optical communication system employing a multiplier phototube receiver. The question at this point is: What should be the numerical values of the parameters in such a system? Or, better yet: What can these parameter values be in light of restrictions placed on the system by the state of the art? In other words: Just what is the optimum system we can produce?

We will attempt to answer these questions by considering three possible systems:

(1) An experimental system which is designed to make efficient use of readily available components. (The parameter values for this system do not represent the best that can be realized at the present state of the art, but simply what can be realized with a readily-built laboratory model of the system. Increasing the diameter of the receiving dish, cooling the photodetector, and reducing channel bandwidth, among other things, would result in extended communication range with this system.) Although this experimental system does not represent an optimum design, it serves as a model for providing experimental verification of theory. As such it is worthwhile to predict the communication range it can realize. Experimental verification of the theoretical predictions will provide the basis for assessing the capabilities of an optimum system.

(2) An optimum system which is designed to make efficient use of state-of-the-art components for the year 1966.

(3) An optimum system which is designed to make efficient use of state-of-the-art components for the year 1970.

In addition to considering the analysis of the optical communication problem from the point of view of these three systems, we must also consider the possible environments for these three systems. Since the laboratory system will be Earth-bound, its performance will be based on both its transmitter and receiver existing in Earth's atmosphere at sea level. The optimum system of 1966 has one end of the transmission link on Earth and the other in space. The performance of the optimum system of 1970 will be considered from two points of view: (1) one end of the transmission link on Earth and the other in space, and (2) both ends of the communication link in space.

The effect of Earth's atmosphere on system performance is accounted for by the atmospheric transmission factor T_a in Eqs. 2 and 3. Values of this factor are dependent on meteorological conditions and the height above Earth's surface at which the receiver and/or transmitter is located. A thorough discussion of T_a is given in Appendix E where it is shown that on a clear day, $T_a \approx 80$ per cent

for a range from sea level to outer space. Even more significant than this is that the value of T_a for one end of the communication link located at a height of 1.5 miles ($\approx$ 8,000 ft or 2.4 km) and the other end in space is approximately 96 per cent. This means that if one end of the range is located at a high cloudless elevation, atmospheric attenuation of the communication beam is negligible. Of course, T_a = 100 per cent if both ends of the communication link are located in space.

We now have enough information to specify some values for the system parameters which appear in Eqs. 1, 2, and 3. A summary of these values is given in Table I for the several systems postulated above. Note in Table I that the several systems are identified as System 1, System 2, System 3 and System 3'. The following analysis will describe the performance of these systems in detail.

SYSTEM 1

In this system, both the transmitter and receiver are located in Earth's atmosphere. This fact should be kept in mind at all times when viewing the results of this analysis, for the Earth-Space and Space-Space systems (which illustrate the ultimate application) will differ significantly from the Earth-Earth system.

Substituting the System 1 parameter values into Eq. 3 yields $P_b = 7.1 \times 10^{-9} T_a$. Assuming $T_a = 1$ (a worst-case condition as far as background noise is concerned), $P_b = 7.1$ nW. Compared to I_d/ρ, this value of P_b can be neglected in Eq. 5, since $I_d/\rho = 3.3 \times 10^{-7}$ W for System 1. Thus, for System 1, Eq. 5 yields a received signal threshold power of $P_s^* = 1.81 \times 10^{-7}$ W. With this value of P_s, Eq. 2 becomes

$$P_t = 4.5 \times 10^{-11} \frac{R^2}{T_a} \tag{6}$$

Table I. SYSTEM PARAMETERS AND SPECIFICATIONS

	System 1	System 2	System 3	System 3′
Parameter	Present Experimental System Earth-To-Earth	Optimum System 1966 Earth-To-Space At Mars	Optimum System 1970 Earth-To-Space At Mars	Optimum System 1970 Space-To-Space
λ	1.06 μm	1.06 μm	1.06 μm	1.06 μm
Collector Mirror Diameter D_c	31 in (= 0.787m)	31 in (= 0.787m)	45 in (= 1.143 m)	45 in (= 1.143 m)
Pump Power Collected	300 W	100 W	200 W	200 W
Laser Efficiency	0.1%	1%	5%	5%
Laser Output Power P_t	300 mW	0.5-1 W	5-10 W	5-10 W
D_r	12 in (= 0.305m)	30 in (= 0.762 m)	60 in (= 1.52 m)	60 in (= 1.52 m)
A_r	0.073 m^2	0.456 m^2	1.81 m^2	1.81 m^2
B	6 MHz	6 MHz	6 MHz	6 MHz
ρ	$3x10^{-4}$ A/W	$3x10^{-4}$ A/W	$3x10^{-1}$ A/W	$3x10^{-1}$ A/W
M	0.1 W/m^2/Å	0.1 W/m^2/Å	0.1 W/m^2/Å	NA
T_a	Depends on Transmission Range (see Figs. E-1, 2)	0.9	0.9	1.0
T_o	0.4	0.4	0.4	0.4
B_{opt}	100 Å	100 Å	100 Å	100 Å
S/N	10	10	10	10
α_t	3 mrad	100 μ rad	10 μrad	10 μrad
α_r	1 mrad	1 mrad	10 μrad	10 μrad
I_d	10^{-10} A	10^{-12} A	10^{-12} A	10^{-12} A
k	$1.38x10^{-23}$ J/°K	$1.38x10^{-23}$ J/°K	$1.38x10^{-23}$ J/°K	$1.38x10^{-23}$ J/°K
T	300°K	300°K	300°K	300°K
ξ	0.1	0.1	0.1	NA

for System 1. According to the atmospheric transmission values T_a given by Fig. E-2, transmitter power and transmission range, as given by Eq. 6, are related as shown in Fig. 2 for several meteorological conditions (i.e., very clear, clear, and hazy atmosphere). From this information it is evident that a transmission range of 10 miles can be achieved on a clear day with the laser output power of 300 mW specified for System 1. On a very clear day, the transmission range is extended to about 17.5 miles for the same laser output power.

System 2

In this system, one of the terminals of the transmission link is located in space and the other on Earth at a high altitude. Other differences from System 1 include the magnitude of the power output (0.5 to 1 W compared to 300 mW), the diameter of the receiving dish (30 inches instead of 12 inches), the beamwidth (100 μrad compared to 3 mrad), and the receiver dark current magnitude which is made two orders of magnitude smaller by cooling the multiplier phototube.

Substituting the System 2 parameters into Eq. 3 yields $P_b = 4.1 \times 10^{-8}$ W. Substituting this value of P_b into Eq. 5 yields received signal threshold powers P_s^* of 9.4×10^{-8} W and 1.6×10^{-9} W for information bandwidths B of 6 MHz and 5 kHz, respectively. With these values of P_s for System 2, Eq. 2 becomes

$$P_t = 4.4 \times 10^{-7} \alpha_t^2 R^2, \quad B = 6 \text{ MHz} \tag{7}$$

$$P_t = 5.5 \times 10^{-9} \alpha_t^2 R^2, \quad B = 5 \text{ kHz} \tag{8}$$

Required transmitter power and transmission range, as given by Eqs. 7 and 8, are related as shown in Figs. 3 and 4 for several values of α_t. Note from Fig. 3 that the specified values of α_t and P_t for System 2 (i.e., $\alpha_t = 100$ μrad and $P_t = 1$ W from Table I) permit a transmission range of 10,000 miles.

SYSTEM 3

System 3 differs from System 2 in that laser power output is greater (5-10 W compared to 0.5-1 W), the receiver dish diameter is greater (60 inches instead of 30 inches), the responsivity is greater (0.3 A/W compared to 3×10^{-4} A/W), and the transmitter and receiver beamwidths are narrower (10 μrad compared to 100 μrad, and 10 μrad compared to 1 mrad, respectively). The location of the terminals is the same as in System 2: one in space and one on Earth at a high altitude.

The significant increase in responsivity predicted for System 3 in 1970 is based on rapidly developing improvements in the performance of photoconductive devices. Recent studies[1] have indicated that although the photomultiplier (in its appropriate spectral region) is the most sensitive device for low modulation frequencies (its very high current gain more than offsetting the signal loss from a somewhat inferior quantum efficiency), in the megahertz region, the photoconductor can have sufficient gain to outperform the photomultiplier. In particular, for the case of wavelengths in the region of 1.06 μm, photomultiplier surfaces are very low in sensitivity (e.g., the best surface presently available at 1.06 μm yields a responsivity of only 3×10^{-4} A/W). Thus, with the rapid advances being made in the development of the photoconductive optical detector, it is entirely conceivable that a high-gain, low-noise, good-responsivity device will be available for practical use in 1970.

As listed in Table I above, a reasonable value for the responsivity of a practical photoconductive device is 0.3 A/W. This value, along with present-day values of gain and noise equivalent power equal to 100 and 10^{-15} W, respectively, again allows the receiver to be considered as shot noise limited — just as in the case of the photomultiplier receiver. Therefore, the approximation of Eq. 1 to Eq. 4 is still valid.

---

[1] H. S. Sommers, Jr. and E. K. Gatchell, "Sensitive Broad-Band Photoconductive Detector," 1965 Annual Meeting of the Optical Society of America, 5-8 October 1965.

Thus, substituting the System 3 parameters into Eq. 3 yields $P_b = 1.63 \times 10^{-11}$ W. Substituting this value of P_b into Eq. 5 yields received signal threshold powers P_s^* of 8×10^{-11} W, 5.8×10^{-11} W, and 1.1×10^{-12} W for information bandwidths B of 6 MHz, 4 MH_z, and 5 kHz, respectively. With these values of P_s^* for System 3, Eq. 2 becomes

$$P_t = 9.6 \times 10^{-11} \alpha_t^2 R^2, \quad B = 6 \text{ MHz} \tag{9}$$

$$P_t = 6.9 \times 10^{-11} \alpha_t^2 R^2, \quad B = 4 \text{ MHz} \tag{10}$$

$$P_t = 1.3 \times 10^{-12} \alpha_t^2 R^2, \quad B = 5 \text{ kHz} \tag{11}$$

SYSTEM 3'

System 3' differs from System 3 in only two respects: Since operation is entirely in space, background radiation is negligible and atmospheric attenuation does not exist. Therefore, $P_b = 0$ and $T_a = 1$ for System 3'. Thus, substituting the System 3' parameters into Eq. 5 yields $P_s^* = 6.7 \times 10^{-11}$ W, 4.6×10^{-11} W, and 4.5×10^{-13} W for bandwidths of 6 MHz, 4 MHz, and 5 kHz, respectively. With these values of P_s^* for System 3', Eq. 2 becomes

$$P_t = 7.3 \times 10^{-11} \alpha_t^2 R^2, \quad B = 6 \text{ MHz} \tag{12}$$

$$P_t = 4.9 \times 10^{-11} \alpha_t^2 R^2, \quad B = 4 \text{ MHz} \tag{13}$$

$$P_t = 5.9 \times 10^{-13} \alpha_t^2 R^2, \quad B = 5 \text{ kHz} \tag{14}$$

Required transmitter power and transmission range, as given by Eqs. 12, 13, and 14, are related as shown in Figs. 5, 6, and 7, respectively, for several values of α_t, the transmitter beamwidth.

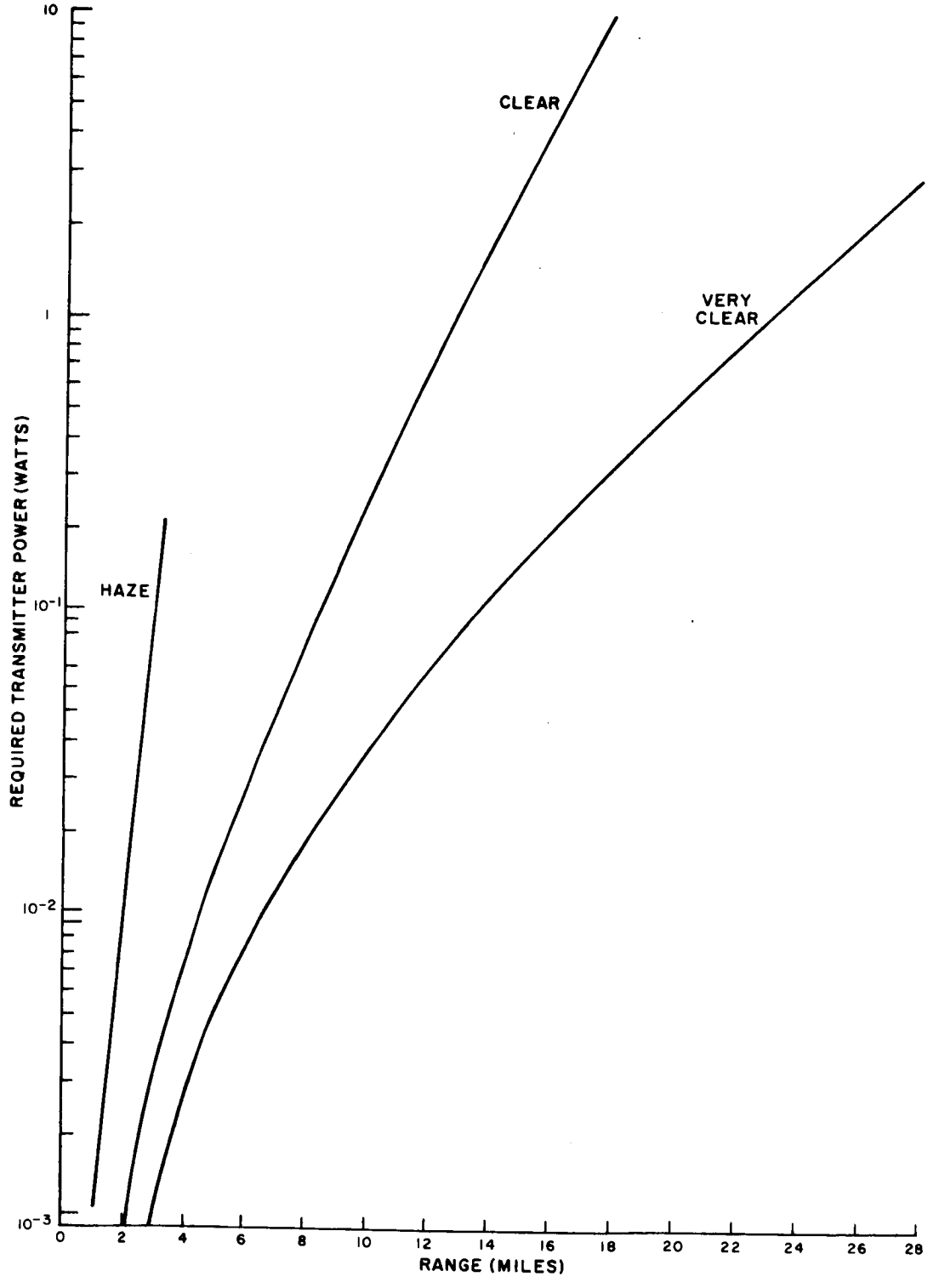


Fig. 2. Required transmitted power vs range through the atmosphere for System 1.

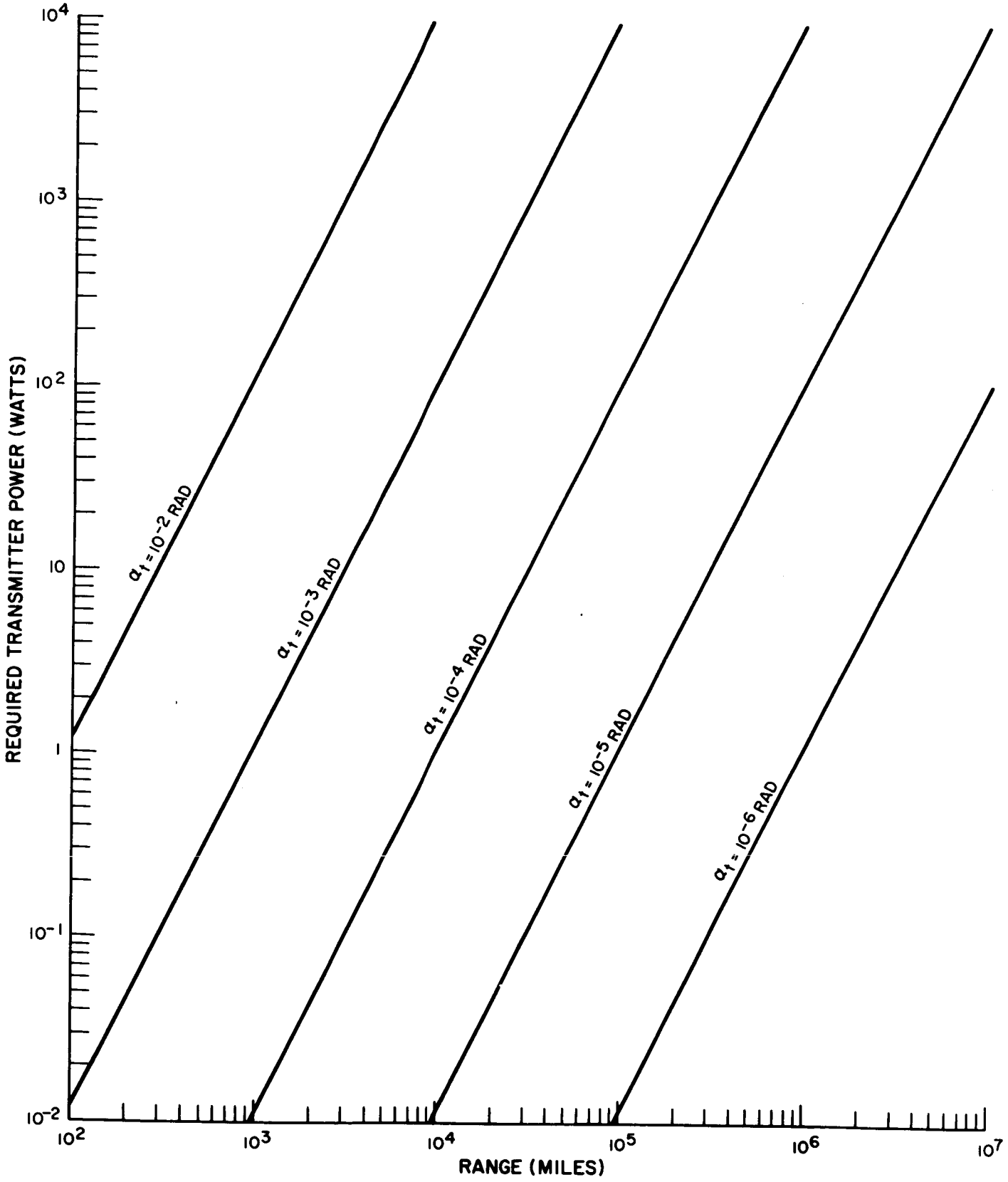


Fig. 3. Required transmitter power vs range between space and Earth for System 2 with B = 6 MHz.

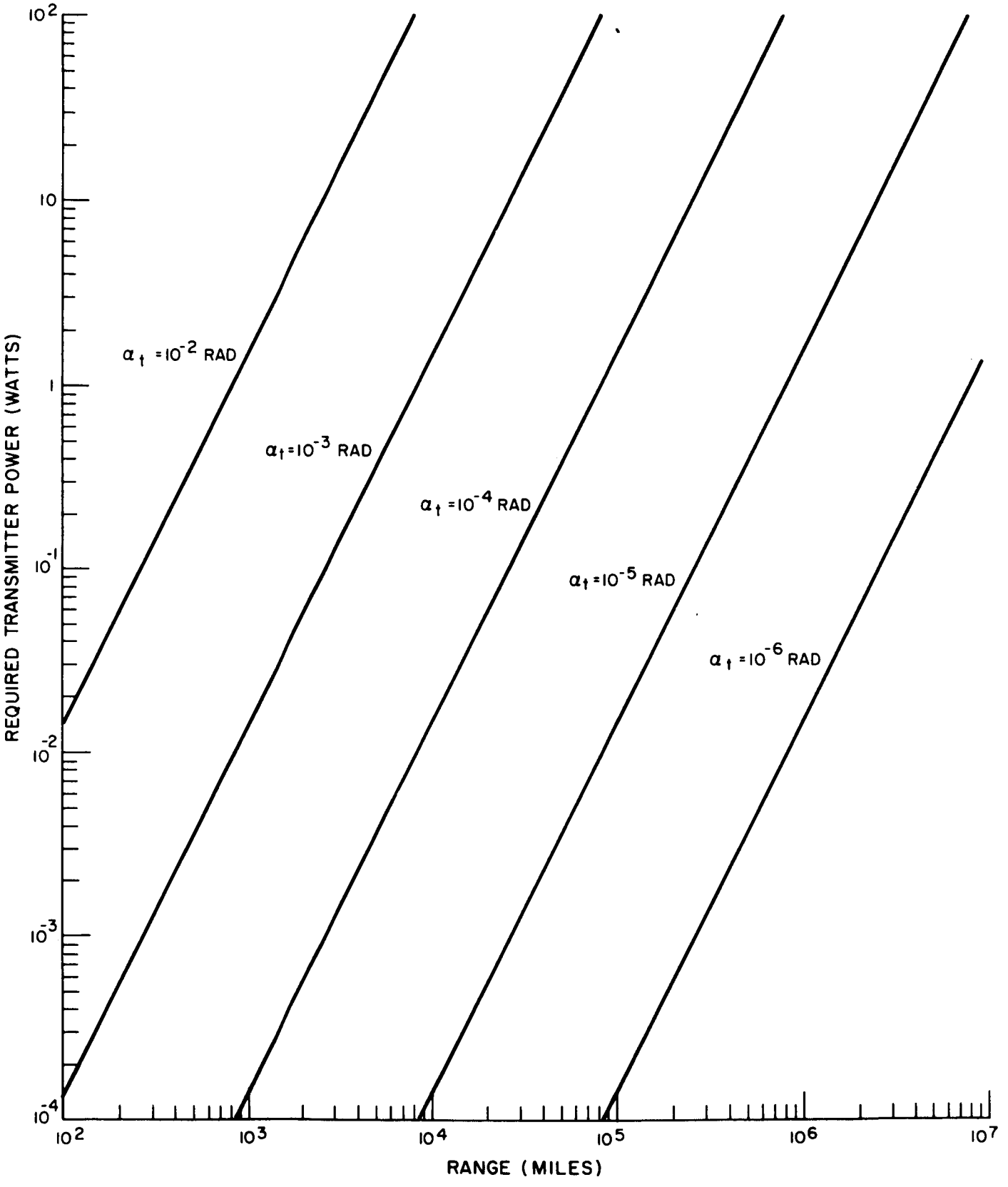


Fig. 4. Required transmitted power vs range between space and Earth for System 2 with B = 5 kHz.

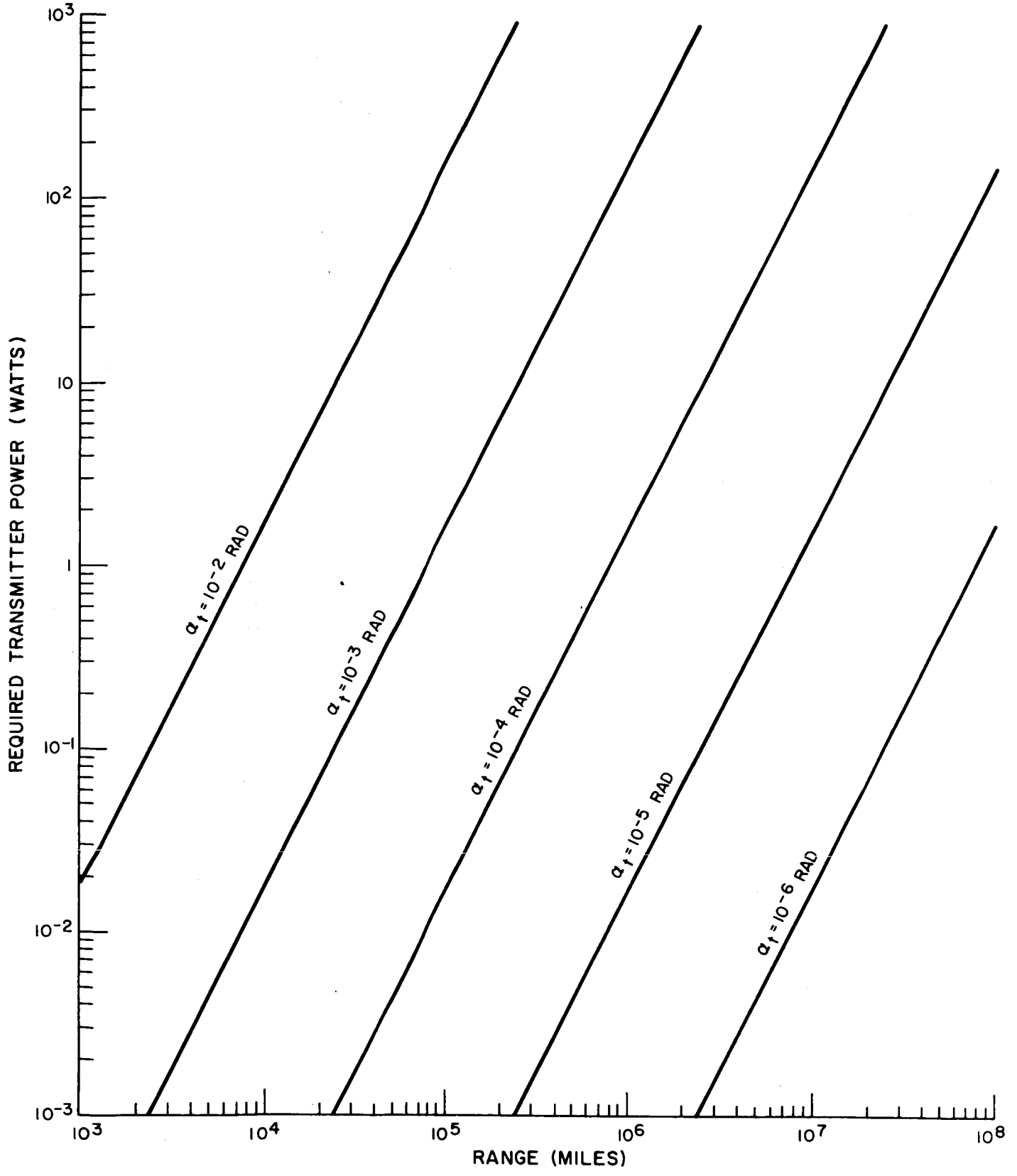


Fig. 5. Required transmitter power vs range in space for System 3' with B = 6 MHz.

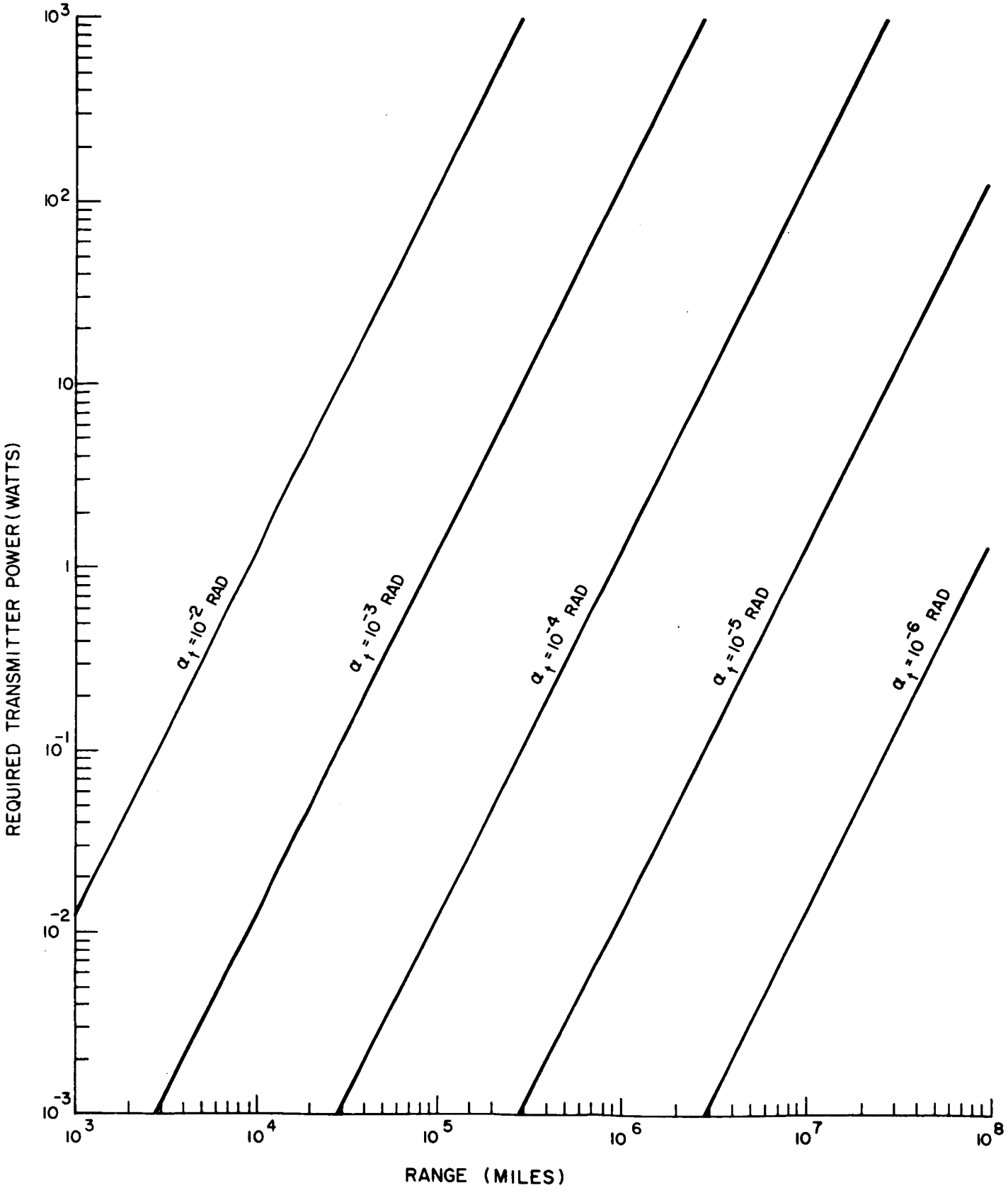


Fig. 6. Required transmitter power vs range in space for System 3' with B = 4 MHz.

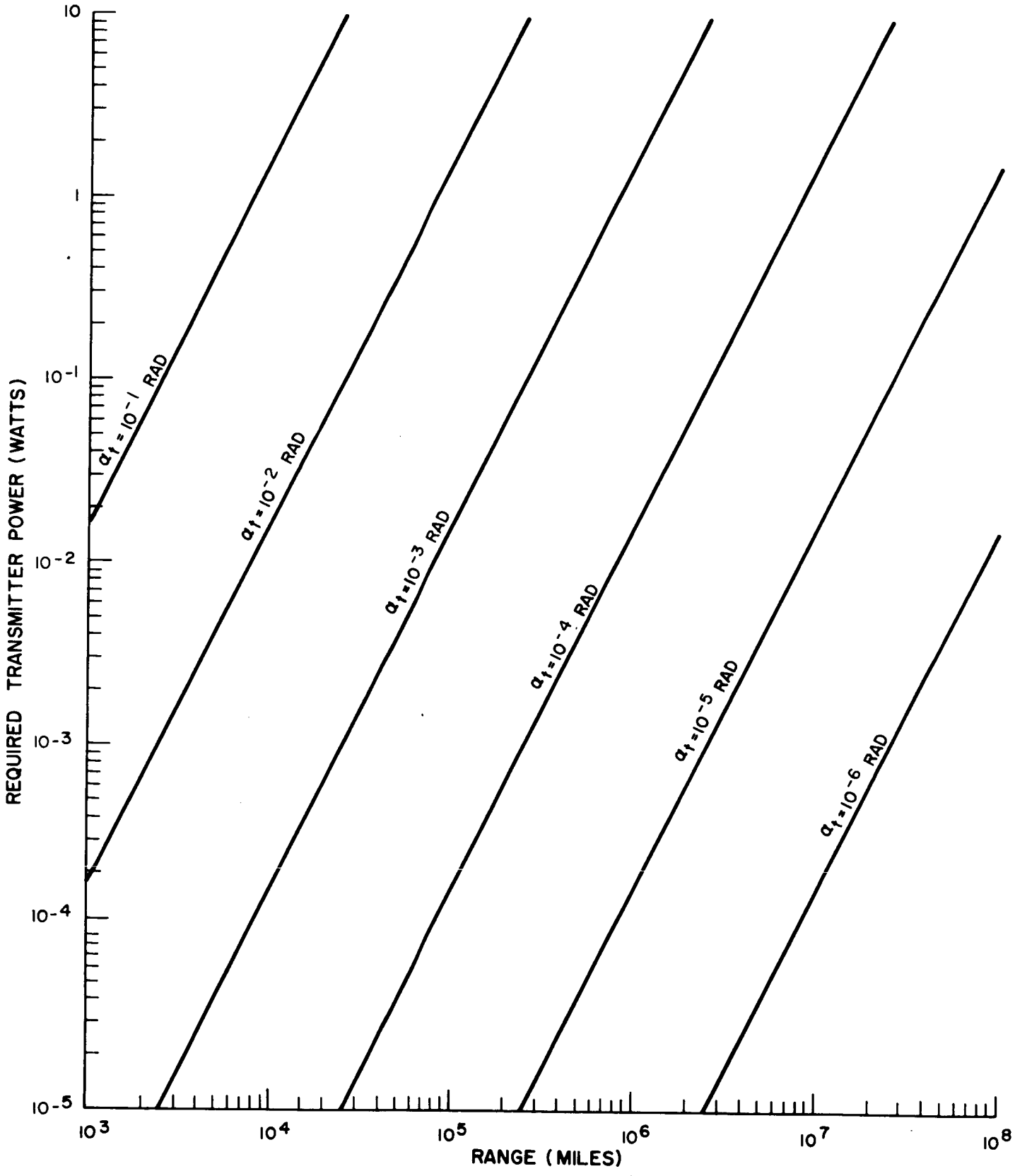


Fig. 7. Required transmitter power vs range in space for System 3' with B = 5 kHz.

Note that Eqs. 9, 10, and 11 are approximately equal, respectively, to Eqs. 12, 13, and 14. Thus, there is very little difference in performance between Systems 3 and 3', and Figs. 5, 6, and 7 apply to both.

Note from Fig. 5 that the specified values of α_t and P_t for System 3' (i.e., $\alpha_t = 10$ μrad and $P_t = 15$ W from Table I) permit a transmission range of 24 million miles. For a 50-million-mile range at $P_t = 10$ W, a beamwidth somewhat less than 10 μrad (say, about 9 μrad) is required. Of course, if the information bandwidth B is only 5 kHz (a real-time voice channel), then from Fig. 7, a range of 80 million miles can be achieved with a power output of 1 W at a beamwidth of 10 μrad, or 24 million miles with a power output of 10 W and a beamwidth of 100 μrad. Note also from Fig. 7 that a range of 70,000 miles can be achieved with a relatively large beamwidth of 10 mrad at 1 W.

From the range-power output data compiled above for Systems 3 (Earth to Space) and 3' (Space to Space), it is evident that (1) there is not much difference between the two systems, and (2) a transmitter beamwidth of 10 μrad (or less) is necessary to achieve a communication range of 50 million miles at the predicted laser power outputs available in 1970. To satisfy these requirements, a configuration such as that illustrated in Fig. 8 can be implemented. In this system, the laser receiver is slaved to a powerful laser beacon positioned at a high altitude on Earth's surface or on board an Earth-orbiting satellite.[1] In operation, the beacon monitor on the spacecraft tracks the beacon and directs the narrow-beamwidth, wide-bandwidth laser communication signal toward the laser receiver. In effect, this is an open-loop control system. Implementation of this system involves the careful consideration of factors such as Earth rotation (the beacon must not be lost behind Earth at any time) and translation, signal transmission time

[1] If the beacon is on board an Earth-orbiting satellite, the orbit of the satellite can be fixed to avoid occultation by Earth at any time during an Earth-Mars mission.

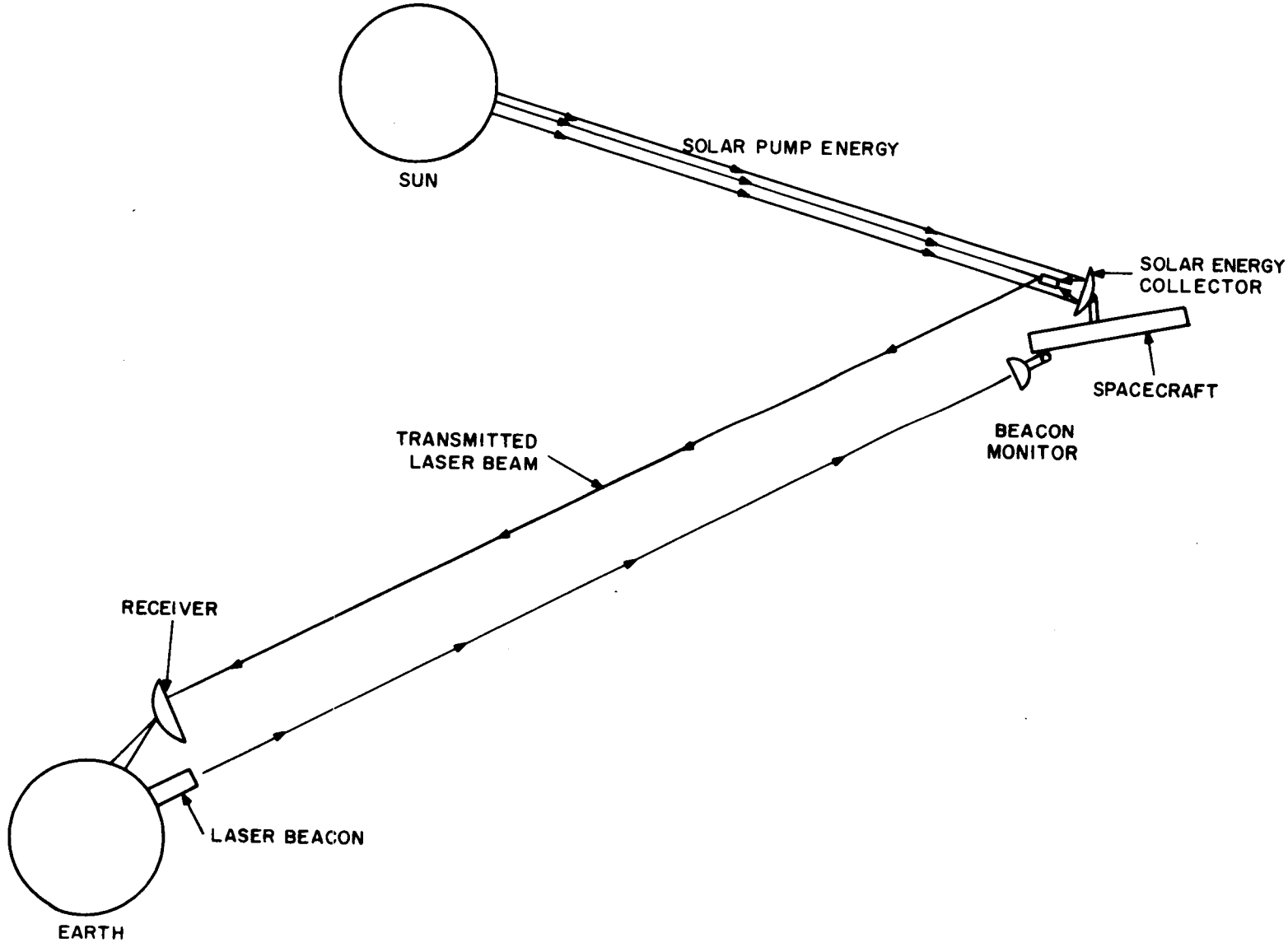


Fig. 8. Deep space to Earth communication system.

($\approx$ 530 seconds two-way transmission time between Earth and Mars), and maneuvering capabilities of the spacecraft (i.e., how quickly the spacecraft can change course). A suggested system (in which the beacon is located on Earth's surface)[1] which employs compensation for these factors is shown schematically in Fig. 9. In this system, communication between Earth station (actual ground base or satellite) and spacecraft is initiated before launch or at some time early in the mission, and is maintained continuously throughout the mission in the following manner: Observe in Fig. 9 that the first signal is received at Earth station at time "t". This time is arbitrary and may be chosen to be the time of initiation of the continuous communication process.

This signal will be assumed to be of duration t_{m1} and to contain information which will allow the determination of the spacecraft course for a time Δt, where Δt is a function of the information processing time, signal transmission time, and the time required to perform any necessary mechanical positioning of Earth (or satellite) and spacecraft antennas. (Δt will assume its maximum value when the spacecraft is at its maximum distance from the Earth station.) The received signal is sent to the Earth computer where:

(1) The future course of the spacecraft is plotted.

(2) Aiming information for the Earth beacon and receiver is generated.

(3) Aiming information for the spacecraft transmitter and receiver is generated.

The time required to perform these functions will be denoted as the information processing time t_p for the Earth computer. After the functions have been performed, the Earth beacon and receiver are aimed to a point on the projected course of the spacecraft, this point being that at a time given by $t + t_{m1} + t_p + t_a$, where t_a is the time required to maneuver the Earth beacon and receiver. After this time,

[1] A similar system applies if the beacon is carried on board an Earth-orbiting satellite.

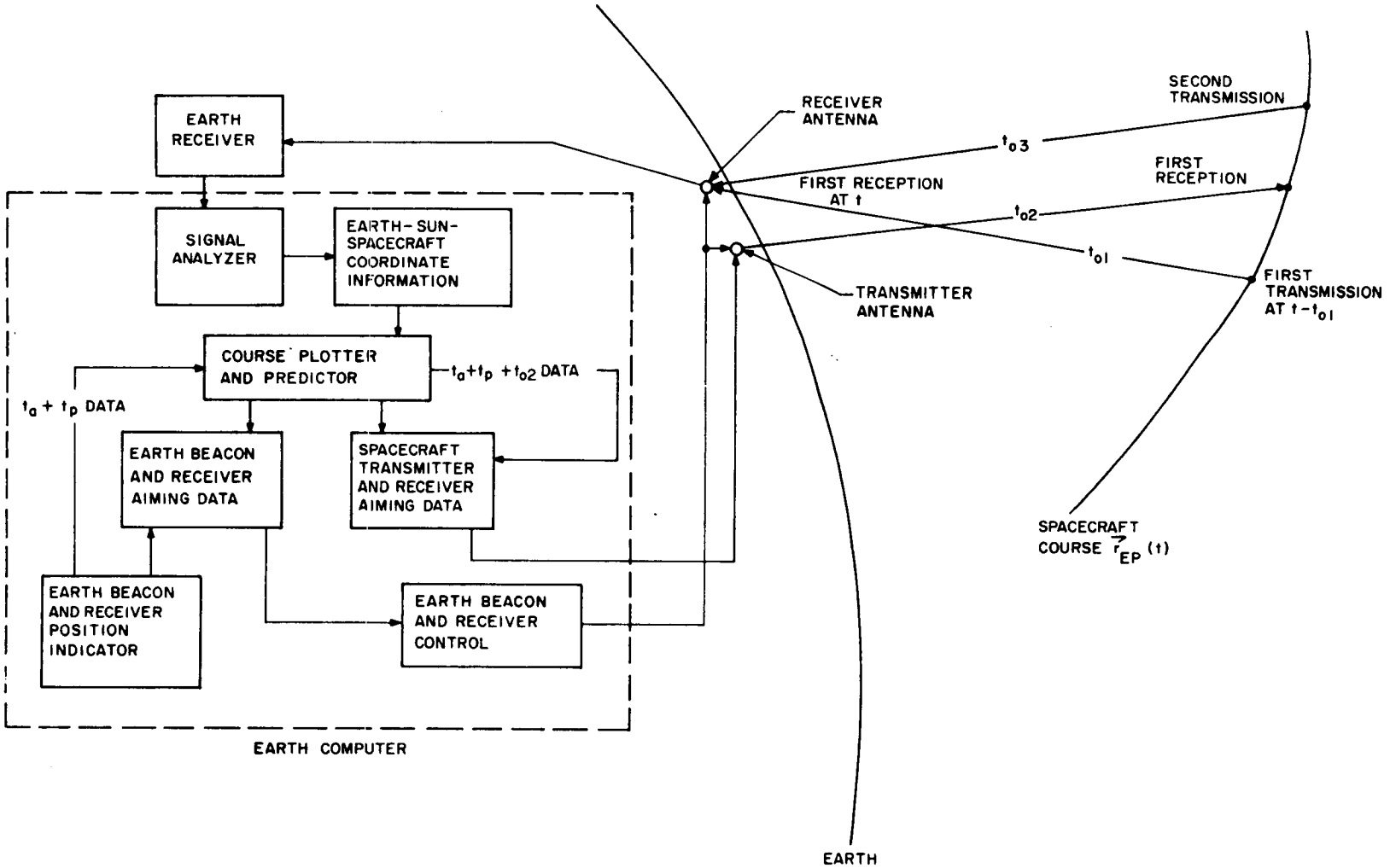


Fig. 9. Scheme for continuous Earth-spacecraft communication.

the beacon and receiver follow the course plotted by the Earth computer. When the beacon is aligned to follow the projected spacecraft course, information for aiming the spacecraft transmitter at Earth is sent to the spacecraft. This information is received there at a time $t + t_{m1} + t_p + t_a + t_{02}$, with:

$$t_{02} = \left| \frac{\vec{r}_{EP}}{c} \right|$$

where

$|\vec{r}_{EP}|$ = the distance between the Earth station and the spacecraft. (It is seen that t_{02} will be a function of message duration if the spacecraft is moving. The Earth computer must be capable of taking this into account. If initial contact is assumed made before launch, t_{02} is constant.)

c = the velocity of light

The duration of this message is t_{m2}, where t_{m2} is sufficient time to allow the spacecraft receiver to be locked onto the Earth beacon. After this initial reception, the spacecraft transmitter and receiver are aimed (using the information received during the first spacecraft reception) to track the Earth station.

Spacecraft transmitter and receiver aiming can be controlled by a relatively simple computer carried on board the spacecraft. When this aiming is complete, the spacecraft can again transmit to the Earth station, and is ready to receive continuous transmission from Earth. After the procedure above has been completed, as long as the spacecraft can supply position information continuously to Earth, the Earth computer can continuously update the plotted course of the spacecraft so that the alignment at each end of the communication link can be continually adjusted as the spacecraft follows its course. The realignment necessary is accomplished by repeating the procedure used for initial alignment with the exception that the receiver need no longer be required to lock on the beacon beam.

Due to the fact that the course information is in fact delayed during its transmission, the question arises as to how this affects the alignment of the communication link. Consider, for example, the case of communication to any point $|\vec{r}_{EP}|$ away where signal transmission time (two way) is approximately t_0 seconds. Let the other delays (processing time, message duration, aiming times) be denoted by t'. After the spacecraft transmits position information, it must wait at least $t_0 + t'$ seconds for instructions to readjust its communication system alignment. It becomes reasonable to ask whether the spacecraft can, in $t_0 + t'$ seconds, move in such a way that its receiver will no longer intercept the beacon beam or that its transmitter will no longer be pointed at the Earth station. It is clear that only movements in a direction normal to $\vec{r}_{EP}$ are of interest in this connection, and, in fact, only relatively "sudden" movements which are the effects of "sudden" accelerations. For such a movement to cause a loss in communications, it is necessary that the spacecraft transmitter move an angle equal to one-half its beamwidth or for the Earth beacon to move one-half its beamwidth. Such a movement as the former may cause temporary loss of communication, since the Earth computer is still able to track the spacecraft for some time after the link is broken. A movement of the latter type will probably cause permanent loss of communication, since the Earth computer will now direct the beacon on a course which the spacecraft no longer follows. The uniform[1] acceleration of a spacecraft during a time $t_0 + t'$ required to cause loss of communication is given by:

$$a = \frac{2x}{(t_0 + t')^2} \tag{16}$$

where

$2x$ = the diameter of the beam of the transmitter considered (miles)

$t_0 + t'$ = the total two-way transmission time (seconds)

[1] Uniform accelerations are considered since they represent minimum accelerations required to move a given distance in a given time.

This acceleration may be interpreted as a limit to the uniform spacecraft acceleration which may be applied for a time $< (t_0 + t')$. The spacecraft transmitter beam diameter is given by $\alpha \left| \vec{r}_{EP} \right|$ where α is the beam divergence. For $\alpha = 10^{-6}$ rad, the permitted maximum uniform acceleration for the spacecraft a_s is given by:

$$a_s = \frac{\left| \vec{r}_{EP} \right| \times 10^{-6}}{(t_0 + t')^2} \text{ mi/s}^2 \tag{17}$$

For loss of the Earth beacon the corresponding spacecraft acceleration is (beamwidth of Earth beacon assumed $= 10^{-5}$ rad)

$$a_b = \frac{\left| \vec{r}_{EP} \right| \times 10^{-5}}{(t_0 + t')^2} \text{ mi/s}^2 \tag{18}$$

It is known that $t_0 = 2 \left| \vec{r}_{EP} \right| /c$. In terms of this time, the equations above can be written:

$$a_{s\,max} = \frac{\left| \vec{r}_{EP} \right| \times 10^{-6}}{\left(\frac{2\left| \vec{r}_{EP} \right|}{c} + t' \right)^2} \text{ mi/s}^2 \qquad a_{b\,max} = \frac{\left| \vec{r}_{EP} \right| \times 10^{-5}}{\left(\frac{2\left| \vec{r}_{EP} \right|}{c} + t' \right)^2} \tag{19}$$

It is observed that the permitted acceleration increases with range until $\left| \vec{r}_{EP} \right| = (1/2)\, ct'$ where it reaches a maximum, and then begins to decrease. Fig. 10 shows the maximum permitted uniform accelerations (such that there is no permanent loss of communications) of the spacecraft during a time $t_0 + t'$ vs $\left| \vec{r}_{EP} \right|$ for large values of $\left| \vec{r}_{EP} \right|$. Also, based on the beamwidths assumed above, it is noted that a "temporary loss" of communication could result if an acceleration one-tenth that required for "permanent loss" occurred. As an example, consider communications between Earth and near Mars where $\left| \vec{r}_{EP} \right| \approx 50 \times 10^6$ mi. Assume $(2\left| \vec{r}_{EP} \right| /c) = 528$ s and $t' = 5$s. Then, an acceleration of 0.00173 mi/s^2 for ≈ 530

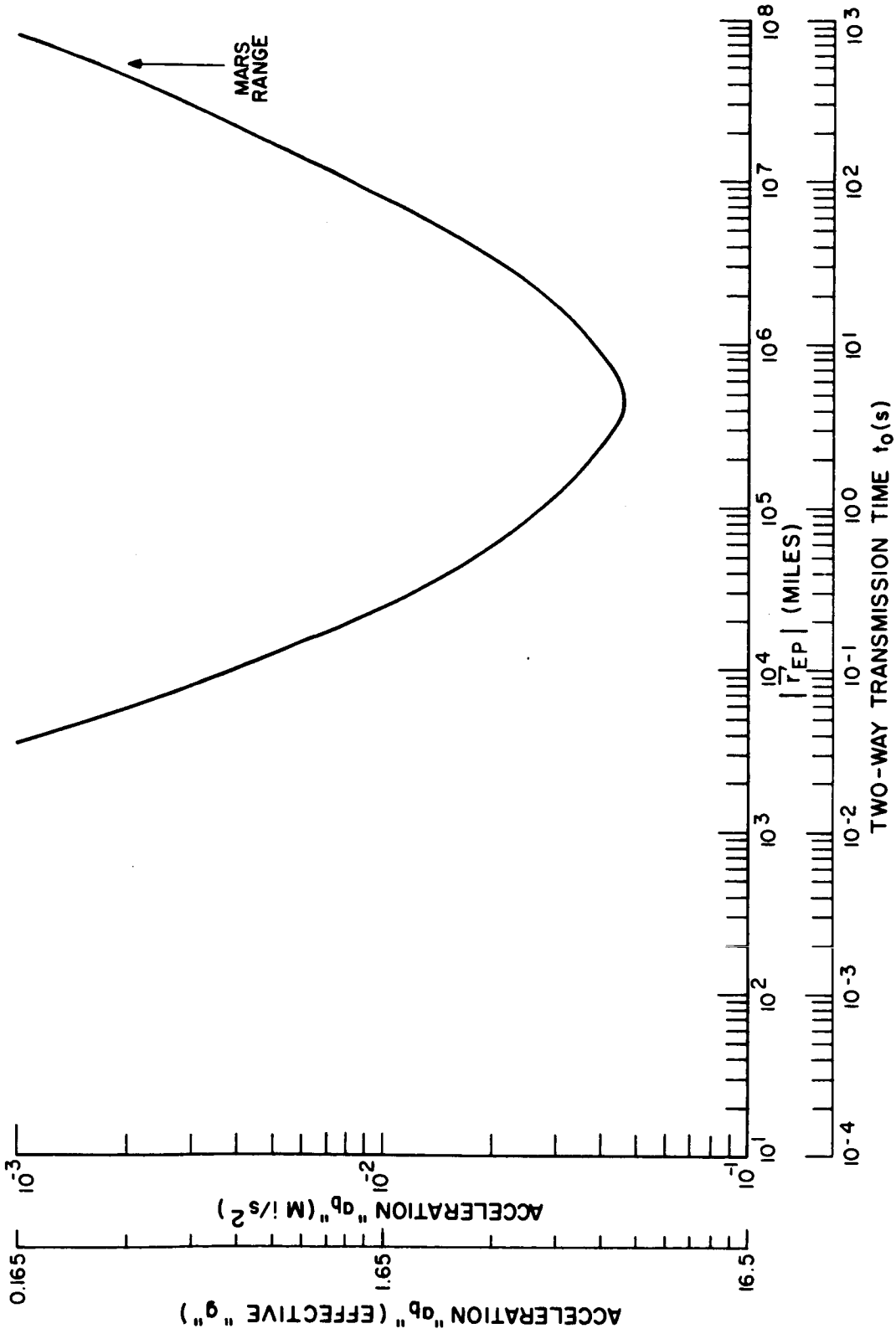


Fig. 10. Uniform acceleration permitted for time $\leq t_0 + t'\left(\text{with } t' = 5s\right) \text{vs } \left|\vec{r}_{EP}\right|$.

seconds will result in permanent loss of communication. This acceleration corresponds to approximately 0.285 g, a condition which will probably never exist due to the tremendous amount of energy required to cause it (for the time involved). In this case a_s (acceleration permitted so communication is never lost) $\approx$ 0.0285 g, a condition still not likely to obtain. This indicates that the communication link can be very reliable at large distances. However, consider the case of operation at small values of $\left|\vec{r}_{EP}\right|$, for example, 10,000 miles. Here $(2\left|\vec{r}_{EP}\right|/c) \approx 0.1$s. If the time t' is of the order of 5 seconds, the values of the "loss accelerations" are

$$a_b \approx \frac{(10^4) \times 10^{-5}}{25} \text{ mi/s}^2 = 0.004 \text{ mi/s}^2$$

$$a_s \approx \frac{(10^4) \times 16^{-6}}{25} \text{ mi/s}^2 = 0.004 \text{ mi/s}^2 \qquad (20)$$

These correspond to $a_b \approx 0.66$ g and a_s 0.066 g, respectively. It is reasonable to expect that accelerations of this magnitude could be applied during the 5-second total communication time at 10,000 miles. This implies that in this region of range, communication requirements must limit the maneuvering of the spacecraft. At distances less than about 5000 miles, the graph of Fig. 10 does not yield accurate values for the permitted acceleration in the 5-second communication time. In that region the beam diameters approach the diameter of the receiver antennas ($\approx$ 4 ft) or become much smaller and the antennas can no longer be considered as points. In this region an estimation may be made to find an "equivalent uniform acceleration," the counterpart of that defined above. It may be assumed that when beam diameter $<<$ antenna diameter, the allowed motion in the 5-second interval is that required to move the beam a distance equal to the radius of the antennas. For both antennas with a diameter of 4 feet the "equivalent uniform accelerations" are:

$$a = \frac{4 \text{ ft}}{25 \text{ s}^2} = 0.16 \text{ ft/s}^2 = 0.005 \text{ g} \qquad (21)$$

In the intermediate range where beam diameter $\approx$ antenna diameter, the actual permitted acceleration depends upon the amount of overlap between beam and antenna. The allowable uniform accelerations in this range for t' = 5 seconds should be of order 10^{-1} g. It thus appears that maintaining contact at close range can be expected to be more difficult than at great distances. The solution to this problem can be obtained in several ways:

(1) Do not allow the spacecraft to maneuver at close range if early acquisition is used to initiate communications.

(2) Allow close maneuvers but perform acquisition at a great distance ($\approx 10^6$ mi).

(3) Provide a method of varying the transmitter beamwidths so that larger values of α can be obtained at close ranges.

It is desirable at this point to investigate the nature of a beacon which could be used in an "open-loop" tracking system.

Assuming a 31-inch beacon monitor receiving mirror with a 3-milliradian field of view (angle size chosen for tracking ease), the Earth subtends an angle of $1/20\ \alpha_r$ at 50 million miles. With this value of $1/20\ a_r$, Eq. 3 yields $P_b = 9.3 \times 10^{-10}$ W, where the parameters $T_a = 0.9$, $T_o = 0.4$, $\xi = 0.1$, $M = 0.1$ W/m^2/Å, and $B_{opt} = 100$ Å have been used.

Now, in order to calculate P_t we must first decide on the kind of laser to use as the beacon. Certainly, in 1970 there will be a much wider choice than at present. Therefore, what we will do here is give an example, using a good present-day laser. Specifically, let us postulate the use of an argon laser, which emits radiation at 5000 Å. According to Fig. A-1, photosensitive surfaces at this wavelength have responsivities on the order of 0.07 A/W.

Thus, with $\rho = 0.07$ A/W and a narrow information bandwidth of B = 10 Hz (which is sufficient for tracking), Eq. 5 yields a value of minimum necessary received signal power (i.e., threshold signal power) of $P_s^* = 1.3 \times 10^{-11}$ W. With this value of P_s, Eq. 2 becomes

$$P_t = 7.1 \times 10^{-13} \alpha_t^2 R^2 \tag{22}$$

which indicates that at a range of 50 million miles and a beacon beamwidth of 1 mrad the power output of the argon laser beacon would have to be 0.5 watt.

Section III

COMPARISON OF POSSIBLE DEEP-SPACE LASER PUMP POWER SOURCES

The suitability of a given type of pump power source for an optically pumped laser on board a deep-space vehicle is determined by the amount of power to be supplied as well as other practical considerations such as size and weight. For the deep-space laser communication system postulated for the year 1970 (see Table I), a laser power output of 5 to 10 watts at 2 per cent efficiency is predicted. A 10-watt laser with an operating efficiency of 2 per cent requires a pump power of 500 watts.

Pump sources which can supply this amount of power in a deep-space environment are of two types: those external to the spacecraft and those carried by it. The only external source capable of supplying 500 watts of pump power is the Sun. An example of the second type is the high-intensity lamp which is, at present (and in the near future), the only practical "artificial" source capable of supplying optical energy at a high power level.[1] Thus, lamp pumping will be considered as the alternative to solar pumping. A comparison between these two sources will serve to illustrate the desirability and practicality of solar pumping.

A. HIGH-INTENSITY LAMP PUMPING

While it is true that different types of lasers require different types of pump lamps, it is possible to consider some properties of lamp-pumped systems in general.

---

[1] The possibility of pumping the transmitting laser with another laser is not considered practical here. The 500-watt output requirement precludes such operation unless the pump laser is almost 100 per cent efficient. A 500-watt 100 per cent efficient laser is far beyond the present state of the art.

One important property of these systems is that the laser crystal must be side pumped, since intense sources needed for end pumping are not available. In certain cases, side pumping can be considerably less efficient than end pumping for a specific pump source. The power absorbed in a crystal is an exponential function of the path length for the pump radiation. Clearly, for end pumping, this absorption path length is at least equal to the length of the crystal. For side pumping, the absorption path length varies depending on the angle of incidence of pump light rays on the crystal. An average absorption path length can be calculated easily. Thus, by Snell's law,

$$n \sin \varphi = n' \sin \varphi'$$

where

φ = angle of incidence of light upon crystal

φ' = angle of refraction of light as it passes through crystal surface

n = index of refraction of medium outside crystal

n' = index of refraction of crystal material

Now the absorption path length d inside the crystal is a function of φ' and the crystal diameter D, viz. ,

$$d = D \cos \varphi'$$

Solving for φ' from Snell's law and substituting in the relation for d yields,

$$d = D \cos \left[\sin^{-1} \left(\frac{n}{n'} \sin \varphi \right) \right]$$

If it is assumed that the light inside the crystal is not reflected at the crystal boundaries, the average absorption path length $\bar{d}$ is, therefore,

$$\bar{d} = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} D \cos \left[\sin^{-1} \left(\frac{n}{n'} \sin \varphi \right) \right] d\varphi$$

$$= \frac{2D}{\pi} \int_{o}^{\frac{\pi}{2}} \sqrt{1 - \left(\frac{n}{n'}\right)^2 \sin^2 \varphi} \; d\varphi$$

$$= 0.78834D$$

Thus, the average absorption path length for side pumping is 0.79 times the crystal diameter. Therefore, if most of the pump power available is not absorbed in a distance 0.79 times the crystal diameter, end pumping can be a more efficient process. If the opposite is true, the side pumping efficiency can approach the end pumping efficiency as a limit. This will be the case when pump radiation is confined to the absorption bands of an optically dense laser material. High-pressure mercury-lamp pumping of YAG:Nd^{3+}-Cr^{3+} is an example of this.

If a lamp-pumped laser operates with a 10-watt power output and an efficiency of 2 per cent, a minimum of 500 watts of pump power is needed in the volume of the crystal. Under the assumptions (1) that the lamp converts 80 per cent of the electrical input power to light and (2) that it is possible to couple 50 per cent of the available light into the crystal, it follows that for side pumping a minimum electrical power input of 1250 watts is required. Thus, this type of laser system requires that a spacecraft carry at least a 1250-watt electrical power source.

The above laser system also requires a laser cooling supply with a minimum cooling capacity such as to maintain a proper operating temperature for the laser under conditions of 500 watts input power. Beyond this is the additional requirement of a lamp cooling system.

Another important consideration in the use of lamp pumping is the limited lifetime of the lamp used, typically 30 to 70 hours of operation. This may, in fact, preclude the use of lamp pumping on an unmanned space mission.

B. SOLAR PUMPING

The use of the solar pump source eliminates the need of an electrical power source from which pump power is obtained after energy conversion. Also, because of the brightness of the source, it is possible to end pump the laser crystal and thereby make use of the most efficient pumping method for a given laser with any pump source. The overall efficiency of solar pumping is therefore determined by the general absorption characteristics of the laser only (match between solar emission spectrum and laser absorption spectrum). In this case then, for a laser which operates at 2 per cent efficiency with a solar pump source, 500 watts of pump power must be supplied to the crystal. This system also requires cooling for maintaining a proper operating temperature with an input of 500 watts to the crystal.

An additional requirement for solar pumping is a tracking system for aiming the collector mirror at the Sun. This adds some complexity to the overall system but, since the aiming system for the transmitter must be accurate to ± 0.5 microradian, whereas the solar tracker must be accurate to ± 3 milliradians, its addition does not impose serious limitations.

C. COMPARISON OF SOLAR AND LAMP PUMPING

An exact comparison of solar and lamp pumping can be made only for a given laser crystal (known material, size, dopant concentration, etc.). However, it is possible to make a useful comparison between two 10-watt laser systems, both two per cent efficient under conditions of either solar or lamp pumping. Such a comparison for a space vehicle in the vicinity of Mars is shown in Table II. (The estimated weight of the lamp-pumped laser power supply in this table was determined by entering Fig. G-3 with a converter electric power of 1250 watts.)

Table II. CHARACTERISTICS OF SOLAR AND LAMP PUMPED LASER SYSTEMS

Characteristic	Lamp-Pumped Laser	Solar-Pumped Laser
Output Power	10 W	10 W
Laser Efficiency	0.02	0.02
Pump Power Required	500 W	500 W
Source of Power	Electrical (For example, from Solar Cells)	Solar Energy Collected by Mirror
Conversion Efficiency: Source Power to Pump Power	0.4 (DC to Optical Power)	1.0 (Direct Pumping)
Total Input Power	1250 W	500 W
Power Supply Weight (Estimated)	730 lb Plus Converter Weight (Solar Cell Source)	100 lb (Total including Solar Collector and Mount)
Power Source Size (Estimated)	245 ft^2 (Area of Solar Cell Array)	11 ft^2 (Collector Dish Area)
Crystal Cooling Capacity	Sufficient for Stable Temperature at 500 W Input	Sufficient for Stable Temperature at 500 W Input
Other Special Properties and Characteristics	Lamp Cooling Required. Pumping Cavity Required. Limited Lamp Life.	Mirror Aiming Required. Laser Mount can Incorporate Cooling.

From the chart it follows that when a given laser is pumped with equal efficiency by solar or lamp pump sources, the choice between the two possible pump sources must then be based on the desirability of using a relatively inefficient, heavy, power source and its associated system components to operate a limited life lamp or a collector mirror with its tracker to pump the laser directly. It is seen that, although both pump sources will operate the laser well, the solar pump source has more advantages for a deep-space application.

The comparison given in Table II does not remain strictly valid if the efficiency of a given laser depends on the pump source used (which is usually the case). If the lamp pumping is less efficient than solar pumping, the undesirable features of the lamp pump source (weight, size, inefficiency) are emphasized more than in the equal efficiency comparison of Table II. If lamp pumping is more efficient, the weight and size requirements of the two systems become more comparable. If lamp pumping is seven times as efficient, these requirements become similar. Since this extremely high relative efficiency has not been realized with lamps used to pump YAG lasers, it is not yet possible to construct a lamp pump source which can match the solar pump source with respect to weight and size.

Section IV

TECHNICAL DETAILS OF A SOLAR-PUMPED LASER COMMUNICATION SYSTEM

A. LASER CRYSTALS

At present, much effort is being devoted to the development of new laser crystals and the improvement of existing laser materials. Several of these have already been found suitable for use as generators of a CW carrier which can be modulated with wideband information. In particular, $CaF_2:Dy^{2+}$ (calcium fluoride doped with divalent dysprosium), $YAG:Nd^{3+}$ (yttrium aluminum garnet doped with trivalent neodymium), and $YAG:Nd^{3+}-Cr^{3+}$ (yttrium aluminum garnet doped with trivalent neodymium and trivalent chromium) are most useful at the present state of the art. The physical properties of these three crystals, their operating characteristics, and their usefulness as CW energy sources are described below. In addition, the optimization of crystal size for solar-pumped operation is discussed in a general way so as to apply to any crystal material used.

1. Physical Characteristics of Laser Crystals

a. Properties of $CaF_2:Dy^{2+}$ [1]

In this crystal the Dy^{2+} dopant ions are substituted for Ca^{2+} ions of the CaF_2 host lattice. The dysprosium enters the lattice in a trivalent state and is reduced to a divalent state at the lattice sites by a process such as photoreduction or electrolytic reduction. In the former, electrons for the $Dy^{3+} \rightarrow Dy^{2+}$ reduction are freed from fluorine atoms by the action of γ radiation. This process

---

[1] J. P. Wittke et al., "Solid State Laser Explorations", Final Report, Contract AFL-TR-64-334, January 1965. (General reference for $CaF_2:Dy^{2+}$ material.)

reduces about 20 per cent of the available Dy^{3+} ions and is reversible under the influence of high temperatures or a high level of optical radiation. Crystals produced by this method are therefore unstable over long periods of high pump level operation. In the electrolytic reduction process, electrons are supplied to Dy^{3+} and fluorine atoms are released from the crystal in gaseous form. Crystals produced by this method are stable and do not exhibit the "bleaching" effects described above.

The process of absorption and consequent laser emission can be shown by the energy level diagram in Fig. 11.

The presence of pump energy in the $CaF_2:Dy^{2+}$ crystal causes transitions from the 4f to the 5d electronic energy levels of Dy^{2+}. The pump bands for this crystal extend from 0.3 μm to 1 μm and thus lie in the visible and infrared portions of the spectrum. Hence, an efficient pump source would be a blackbody

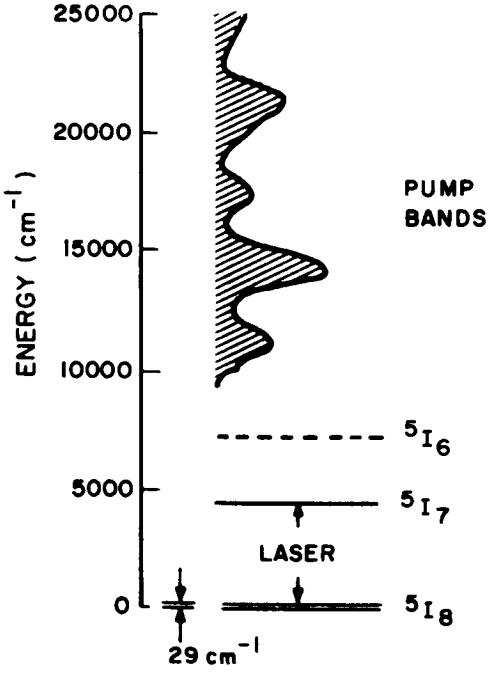


Fig. 11. Energy levels of $CaF_2:Dy^{2+}$.

at several thousand degrees Kelvin. Both tungsten lamp (3400° K) and solar (6000° K) pumping have been successfully used.

Laser emission accompanies a de-excitation transition between the triply degenerate $^5I_7$ and $^5I_8$ states, the latter being on the order of 30 cm^{-1} above the ground state. This emission occurs at temperatures below 77° K. At a temperature of 70° K the energy of the terminal state is 0.61 kT, indicating that the system functions as a three-level laser. The fluorescent lifetime of the $^5I_7$ state is 3 ms and the fluorescent linewidth is about 0.3 cm^{-1}. This ensures that CW laser operation can be achieved with moderate pump power levels. (The "long" lifetime ensures that the pump source can excite atoms at the necessary rate, and the narrow linewidth contributes to requiring a small value for the critical population inversion.) The narrowness of the linewidth also allows single mode operation. (The mode spacing is 0.15 cm^{-1} for a laser 1 inch long.) The narrowness is important if a magnetic modulation technique is to be possible.

The wavelength of the laser emission is 2.36 μm. The temporal characteristics of the output depend on the type of reflectors used at the crystal ends. If flat-end crystals are used, the output consists of a series of 2-μs pulses. The amplitudes of the pulses are relatively independent of pump power, but the pulse frequency is strongly dependent on the pump power. Frequencies range from a few thousand pulses per second at threshold to 10^5 p/s at high levels (i.e., 1 kW into a pump lamp used in an elliptical cavity). For spherical-end crystals, spikes are not observed. The output consists of a large CW component with small ripple content (typically 10 to 30 per cent).

As described in Sec. V. B below, a $CaF_2:Dy^{2+}$ laser has been operated under solar-pumped conditions. Inhomogeneous magnetic modulation was used to transmit audio information, and operating characteristics were found to be well-described by the above theory. Pertinent data for the specific crystal used in this experiment are:

Crystal dimensions: 1/8 inch diameter, 1 inch long

Threshold power $\approx$ 20 W solar energy

Power output = 40 mW

Ripple frequency = 6×10^4 c/s (at pump level attainable with solar collector used)

b. Properties of YAG:Nd^{3+}[1]

Yttrium aluminum garnet ($Y_3Al_5O_{12}$) has proven to be an ideal host for the rare earth series ions. In the crystal under consideration the Nd^{3+} ions enter substitutionally into the $Y_3Al_5O_{12}$ lattice at the yttrium sites. Since the yttrium ion is trivalent, the substitution of Nd^{3+} does not necessitate the use of any charge compensation.

The energy level diagram for this system is shown in Fig. 12. In this crystal, pump energy causes transitions between the 4f-5d electronic energy states of Nd^{3+}. In the trivalent rare earths the 4f-5d transitions typically correspond to the absorption of short-wavelength radiation. For YAG:Nd^{3+}, the pump bands lie in the ranges 0.50 μm - 0.54 μm, 0.57 μm - 0.60 μm, and 0.73 μm - 0.83 μm. Thus, the pump bands are quite narrow. This implies that efficient use cannot be made of broad pump sources such as a high-temperature blackbody or even a broad non-blackbody source. However, even though inefficient, it is still possible to pump YAG:Nd^{3+} with broad pump sources and, in fact, crystals have been pumped with a tungsten lamp and the Sun.

Infrared fluorescence in YAG:Nd^{3+} occurs between the $^4F_{3/2}$ level and the members of the ^4I multiplet. Laser action occurs in the transition $^4F_{3/2} \rightarrow {}^4I_{11/2}$. The $^4I_{11/2}$ state is 2000 cm^{-1} above the lowest member of the

[1] R. J. Pressley et al., "Solid State Laser Explorations," Interim Technical Report No. 1, Contract No. AF33(615)2645, p. 1, RCA, Princeton, N. J.

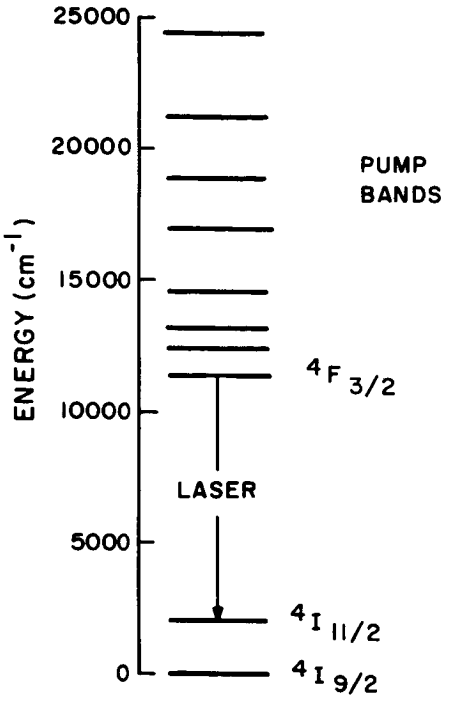


Fig. 12. Energy levels of YAG:Nd^{3+}.

^4I multiplet ($^4I_{9/2}$). This means that at a temperature of 300° K the energy of the terminal laser state is equal to 9.5 kT (above ground state). Thus, YAG:Nd^{3+} is a four-level laser system even at T = 300° K. With E = 9.5 kT the system approaches the ideal of terminal state population limited by the non-radiative lifetime of the terminal state. The fluorescent lifetime of the laser transition at 300° K is 210 μs and the linewidth is approximately 2 cm^{-1}. The lifetime is dependent on the dopant concentration.

The wavelength of the laser radiation for this crystal is 1.06 μm, As in the case of the CaF_2:Dy^{2+} laser system, the temporal characteristics of the laser emission depend upon the type of cavity reflectors used. When flat-end crystals are used, the output is a series of spikes. When spherical-end crystals are used, the output is CW with superimposed damped spiking.

As described in Sec. V. B below, a YAG:Nd^{3+} laser has been operated under solar-pumped conditions. Wideband electro-optic modulation was used to transmit a television signal. Pertinent data for the specific crystal used in this experiment are:

Crystal dimensions: 0. 094 inch dimater, 1. 25 inches long

Threshold power $\approx$ 100 W solar energy

Power output = 90 mW

Ripple frequency $\approx 10^4 - 10^5$ Hz (depending on pumping level)

c. Properties of YAG:Nd^{3+}-Cr^{3+} [1,2]

In this double-doped laser system, both dopants enter substitutionally into the host lattice of YAG. As in YAG:Nd^{3+}, the Nd^{3+} ions enter directly into the yttrium sites. The Cr^{3+} ions enter into aluminum sites which have an eightfold coordination. (There exist Al^{3+} sites with eightfold and fourfold coordination. Cr^{3+} substitutes into an eightfold coordinated site due to crystal field stabilization at these sites.) In both cases it is not necessary to add any charge compensation. Fig. 13 illustrates the energy level scheme for this laser system.

The purpose of introducing Cr^{3+} into the YAG laser system is to provide a mechanism for energy exchange from Cr^{3+} excited levels to Nd^{3+} levels. Such an exchange allows the use of the pump bands of Cr^{3+} for inducing the laser transition between the $^4F_{3/2}$ and $^4I_{11/2}$ levels of Nd^{3+}. Pump radiation causes the Cr^{3+} ions to be excited. The ions then de-excite via a non-radiative transition to the metastable ^2E state of Cr^{3+}. An energy exchange then takes place between this state and the Nd^{3+} ions which places Nd^{3+} ions in the $^4F_{3/2}$ state, the state from which the laser transition originates. It is thus possible to obtain laser output by pumping the system with radiation which is absorbed by Cr^{3+}.

---

[1] R. J. Pressley et al., "Solid State Laser Explorations," Interim Technical Report No. 1, Contract AF33(615)2645, p. 1, RCA, Princeton, N. J.

[2] Z. J. Kiss and R. C. Duncan, "Cross-Pumped Cr^{3+}-Nd^{3+}:YAG Laser System," RCA Internal Report.

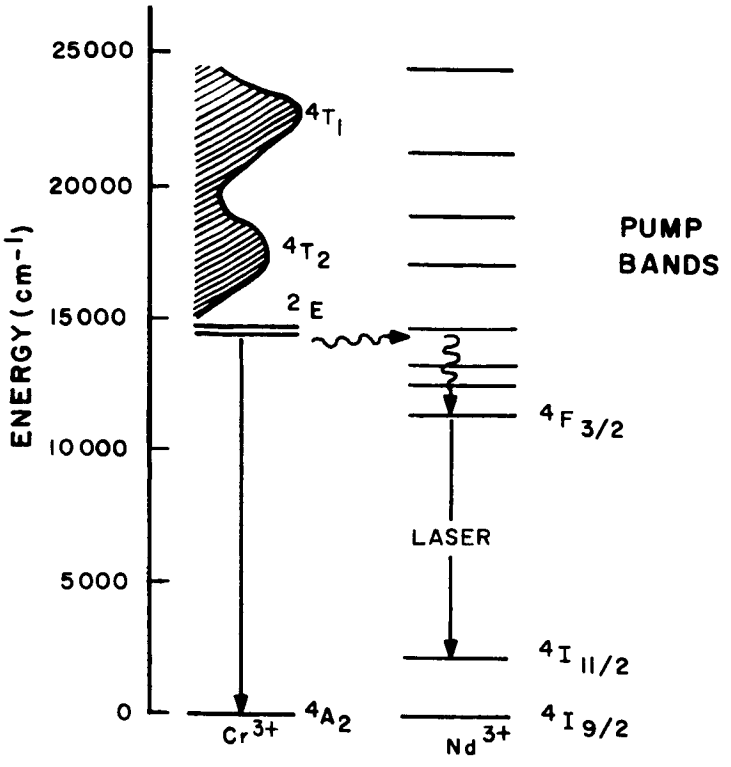


Fig. 13. Energy levels for YAG:Nd^{3+}-Cr^{3+}.

As a result of this energy exchange process, the pump spectrum for YAG:Nd^{3+}-Cr^{3+} consists of the bands 0.40 μm - 0.50 μm, 0.52 μm - 0.66 μm, and 0.73 μm - 0.83 μm. Comparison of these bands with those of singly-doped YAG shows that for tungsten lamp pumping (3400° K) little increase in pumping efficiency is expected since the lamp spectrum does not contain appreciable output in the Cr^{3+} pump bands. As the color temperature of the source increases, however, appreciable output occurs in these bands, and the efficiency of doubly-doped YAG increases above that of the singly-doped material. Thus for solar pumping (6000° K) more laser output is expected for YAG:Nd^{3+}-Cr^{3+} than for YAG:Nd^{3+} if samples of crystals of comparable quality are used. High pressure mercury lamp pumping may also be used for pumping doubly-doped YAG since this source contains appreciable output in the Cr^{3+} pump bands.

It has been observed that the efficiency of the energy transfer process is very dependent upon the concentration of Cr^{3+} in the crystal. The time required for the transfer to take place decreases with increasing Cr^{3+} concentration. This means the efficiency increases with increasing concentration, since if a transfer time which is an appreciable fraction of the lifetime of the metastable ^2E state of Cr^{3+} exists, fluorescence losses deplete the population of this state and thus diminish the number of possible energy transfers. It has been possible to grow YAG:Nd^{3+}-Cr^{3+} with Cr^{3+} concentrations such that the efficiency of pumping in the Cr^{3+} pump bands approaches 100 per cent. It has also been demonstrated that laser output can be achieved by pumping only in the Cr^{3+} bands.

With the exception of the special properties contributed by the addition of Cr^{3+}, the temporal characteristics for CW operation of the 1.06-μm output of YAG:Nd^{3+}-Cr^{3+} are similar to those of the singly-doped material.

Solar-pumped operation of a YAG:Nd^{3+}-Cr^{3+} laser has been achieved, but the output of the sample used did not exhibit "good" CW characteristics. This is due to the fact that the crystal used was not of optimum optical quality. The laser threshold in this case was again on the order of 80-100 W of solar energy.

2. Optimization of Crystal Size

For a wideband communication system, an optimum laser system can be defined as one which furnishes the required power output while operating at maximum efficiency. In this regard, it is shown in Appendix G that there exists an optimum size for the laser crystal used in such a system. The analysis of Appendix G indicates that optimum pumping efficiency is achieved when the crystal diameter equals the size of the solar image and when the length is chosen according to the following criteria:

(a) Since power absorbed (and hence output power) is proportional to crystal length, a length should be chosen such that an appreciable amount of available pump power is absorbed.

(b) Since scattering losses increase with length, the crystal must not be of such a length that losses offset the effects of greater absorption with increasing length.

Experiments performed with the three crystals described above have indicated that cylindrical crystals with a length 20 to 30 times the crystal radius satisfy the above criteria reasonably well.

B. SOLAR COLLECTOR

The area of the solar energy collector required to operate a laser is determined by:

(1) The laser threshold

(2) The efficiency of the collection system

(3) The efficiency of the laser

(4) The magnitude of the solar constant at the position of operation

(5) The required laser power output

Examination of these parameters and their interrelationship establishes the solar collector size required to obtain a specified laser output.

The laser threshold sets an absolute minimum collector size for a given laser crystal. Power can be obtained from the laser if, and only if, the collector used is larger than this minimum size.

The threshold power of a four-level cylindrical laser is given by Eq. G-1, which may be written in the form

$$P_{th} = \frac{4\pi^2 h\nu^3 \Delta\nu \eta^3 V}{c_o^3 t_p} \tag{23}$$

where h = Planck's constant (6.62×10^{27} erg · s)

ν = laser radiation frequency (Hz)

$\Delta\nu$ = laser radiation linewidth (cm^{-1})

η = index of refraction of laser material

t_p = photon lifetime (s)

c_o = velocity of light in free space (cm/s)

V = laser crystal volume (cm^3)

or, since $V = \pi r^2 \ell$ for a cylindrical crystal (where r = crystal radius and ℓ = crystal length) and laser radiation wavelength $\lambda_o = c_o/\nu$,

$$P_{th} = \frac{4\pi^3 h c_o^2 \eta^2 r^2 \Delta\nu\alpha}{\lambda_o^3} \tag{24}$$

The photon lifetime t_p is given by

$$t_p = \frac{\ell\eta}{\alpha c_o} \tag{25}$$

where ℓ = length of laser (cm)

α = loss per pass

if $\alpha << 1$.

Using these equations, the threshold powers for the YAG:Nd^{3+} and YAG:Nd^{3+}-Cr^{3+} can be calculated. As evident from Eq. 17, these calculations are dependent on crystal geometry and loss properties. The cyclindrical crystals used in the experiments described in this report have the following properties:

ℓ = 1.0 in

2r = 0.125 in

α = 0.05

$\Delta\nu$ = 2 cm^{-1}

η = 1.4

λ_o = 1.06 x 10^{-4} cm

Thus, with these values, the threshold power for both YAG:Nd^{3+} and YAG:Nd^{3+}-Cr^{3+} is

$$P_{th} = \frac{4\pi^3(6.62 \times 10^{-27})(3 \times 10^{10})^2 (1.96)(0.317)^2(2)(5 \times 10^{-2})}{(1.06 \times 10^{-4})^3}$$

$$= 1.22 \times 10^7 \text{ erg/s} = 1.22 \text{ W}$$

If the laser is pumped by radiation of wavelength λ (rather than λ_o), the actual pump power required to reach threshold is λ_o/λ times the value of P_{th} calculated above. For example, for pump radiation of wavelength λ = 5900 Å, the actual threshold pump power P'_{th} is

$$P'_{th} = \frac{\lambda_o}{\lambda} P_{th}$$

$$= \frac{10600 \text{ Å}}{5900 \text{ Å}} 1.22 = 2.19 \text{ W}$$

When solar pumping is used, pump radiation within a broad spectrum of wavelengths is supplied to the crystal. For example, the solar spectral irradiance at a distance of 93×10^6 mi from the Sun (mean Earth-Sun separation) appears as shown in Fig. 14.[1] In addition to this broad spectral distribution, laser crystal

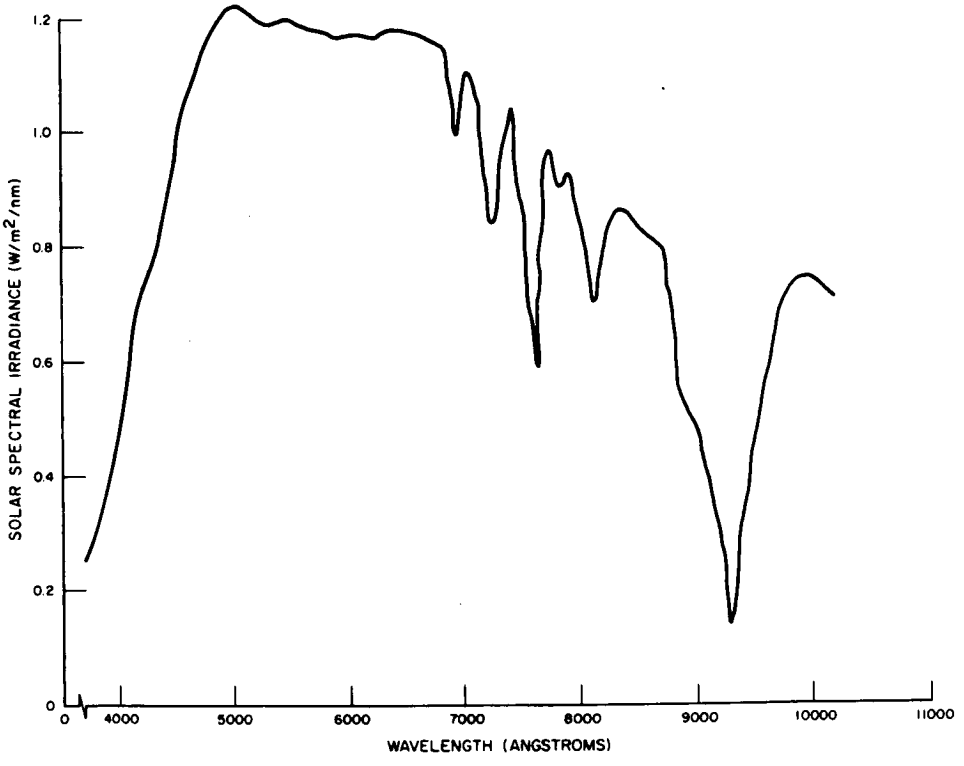


Fig. 14. Solar emission spectrum.

[1] Constructed from data in "Handbook of Geophysics," USAF, p. 16-16, The Macmillan Co., New York, 1960.

wavelength absorption is wavelength dependent; i. e. , absorption within the crystal does not occur only for pump radiation of a single wavelength. Information relating to this effect is given in Figs. 15 and 16[1], where the absorption spectra for YAG:Nd^{3+} and YAG:Nd^{3+} -Cr^{3+} are shown.

The information available in Figs. 14, 15, and 16 is useful for determining how much of the available solar power is actually pumped into the laser crystal as useful pump power. In this regard, it is assumed that any compensation for power requirements resulting from the use of solar pumping (as contrasted to a narrow spectral pump source) can be accounted for by the collection efficiency ζ_1, which is defined as

$$\zeta_1 \triangleq \frac{(\lambda/\lambda_o)(\text{Power in pump bands coupled to laser crystals})}{\text{Total collected solar power}} \qquad (26)$$

$$= \frac{\text{Effective pump power coupled to crystal}}{\text{Total collected solar power}}$$

where

(λ/λ_o) = wavelength conversion factor for use with solar pump source (i. e. , the average conversion factor over all the wavelengths indicated by the solar emission spectrum shown in Fig. 14)

There is still one other phenomenon to be considered before a value of minimum collector size can be determined; that is, not all the power in the pump bands coupled to the crystal excites atoms to the upper laser state. The effect of this process is accounted for by the pump efficiency ζ_2, which is defined as the ratio of the number of crystal atoms in the upper laser state to the number of photons which excite crystal atoms to that state. This efficiency is a function of dopant concentration, optical properties of the laser material, and the wavelength spectrum of the pump source.

---

[1] R. J. Pressley et al. , "Solid State Laser Explorations," Interim Technical Report No. 1, Contract AF33(615)2645, pp. 2-3, May 1965.

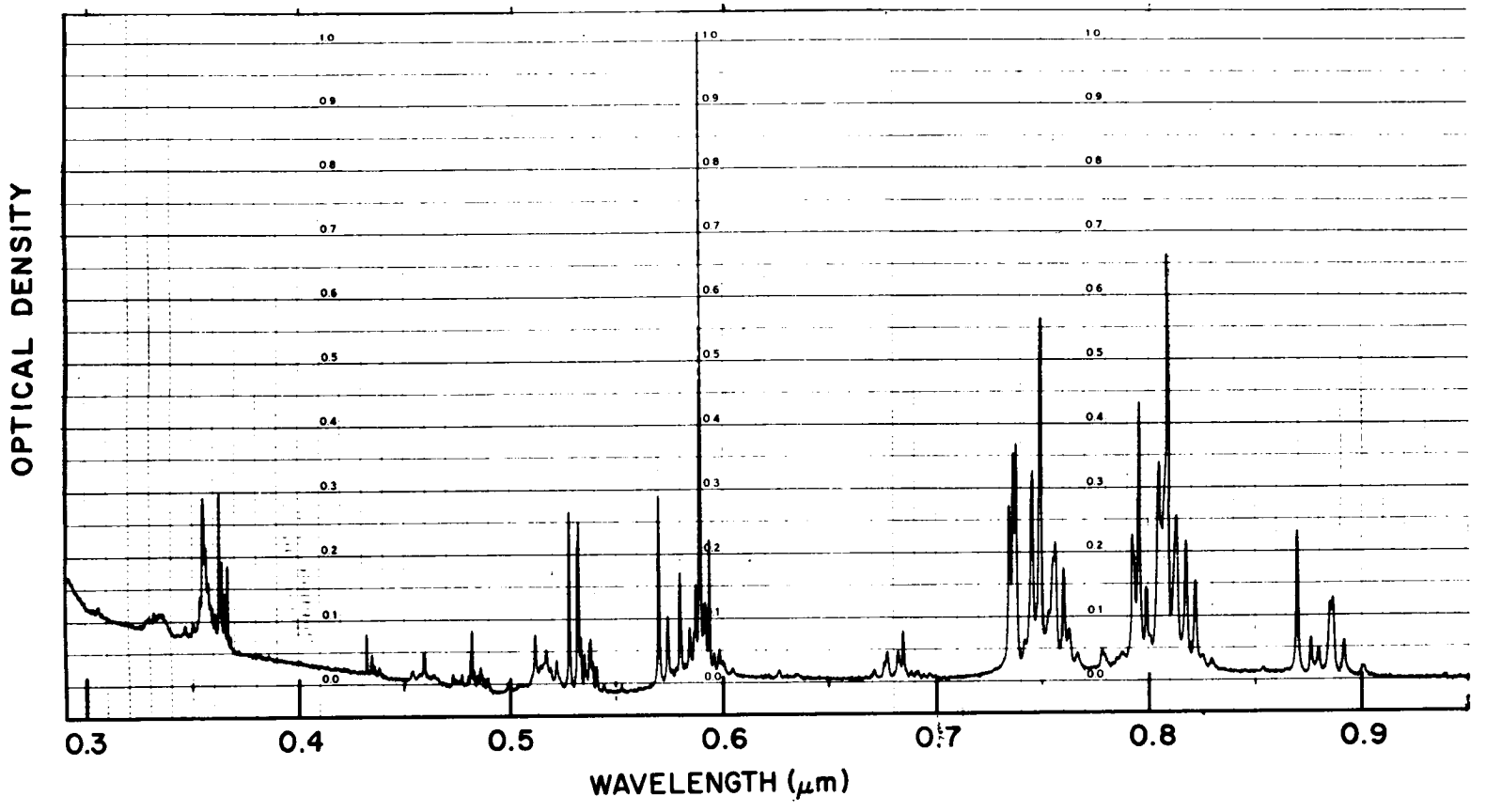


Fig. 15. YAG:Nd^{3+} absorption spectrum.

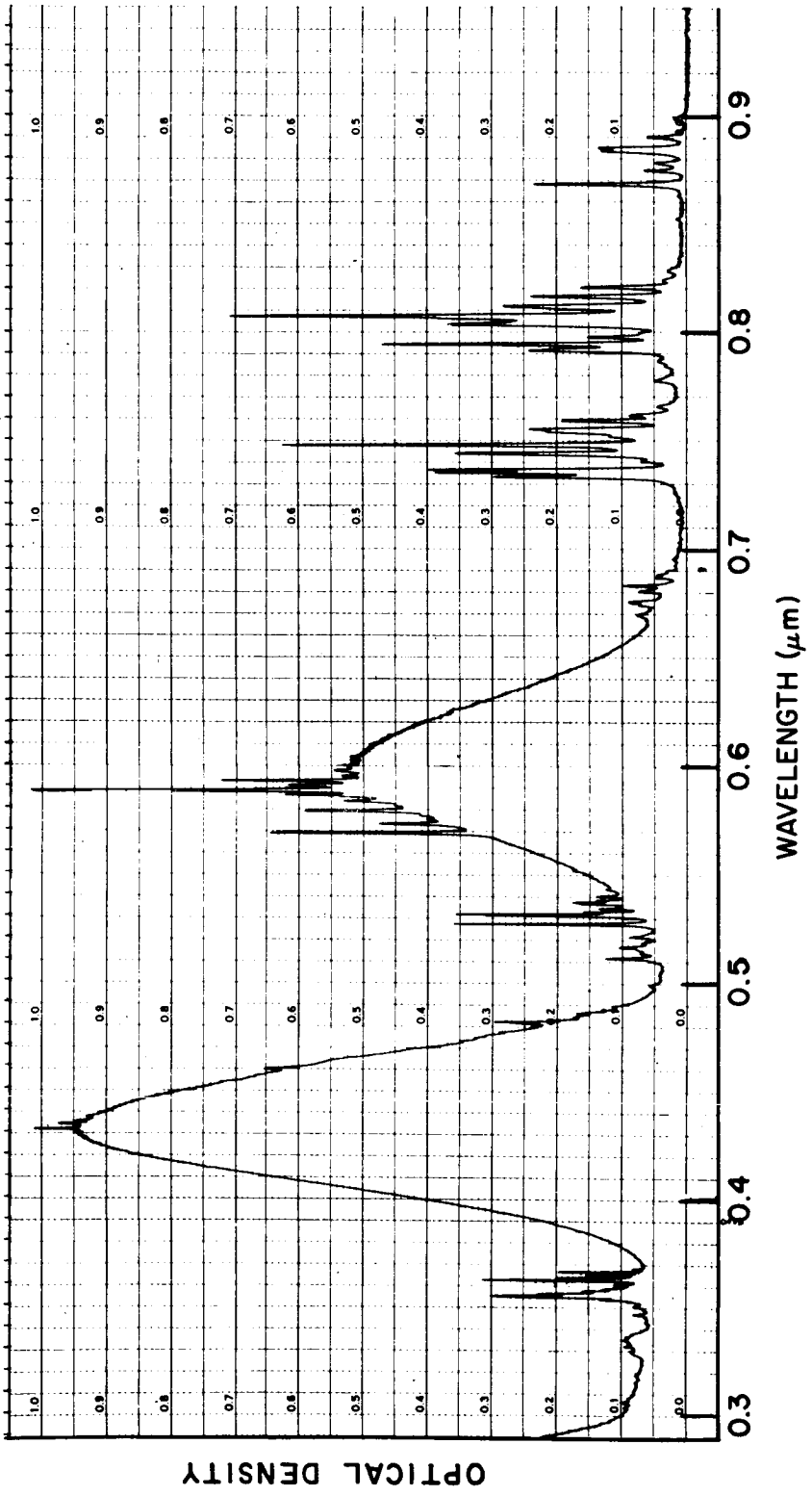


Fig. 16. YAG:Nd^{3+}-Cr^{3+} absorption spectrum.

With the above considerations the total efficiency of a solar-pumped laser can now be determined. Thus, the collected power useful in exciting atoms to the upper laser state is given by $\zeta_1 \zeta_2$ times the total power in the pump bands collected. This collected pump power is some fraction ζ_o of the total solar power collected. Hence, to reach threshold, solar power equal to $P_{th}/\zeta_o \zeta_1 \zeta_2$ must be collected. Also, in terms of the solar power collected, the laser power output P_{out} is given by

$$P_{out} = \zeta_o \zeta_1 \zeta_2 \left[\text{(total solar power collected)} - \text{(solar power required for threshold)}\right] \tag{27}$$

(The assumption is then made here that all atoms placed in the upper laser state contribute to the laser output.) The overall (total) efficiency ζ_t of the laser is found as

$$\zeta_t = \frac{P_{out}}{\text{Total solar power collected}} = \zeta_o \zeta_1 \zeta_2 \left(1 - \frac{\text{solar power required for threshold}}{\text{total solar power collected}}\right) \tag{28}$$

Now, to determine collector diameter for a given power output for YAG lasers, the amount of power in the solar spectrum which can be absorbed as pump power by the laser must be determined. To do this, consider the equation for the pump power P_a absorbed by a laser crystal:

$$P_a = A_c \int_{\lambda_1}^{\lambda_2} P_o(\lambda) \left[1 - e^{-\sigma(\lambda) d}\right] d\lambda \tag{29}$$

where

A_c = collector area (m^2)

$P_o(\lambda)$ = solar intensity at wavelength λ ($W/m^2 \cdot nm$)

$\sigma(\lambda)$ = crystal absorption coefficient at wavelength λ (m^{-1})

d = path length for radiation (m)

λ_1, λ_2 = wavelength at edges of pump source spectrum

This integral can be evaluated numerically using the plots of Figs. 14, 15, and 16. For the two types of YAG lasers, the results are:

$$P_a\,(YAG{:}Nd^{3+}) = A_c \sum_{\lambda = 4000\,\text{Å}}^{\lambda = 9000\,\text{Å}} P_o(\lambda)\left[1 - e^{-\sigma(\lambda)d}\right]\Delta\lambda = 273.6\,A_c \text{ (watts)}$$

$$P_a\,(YAG{:}Nd^{3+}\text{-}Cr^{3+}) = 400\,A_c \text{ (watts)}$$

The fraction ζ_o may now be determined immediately since the integrated solar intensity (at Earth) is known to be 1322 W/m^2. Hence

$$\zeta_o(YAG{:}Nd^{3+}) = \frac{273.6A_c}{1322\,A_c} = 0.206$$

$$\zeta_o(YAG{:}Nd^{3+}\text{-}Cr^{3+}) = \frac{400A_c}{1322\,A_c} = 0.302$$

The collector area required to achieve laser threshold may now be found if ζ_1 and ζ_2 are known. Inspection of Eq. 28 shows that for high pump levels,

$$\frac{\text{Solar power required for threshold}}{\text{Total solar power collected}} << 1$$

and, therefore,

$$\zeta_t \approx \zeta_o\zeta_1\zeta_2 \tag{31}$$

Thus, with $\zeta_o = 0.206$ for the YAG:Nd^{3+} crystal, the product $\zeta_1 \zeta_2$ must equal 0.097 if the laser operates at an overall efficiency ζ_t of 0.02. The corresponding requirement for a 0.02-efficiency YAG:Nd^{3+}-Cr^{3+} laser is $\zeta_1 \zeta_2 = 0.066$ (when $\zeta_o = 0.302$).

As a typical example, assume $\zeta_1 = 0.5$ and $\zeta_2 = 0.194$ for a YAG:Nd^{3+} laser. The minimum collector area A_{min} is then found as

$$A_{c_{min}} (YAG:Nd^{3+}) = \frac{1.22\ W}{\zeta_o \zeta_1 \zeta_2 \left(1322 \frac{w}{m^2}\right)} = 0.046\ m^2$$

The corresponding minimum diameter is

$$D_{c_{min}} (YAG:Nd^{3+}) = 0.242\ m$$

A similar calculation for YAG:Nd^{3+}-Cr^{3+} yields the same result. This implies that double-doped YAG can be pumped more efficiently by the solar spectrum than can singly-doped YAG, since the required efficiency product $\zeta_1 \zeta_2$ is smaller for YAG:Nd^{3+}-Cr^{3+} when both lasers operate at 0.02 (or 2 per cent) overall efficiency at high pump levels. In fact, if $\zeta_1 \zeta_2 = 0.097$ for both lasers, the ratio of overall efficiencies γ at high pump levels approaches

$$\gamma = \frac{\zeta_t (YAG:Nd^{3+}\text{-}Cr^{3+})}{\zeta_t (YAG:Nd^{3+})} = \frac{\zeta_o (YAG:Nd^{3+}\text{-}Cr^{3+})\ \zeta_1 \zeta_2}{\zeta_o (YAG:Nd^{3+})\ \zeta_1 \zeta_2} \qquad (32)$$

$$= \frac{0.302}{0.206} = 1.465$$

This value implies that for identical operating characteristics (i.e., $\zeta_1 \zeta_2$ = constant), the doubly-doped YAG is about 1.5 times as efficient as singly-doped YAG.

From Eq. 27, a plot of power output vs collector diameter can now be generated. Thus, from Eq. 27,

$$P_{out} = \zeta_o \zeta_1 \zeta_2 \, (1322 \, A_c - 1322 \, A_{c_{min}}) \text{ watts} \tag{33}$$

for operation at 93×10^6 miles range from the Sun. In particular, for $\zeta_o = 0.206$, $\zeta_1 = 0.5$, and $\zeta_2 = 0.194$,

$$P_{out} = 26.55 \, (A_c - A_{c_{min}})$$

$$= 26.55 \, A_c - 1.22 \text{ watts} \tag{34}$$

In terms of collector diameter, this becomes

$$P_{out} = 20.8 \, D_c^{\,2} - 1.22 \text{ watts} \tag{35}$$

Equation 33 is shown plotted in Fig. 17 for $\zeta_2 = 0.194$ and $\zeta_1 = 0.3$, 0.5, 0.7, corresponding to overall efficiences of 0.027, 0.02, and 0.012, respectively, thereby illustrating the effect of varying the collection efficiency, ζ_1.

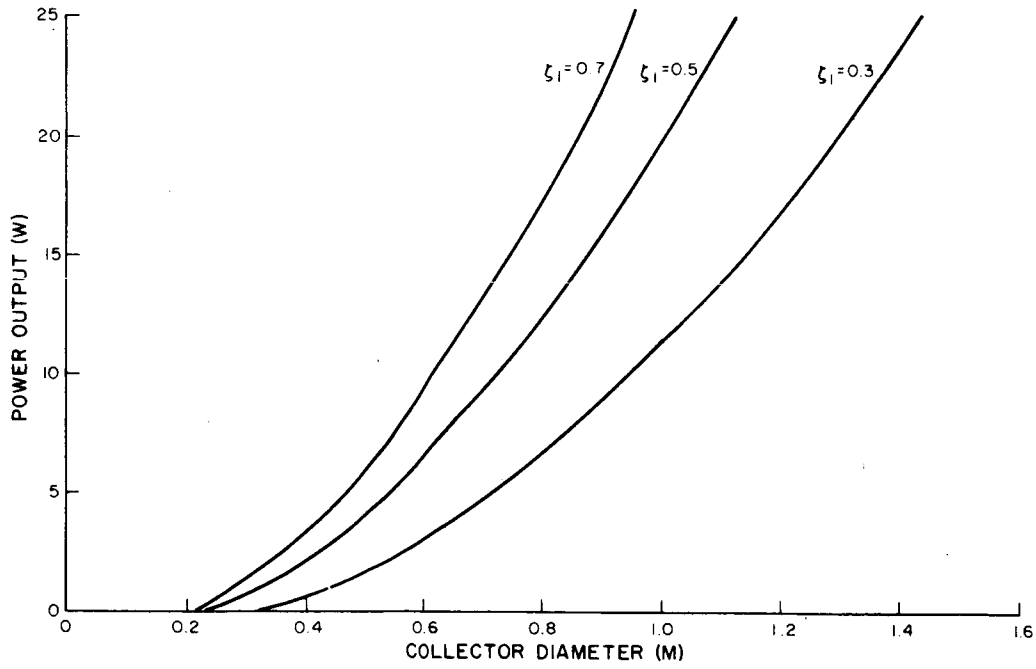


Fig. 17. Laser power output vs collector mirror diameter for YAG:Nd^{3+} laser positioned at Earth with pump efficiency = 0.194 and collection efficiency = 0.3, 0.5, 0.7.

A plot of output power vs collector diameter for a YAG:Nd^{3+} - Cr^{3+} laser, operating under the same conditions as the singly-doped YAG laser, is shown in Fig. 18. The curves in Fig. 18 are obtained by multiplying the coefficient of D_c^2 in Eq. 35 by the relative efficiency factor γ (see Eq. 32).

Plots similar to those in Figs. 17 and 18 can also be obtained for operation at positions other than Earth. For example, consider operation in the vicinity of Mars. Here, the integrated solar irradiance is 555 W/m^2 instead of 1322 W/m^2 (the value at Earth). In this case, Eq. 33 is changed to

$$P_{out} = \zeta_o \zeta_1 \zeta_2 (555 A_c - 555 A_{c_{min}}) \tag{36}$$

for the conditions $\zeta_o = 0.206$, $\zeta_1 = 0.5$, and $\zeta_2 = 0.194$. Here $A_{c_{min}} = 0.1\ m^2$ and $D_{c_{min}} = 0.374$ m.

Plots of Eq. 36 for YAG:Nd^{3+} and YAG:Nd^{3+}-Cr^{3+} laser operation in the vicinity of Mars with $\zeta_o = 0.206$, $\zeta_1 = 0.3, 0.5, 0.7$, and $\zeta_2 = 0.194$ are shown in Figs. 19 and 20.

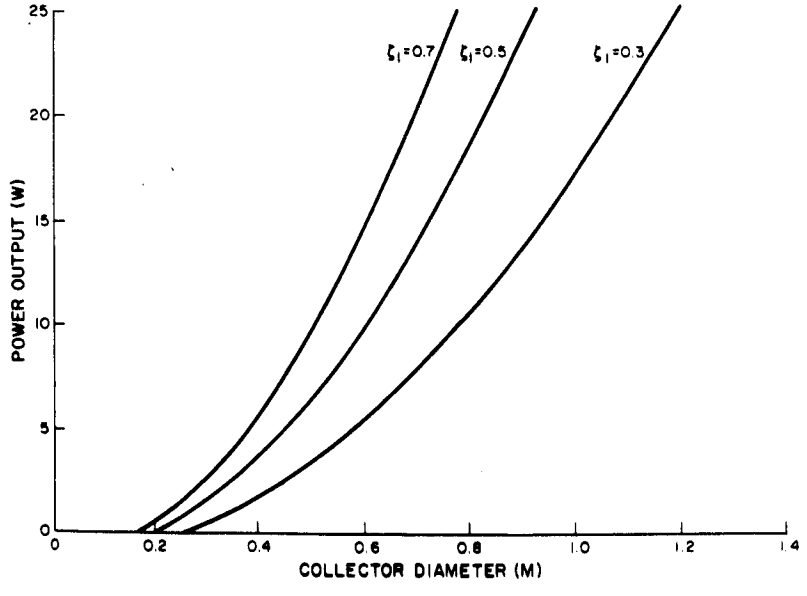


Fig. 18. Laser power output vs collector mirror diameter for YAG:Nd^{3+} - Cr^{3+} laser positioned at Earth with pump efficiency = 0.194 and collection efficiency = 0.3, 0.5, 0.7.

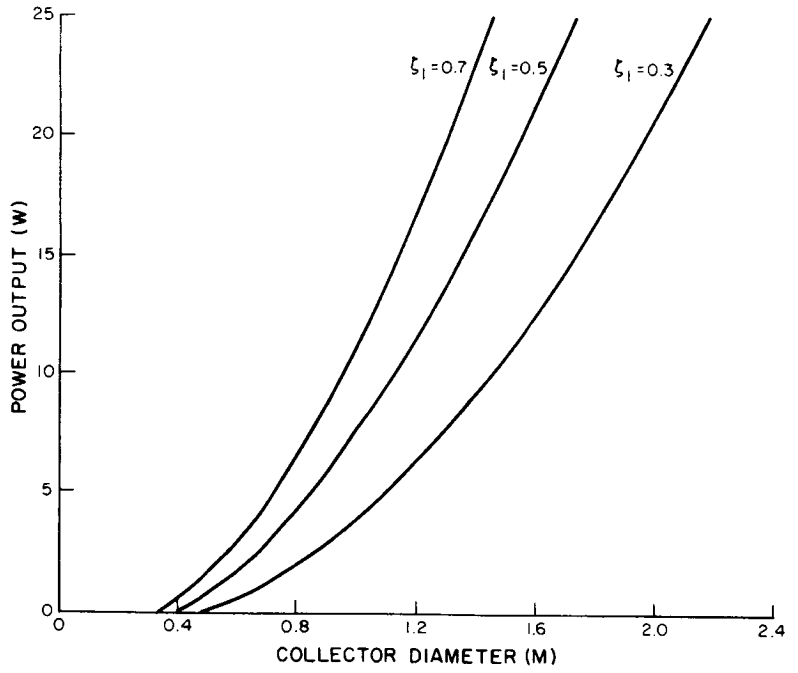


Fig. 19. Laser power output vs collector mirror diameter for YAG:Nd^{3+} laser positioned at Mars with pump efficiency = 0.194 and collection efficiency = 0.3, 0.5, 0.7.

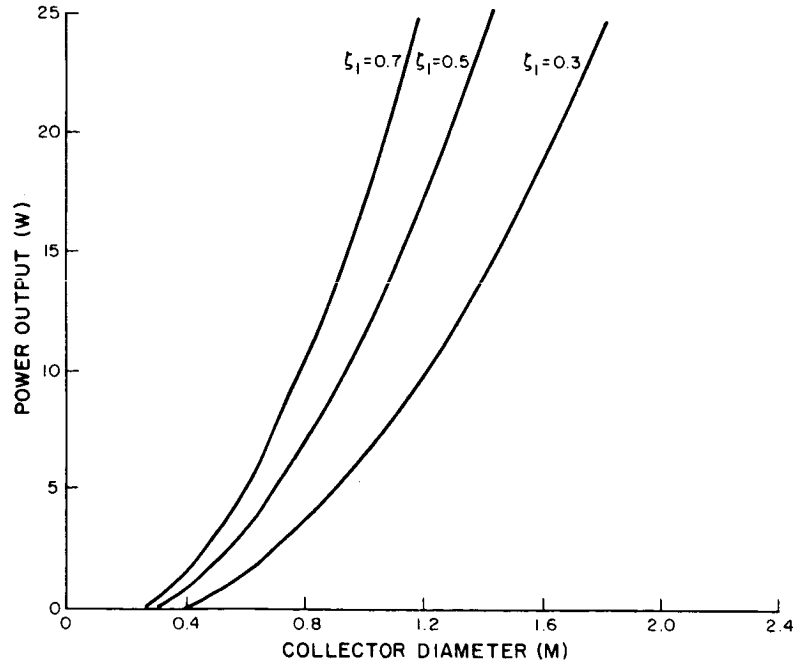


Fig. 20. Laser power output vs collector mirror diameter for YAG:Nd^{3+}-Cr^{3+} laser positioned at Mars with pump efficiency = 0.194 and collection efficiency = 0.3, 0.5, 0.7.

C. THERMAL ANALYSIS

1. Introduction

This analysis determines the thermal profile of the laser crystal (i.e., its temperature distribution) and establishes cooling requirements. A detailed analysis of the optimization of cooling fins for dissipation of heat to space is made also. Consideration is given to the best method for transferring heat from the laser crystal to the radiator fins and a detailed analysis is made of the preferred method.

A critical problem in the realization of a solid-state laser oscillator is cooling the laser crystal so as to maintain it at a suitable operating temperature. Heretofore, this temperature was quite low and liquefied gases were needed for cooling. Now, however, the YAG:Nd^{3+} (and YAG:Nd^{3+}-Cr^{3+}) laser operates at room temperature (300°K), which considerably simplifies the cooling problem. In any case, because of the relatively poor thermal conductivity of the laser crystal material, there is a limitation on laser crystal size and output due to internal temperature rises above the surface and coolant temperatures. In addition, the internal temperature gradients in the laser crystal cause thermal stresses that may be significant.

When a laser is used in a spaceborne communication system, the heat removed from the laser crystal must eventually be radiated into space. Therefore, an optimum design for space radiators is important, and it is also quite important to evaluate, from the system point of view, the radiator area, size and weight required as a function of laser output.

The remaining thermal problem in a space-vehicle laser system is the method used to effect the transfer of heat from the laser crystal surface to the space radiator. This may be accomplished by conduction in a solid (the simplest method, but poor in performance, as well as requiring very heavy implements);

by convection of a circulating liquid (a complex method, requiring a pump and good plumbing, but excellent in performance); or by distillation (excellent performance and light weight, but a technique must be developed for liquid return in a zero-g environment). These considerations are discussed further in the following sections.

2. Laser Crystal Temperature Distribution and Cooling

Of the solar energy absorbed by the laser crystal for "pumping" (i.e., maintenance of the inverted, or negative, temperature state), most must ultimately be removed in the form of heat for continuous operation of the laser, that is, for the maintenance of steady conditions. Accordingly, the temperature distribution of the laser element is analyzed for the geometry assumed — namely, a relatively long right circular cylinder. The following assumptions are made for this analysis:

a. The thermal properties of the laser crystal are temperature independent.

b. The heat to be dissipated is uniformly distributed in the laser crystal.

c. Steady (time-independent) conditions prevail.

d. The laser crystal is uniformly cooled on its lateral surfaces.

In view of these assumptions, the laser crystal temperature distribution is axisymmetric and is a function only of the radial distance from the crystal axis.

The following notation is adopted:

r = radial distance

a = laser crystal radius

K = crystal thermal conductivity

h = thermal surface conductance to crystal surroundings

q = rate of radial heat transport per unit area

q_0 = specific rate of internal heat dissipation (per unit volume)

t = empirical temperature (coolant temperature taken as zero)

Subscripts a and 0 refer to values of quantities at the crystal surface and axis, respectively. Other special symbols will be introduced and defined as needed.

The steady, radial temperature distribution in a conducting, heat-dissipating solid — subject to the assumptions made above — is governed by the relations:

$$q = -K \frac{dt}{dr} \qquad \text{(Fourier's Law)} \tag{37}$$

$$\frac{d}{dr}(rq) = rq_0 \qquad \text{(Energy Relationship)} \tag{38}$$

and is subject to the boundary conditions:

$$q = 0 \text{ at } r = 0 \text{ (Symmetry)}$$
$$ht_a = q_a = -K\left(\frac{dt}{dr}\right)_a \text{ (Newton's Law of Cooling).} \tag{39}$$

Eq. 38 may be integrated from $r = 0$ to $r = r'$, where $0 \leq r' \leq a$; i.e.,

$$rq = \int_0^{r'} rq_0 \, dr + C_1 \tag{40}$$

Since q_0 is a constant, it immediately follows that:

$$rq = q_0 \frac{r'^2}{2} + C_1 \ . \tag{41}$$

By replacing r' by r and using the information that $q = 0$ at $r = 0$, it is found that:

$$rq = q_0 \frac{r^2}{2} \text{ or } q = \frac{1}{2} q_0 r \ . \tag{42}$$

Introducing Eq. 42 in Eq. 37 yields:

$$K\, dt/dr = q_0\, r/2 \quad , \tag{43}$$

and integrating gives

$$t - t_a = \frac{q_0 a^2}{4K} (1 - \xi^2) \qquad \xi = r/a \tag{44}$$

From Eq. 42, $q_a = q_0\, a/2$; hence, from the second boundary condition given, $t_a = q_0\, a/2h$, and so:

$$t = q_0\, a/2h + (q_0\, a^2/4K)\,(1 - \xi^2) \tag{45}$$

or

$$t = q_0\, a^2/4K \left[\left(\frac{2K}{ha} + 1 \right) - \xi^2 \right]. \tag{46}$$

The parameter ha/K is conventionally known as Biot's modulus, denoted Bi. A dimensionless temperature, $t^* = t/(q_0\, a^2/4K)$, may conveniently be defined and then Eq. 46 becomes:

$$t^* = \left(\frac{2}{Bi} + 1 \right) - \xi^2 \quad . \tag{47}$$

The maximum temperature occurs at the crystal axis ($\xi = 0$) and is

$$t_0^* = \frac{2}{Bi} + 1 \quad . \tag{48}$$

The crystal surface temperature ($\xi = 1$) is

$$t_a^* = 2/Bi \tag{49}$$

The ratio $(t_0^* - t_a^*)/t_a^* = Bi/2$ is interesting, and expresses the relative importance of the internal temperature rise (depending on crystal dimensions, heat dissipation rate, and crystal properties) and the surface temperature rise (depending on coolant conditions). Since Bi = ha/K, for specific values of h and K, Bi is proportional to a. This shows that as the crystal diameter increases, the internal temperature rise becomes more and more significant.

In Sec. C. 3 it is necessary to relate the heat to be dissipated by the cooling systems to the laser output. Depending on the laser crystal material, frequency distribution of the "pumping" radiation, and the laser crystal temperature, there will be a certain threshold "pump" rate that will maintain the inverted population, or negative temperature state, required to achieve laser oscillation. For the YAG:Nd laser, at the present time, this threshold power for solar pumping, where solar radiation is accepted over the range of 4000 to 9000 Å is to be regarded as an empirical quantity. This quantity, denoted here by q_{th}, is either ultimately dissipated as heat (including spontaneous radiative heat loss to the surrounding environment) resulting from spontaneous radiative and non-radiative decay processes from the excited states, or contributes to that part of the "pump" radiation absorbed by the laser crystal which plays no role in the laser action.

When the laser is producing power, additional "pump" energy is required to provide power for the stimulated radiation. We define the laser quantum efficiency, ζ_2, as the ratio of laser output to this excess of "pump" power over the threshold pump power. If we denote the laser specific power (output per unit volume) by p, the heat to be removed from the laser crystal is:

$$q_0 = q_{th} + \frac{(1-\zeta_2)}{\zeta_2}\, p \tag{50}$$

or, in dimensionless form,

$$q_0^* = \frac{q_0}{q_{th}} = 1 + \left(\frac{1}{\zeta_2} - 1\right) p^* \tag{51}$$

where $q_0^* = q_0/q_{th}$, and $p^* = p/q_{th}$. (Eq. 51 will be used later.)

Eq. 50 is plotted in Fig. 21 for several values of laser quantum (or pump) efficiency ζ_2. Note from this plot that for a ζ_2 of 20 per cent, 25 to 40 W of heat must be dissipated from the laser crystal for a power output of 5 to 10 W from the 1970 solar-pumped laser communication system.

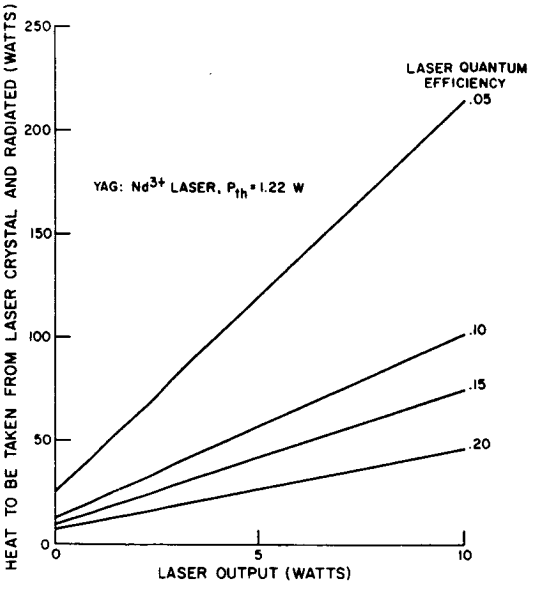


Fig. 21. Heat to be taken from laser crystal and radiated vs laser output.

The total amount of heat to be radiated from the system must include not only the heat to be taken from the laser, but also any heat which might impinge on the surroundings as a result of the pumping process. The total pump power supplied to the laser is equal to the heat to be dissipated plus the power output. According to the analysis of Sec. IV. B above (see paragraph following Eqs. 29 and 30), this pump power is only 0. 206 times the total solar power collected. Thus, the "useless" pump power (that pump power not actually converted into laser output) and the power collected but unsuitable for pumping must be radiated from the system.

The total power to be radiated is given by:

$$\frac{\text{Laser heat to be radiated} + \text{Laser power output}}{0.206} - \text{Laser Power Output}$$

Fig. 22 shows the total power which must be radiated versus laser power output for various values of the laser quantum efficiency, ζ_2.

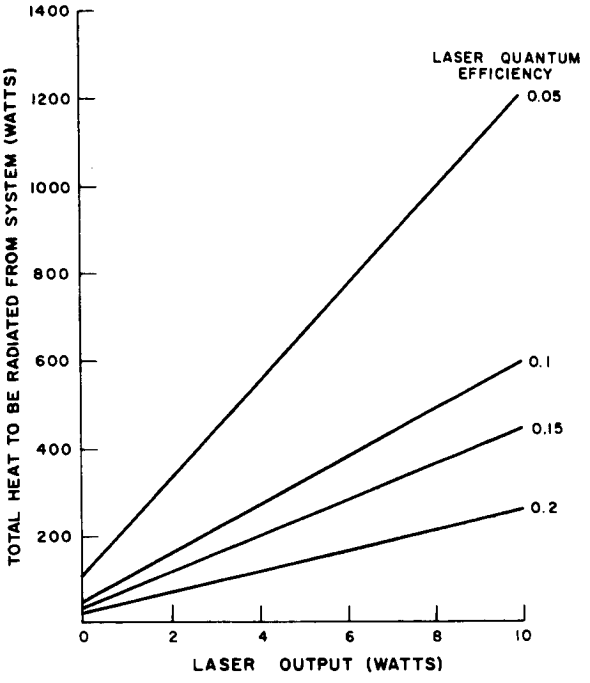


Fig. 22. Total heat to be radiated from system vs laser output.

3. Optimum Design of Uniform, Rectangular Radiating Fins

The minimum surface area required to radiate heat into space at a given rate at temperatures not exceeding a certain maximum value will occur if the entire radiating surface temperature is uniform and at this maximum temperature. If the radiating surface is convex (or flat) so that every element of it can "see" into the complete hemisphere of directions, and no other bodies (planets or parts of the space vehicle, especially the Sun) can be "seen" from any element of its surface, the required minimum radiator area is determined by the relation

$$Q_1 = A \varepsilon \sigma T^4 \tag{52}$$

where Q_1 is the heat dissipation rate, T is the maximum temperature, ϵ is the

surface emissivity, σ is the Stefan-Boltzmann constant, and A is the minimum radiator area. When heat must be dissipated from a concentrated source, as is the case for the laser under consideration here, it is difficult to distribute the heat uniformly over the radiator surface. It is much simpler to distribute it along one edge of a good conducting radiating fin. Therefore, from a practical point of view, optimum fin design will be analyzed by considering fin weight.

The following assumptions are made:

a. The fin is rectangular and of uniform thickness.

b. One edge of the fin is uniformly heated at a temperature T_1.

c. The other edges of the fin are insulated.

d. The fin is one-sided, i. e. , there is only one radiating surface. A two-sided fin can be analyzed using these results by considering it to consist of two one-sided fins back-to-back.

e. Steady conditions prevail.

f. The fin is homogeneous and isotropic in its properties.

g. All other linear dimensions of the fin are large compared with the fin thickness. This permits an accurate, one-dimensional analysis.

The following notation is adopted:

Q = rate of heat transport per unit of fin width

T = $T(x)$ = fin absolute temperature

x = distance along fin

ω = fin thickness

K = fin material thermal conductivity

ϵ = surface emissivity

σ = Stefan-Boltzmann constant

ℓ = $x_0 - x_1$ = fin length

Subscripts 0, 1 refer to the fin end and root, respectively. Other special symbols will be introduced and defined as needed.

The steady temperature distribution in a fin, as described above, is governed by the equations:

$$Q = -K\omega\, dT/dx \text{ (Fourier's Law)} \tag{53}$$

$$dQ/dx = -\epsilon\sigma T^4 \text{ (Energy Relationship)} \tag{54}$$

together with the boundary conditions:

at $x = x_1$ (fin root) $T = T_1$; $Q = Q_1$

at $x = x_0$ (fin end) $T = T_0$; $Q = 0$ (insulated end).

The usual method for solution of a system of equations like Eqs. 53 and 54 is to eliminate Q between them, obtaining a second-order differential equation — which in this case is nonlinear, the solution being rather difficult. Here we will adopt a different scheme.

The independent variable x can be eliminated between Eq. 53 and Eq. 54 yielding:

$$K\omega \frac{dT}{Q} = \frac{1}{\epsilon\sigma T^4} dQ \tag{55}$$

Separating the variables and integrating yields:

$$\frac{1}{5}\epsilon\sigma K\omega(T^5 - T_0^5) = \frac{1}{2}(Q^2 - Q_0^2) = \frac{1}{2}Q^2 \text{ (since } Q_0 = 0) \tag{56}$$

or

$$Q = \left(\frac{2}{5}\epsilon\sigma K\omega\right)^{\frac{1}{2}}\left(T^5 - T_0^5\right)^{\frac{1}{2}} = \left(\frac{2}{5}\epsilon\sigma K\omega T_1^5\right)^{\frac{1}{2}}\left(\tau^5 - \tau_0^5\right)^{\frac{1}{2}} \tag{57}$$

where $\tau = T/T_1$. Thus, to dissipate the quantity Q_1 of heat at T_1, it is required that

$$Q_1 = \left(\frac{2}{5}\epsilon\sigma K\omega T_1^5\right)^{1/2}\left(1 - \tau_0^5\right)^{1/2}. \tag{58}$$

Then, obviously

$$Q/Q_1 = \left(\tau^5 - \tau_0{}^5\right)^{\frac{1}{2}} \Big/ \left(1 - \tau_o{}^5\right)^{\frac{1}{2}}. \tag{59}$$

To determine the fin length satisfying these conditions, we must integrate Eq. 53, namely:

$$dx = -\frac{K\omega dT}{Q} = -\left[\frac{K\omega T_1}{\left(\frac{2}{5}\epsilon\sigma K\omega T_1{}^5\right)^{1/2}}\right]\frac{d\tau}{\left(\tau^5 - \tau_0{}^5\right)^{1/2}} \tag{60}$$

yielding

$$\ell^* = \int_{\tau_o}^{1} \frac{d\tau}{\left(\tau^5 - \tau_o{}^5\right)^{\frac{1}{2}}} \quad \text{where} \quad \frac{\ell^*}{\ell} = \left(\frac{2\,\epsilon\sigma\, T_1{}^3}{5\,K\omega}\right)^{\frac{1}{2}}. \tag{61}$$

This expression looks better, since τ_0 is not known in advance, if we let

$$\xi = \frac{\tau}{\tau_0}.$$

Then

$$\ell^* = \tau_0{}^{-\frac{3}{2}} \int_1^{1/\tau_0} d\xi/(\xi^5 - 1)^{\frac{1}{2}}. \tag{62}$$

In Eqs. 61 and 62, τ_0 can be regarded as a parameter, and ℓ^* can be tabulated as a function of τ_0. The integral converges in spite of the singularity at $\xi = 1$. We evaluate Eq. 62 in three segments, as follows:

(a) Let $\xi = 1 + \delta$. Then for small δ, $\xi^5 \approx 1 + 5\,\delta$, and

$$i = \int_1^{1+\delta_1} d\xi/(\xi^5 - 1)^{\frac{1}{2}} \approx \int_0^{\delta_1} d\delta/\sqrt{5\delta} = \frac{2}{\sqrt{5}}\,\delta_1{}^{\frac{1}{2}} \tag{63}$$

Thus, for $\delta_1 = 0.01$, $i = .2/\sqrt{5} = 0.0895$. For $\delta_1 = 0.04$, $i = 0.1790$.

(b) For ξ large, say

$$\xi > 3, \left(\xi^5 - 1\right)^{\frac{1}{2}} \approx \xi^{\frac{5}{2}}$$

Then,

$$i = \int_3^{\frac{1}{\tau_0}} d\xi \Big/ \left(\xi^5 - 1\right)^{\frac{1}{2}} \approx \int_3^{\frac{1}{\tau_0}} \xi^{-5/2}\, d\xi = \frac{2}{3}\left(3^{-3/2} - \tau_0^{3/2}\right) = \frac{2}{3}\left(0.1924 - \tau_0^{3/2}\right) \tag{64}$$

(c) For $1.01 < 1/\tau_0 < 3$, we compute the integral numerically. The results of these computations are given in Table III.

Table III. EVALUATION OF $i(\xi_0) = \int_1^{\xi_0} d\xi \Big/ (\xi^5 - 1)^{1/2}$

ξ_0	$i(\xi_0)$	ξ_0	$i(\xi_0)$
1.00	0	2.2	0.7170
1.01	0.0895		
1.02	0.1271	2.4	0.7424
1.03	0.1552		
1.04	0.1787	2.6	0.7629
1.05	0.1991		
1.1	0.2787	2.8	0.7798
1.2	0.3837		
1.3	0.4550	3.0	0.7939
1.4	0.5092	4.0	0.8388
1.5	0.5525	5.0	0.8624
1.6	0.5881	6.0	0.8764
1.7	0.6181	7.0	0.8835
1.8	0.6437	8.0	0.8928
1.9	0.6658	9.0	0.8976
2.0	0.6850	10.0	0.9009
		∞	0.9222

These results can now be used to design a minimum weight fin to dissipate Q_1 of heat at a root temperature T_1. From Eq. 58, the fin thickness is:

$$\omega = \frac{5}{2} Q_1^2 \Big/ \epsilon\sigma K T_1^5 \left(1 - \tau_0^5\right), \tag{65}$$

or, denoting $Q_1^* = Q_1/\epsilon\sigma T_1^4$, $Q_1^{**} = Q_1/K T_1$

$$\omega = \frac{5}{2} Q_1^* Q_1^{**} \Big/ (1 - \tau_0^5) . \tag{66}$$

The fin length from Eq. 61 is:

$$\ell = \ell^*(\tau_0) \Big/ \left(\frac{2\epsilon\sigma T_1^3}{5 K\omega}\right)^{\frac{1}{2}} \tag{67}$$

or

$$\ell = \frac{5}{2} Q_1^* \ell^* (\tau_0) \Big/ \left(1 - \tau_0^5\right)^{\frac{1}{2}} . \tag{68}$$

The fin volume (proportional to weight) per unit fin width is then:

$$\bar{V} = \ell\omega = \frac{25}{4} Q_1^{*2} Q_1^{**} \ell^* (\tau_0) \Big/ \left(1 - \tau_0^5\right)^{\frac{3}{2}} . \tag{69}$$

The minimum fin weight (optimum) will have that value of τ_0 which minimizes Eq. 69. This value holds for all fins, there being a single value of τ_0 which minimizes Eq. 69, namely that value for which $\ell^* (\tau_0) \Big/ \left(1 - \tau_0^5\right)^{3/2} = \text{min.}$

Rewriting this in terms of $\xi_0 = 1/\tau_0$, we have to find ξ_0 for which

$$\left(\frac{\xi_0^{\,5}}{\xi_0^{\,5}-1}\right)^{\frac{3}{2}} \ell^*(\xi_0) = \text{min.} \tag{70}$$

This computation has been carried out numerically, yielding $\xi_0 = 1.265$, which makes $\tau_0 = 0.790$. For an optimum fin, then, the fin thickness is

$$\omega = \frac{5}{2(0.692)} Q_1^* Q_1^{**} = 3.615\, Q_1^* Q_1^{**}. \tag{71}$$

The fin length is

$$\ell = \frac{5(.612)}{2(.832)} Q_1^* = 1.84\, Q_1^* . \tag{72}$$

The fin volume (per unit of fin width) is then

$$V = \omega\ell = 6.67\, Q_1^{*2} Q_1^{**} , \tag{73}$$

and the fin weight is

$$M = \gamma V = 6.67\, \gamma\, Q_1^{*2} Q_1^{**} ,$$

where γ is the fin material density. Note from Eq. 73 that the radiating surface of an optimum edge-heated rectangular fin must be 1.84 times the minimum radiating areas to dissipate Q_1 of heat at a uniform temperature T_1, as given in Eq. 52.

Radiator area versus laser power output is plotted in Fig. 23. Fig. 24 shows the corresponding required radiator area for radiating the total heat which must be dissipated from the laser system.

The product βM, where

$\beta \triangleq$ fin aspect ratio $= \dfrac{\text{root dimension of radiator}}{\text{edge dimension of radiator}}$

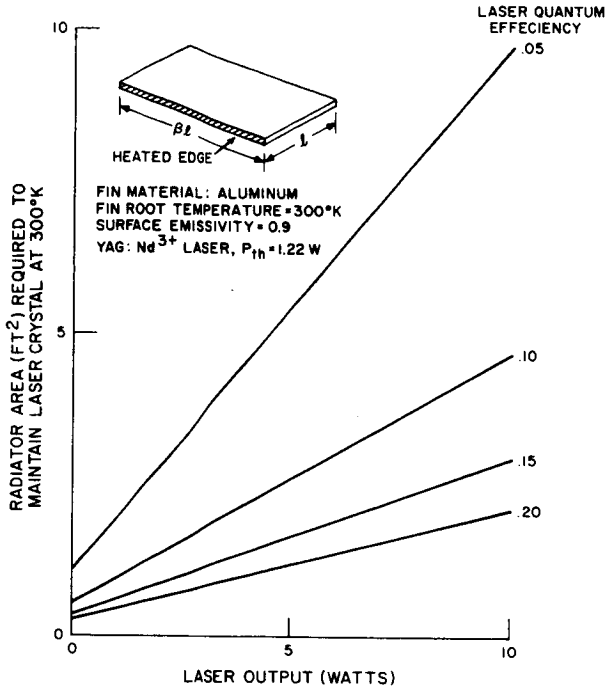


Fig. 23. Radiator area required to maintain laser crystal at 300°K versus laser output.

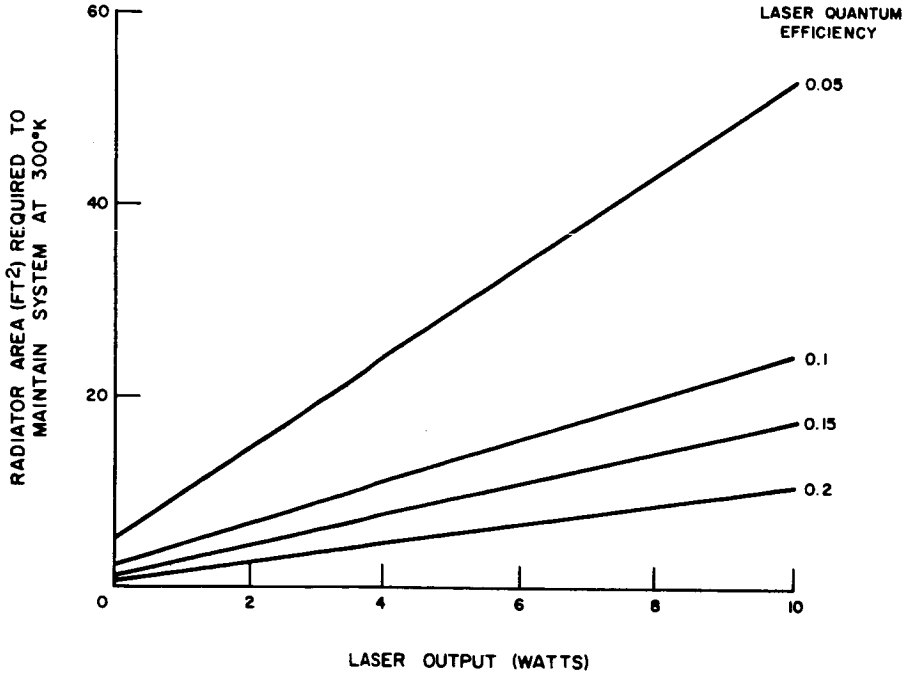


Fig. 24. Radiator area required to maintain system at 300°K versus laser output.

and

M = fin weight,

is a useful design parameter. A plot of βM vs power output for aluminum radiating fins (γ = 173 lb/ft^3), with fin root temperature T at 300°K, surface emissivity ϵ equal to 0.9, and several values of laser quantum (or pump) efficiency ζ_2, is given in Fig. 25. Note from this plot that for 5 to 10 W of power output from the 1970 solar-pumped laser communication system, the radiator weight can be made small by adjusting the ratio β. If β is made large by using physically large radiators, or by appropriate coolant flow routing, large amounts of heat can be radiated by a relatively light radiator. For example, reference to Fig. 25 shows that for β = 10, a radiator weighing 40.7 to 140.5 pounds would suffice to accomplish the necessary cooling. For larger β, even lighter fins would be adequate.

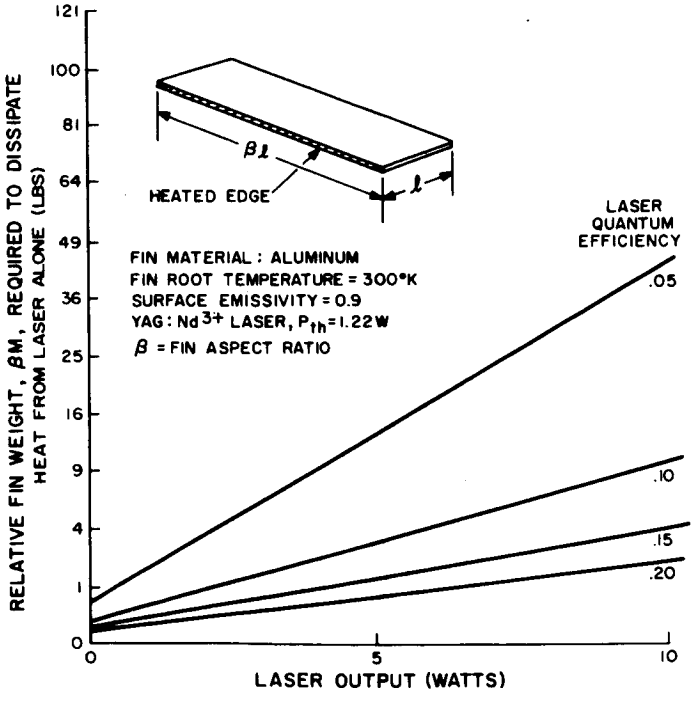


Fig. 25. Relative fin weight, βM, required to dissipate heat from laser alone versus laser output.

D. MODULATION

As discussed in Sec. II. A where introductory remarks were made concerning the analysis of the performance of a laser communication system, it was pointed out that, among other things, a choice of modulation method depended on bandwidth required. Essentially, the choice divides between narrow-band and wideband systems, depending on the information capacity requirements. Also, once the narrow-band or wideband choice is made, accomplishment of the modulation method selected depends on the devices available for implementation. In the discussion to follow, these factors will be analyzed, and the choices made for the communication system proposed in Sec. II. B will be supported.

1. Magnetic Modulation

An effective technique for accomplishing narrow-band modulation of a laser source is magnetic modulation. As described in Appendix I, magnetic modulation of a laser beam is based on the energy (or equivalently, frequency) shifts which occur as the result of application of a magnetic field to an atomic system. By virtue of the normal Zeeman effect, it is possible to shift the center frequency of the fluorescent line or to effectively broaden the fluorescent line. The former takes place when a homogeneous field is applied to the laser crystal and the latter when a an inhomogeneous field is used. Both methods control the power output of the laser crystal and thereby accomplish modulation.

Experimental data described in Sec. V below shows the inhomogeneous field method to be effective in accomplishing 100 per cent modulation. As shown in Appendix I, the drive current needed for 100 per cent modulation with this technique is given by

$$I = \frac{4\pi k N r^2 h \Delta\nu_{eff} (\xi - 1)}{L g \beta} \times 10^{-4} \text{ (amperes)} \qquad (74)$$

where

k = 0.53 (finite solenoid compensation factor)

N = number of turns in coil

r = radius of coil (meters)

h = Planck's constant

$\Delta\nu_{eff}$ = effective linewidth (Hz)

ξ = pump power ratio

L = coil inductance (henrys)

g = gyromagnetic ratio

β = Bohr magneton (9.3×10^{-21} erg/gauss)

The solenoid used in the experiments described in Sec. V. A. 1 below had the following parameter values: $N = 20$ turns, $r = 1/8$ inch, $L = 6\ \mu H$. The laser crystal used in the experiments had an effective linewidth $\Delta\nu_{eff}$ ($= \Delta\nu/4$) of 7.5×10^8 Hz. Substituting these values in Eq. 74 yields

$$I = 1.28(\xi - 1) \tag{75}$$

This relationship is plotted in Fig. 26. From this plot we see that if the pump power is twice the threshold value (as it was for the experiments conducted), the current needed for 100 per cent modulation should have been 1.28 amperes. The actual current needed for 100 per cent modulation was found to be about 1 ampere, which is slightly lower than the theoretical value. This is not very surprising since all phases of laser operation were not taken into account by the theory, and several approximations were made. The theoretical result is, however, accurate enough for design purposes. Also, it provides a basis for evaluating the effect of variation of parameters such as pump power, linewidth, doping levels, etc.

A model of the basic magnetic modulation circuit is shown in Fig. 27. Here a DC current is established in L', its magnitude being sufficient to create a magnetic field in L that will inhibit laser action. Modulation is accomplished by sending a pulse through L in order to create an opposing magnetic field which, if large enough, will initiate laser action.

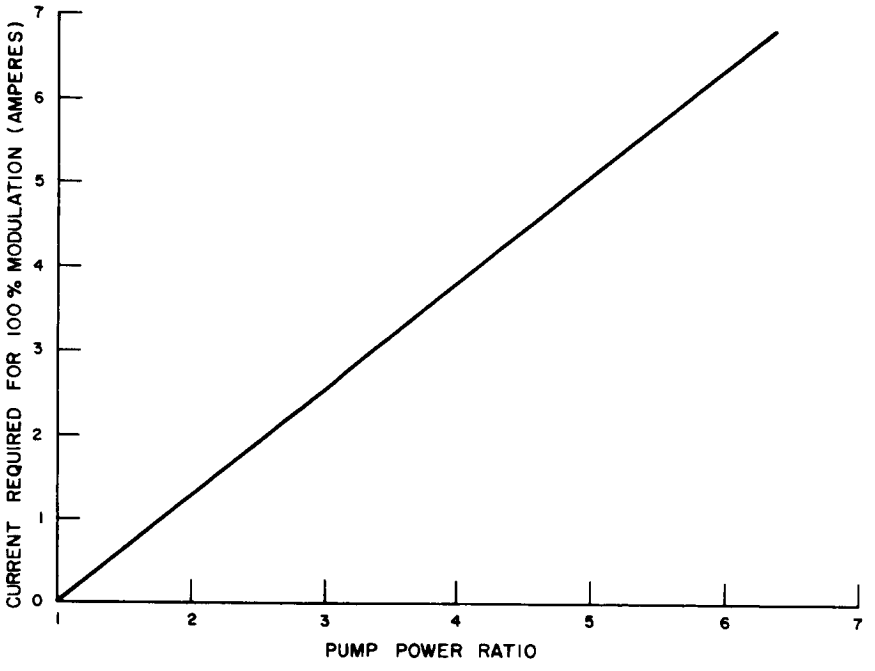


Fig. 26. Current I required for 100 per cent modulation vs pump power ratio ξ.

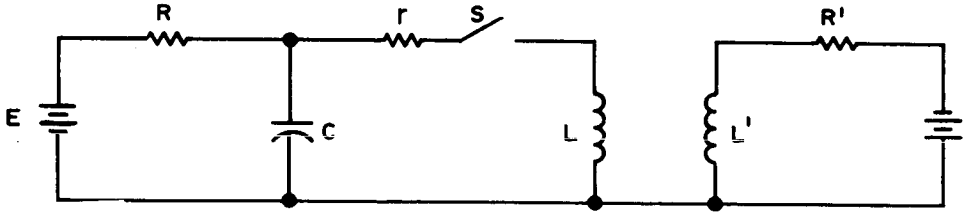


Fig. 27. Magnetic modulation circuit.

Using the model of Fig. 27, the power requirements for magnetic modulation of a crystal laser were analyzed in Appendix I for the two cases of (1) pulse modulation and (2) analog modulation. From this analysis, graphs of modulation power vs information bandwidth were determined and are shown in Figs. 28 and 29.

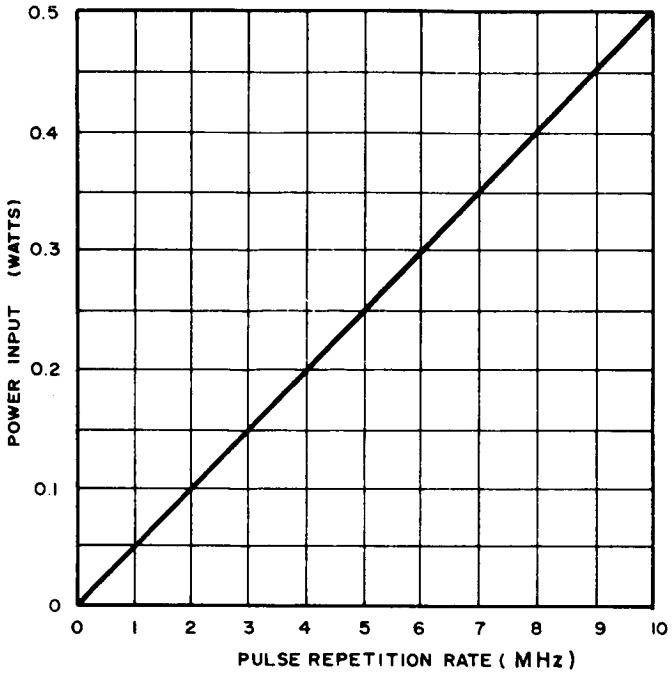


Fig. 28. Power input vs pulse repetition rate for pulse modulation.

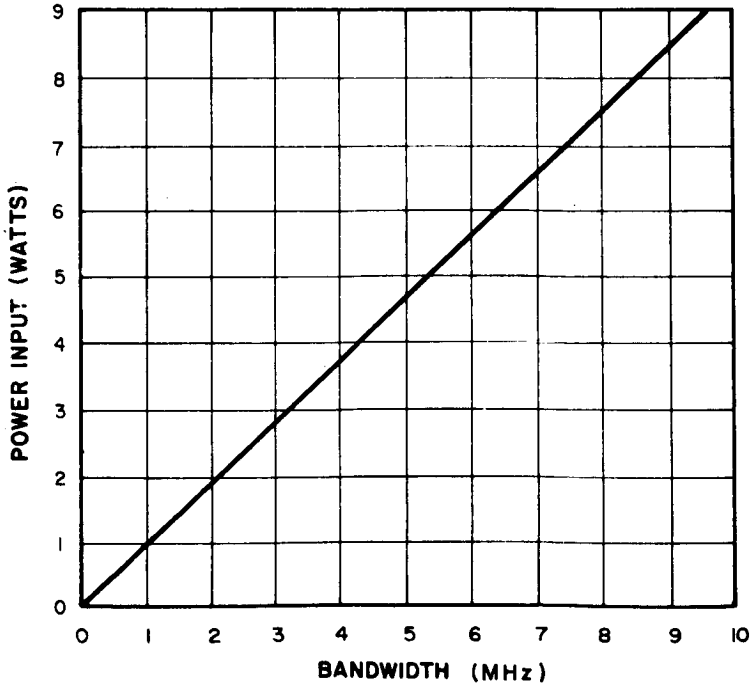


Fig. 29. Power input vs modulation bandwidth for analog modulation (L fixed at 6 μH).

Note from Fig. 29 that only about 4 watts of input power is needed for an information bandwidth of 4 MHz; however, though the power input required is moderate, it has been observed experimentally that output power decreases rapidly for modulation frequencies greater than about 300 kHz. This apparent maximum rate is thought to be a result of the spiking repetition rate of the crystal. Thus, any video information sent by this method would necessarily be sent at a low frame rate.

2. Electro-Optic Modulation

For some space applications, low frame rate video information may be adequate, in which case magnetic modulation would be sufficient. However, if rapid scanning rates are necessary, as they certainly will be in 1970, the laser radiation can be externally wideband modulated by means of an electro-optic crystal modulator. The major components of such an optical communication system are shown in the block diagram of Fig. 30.

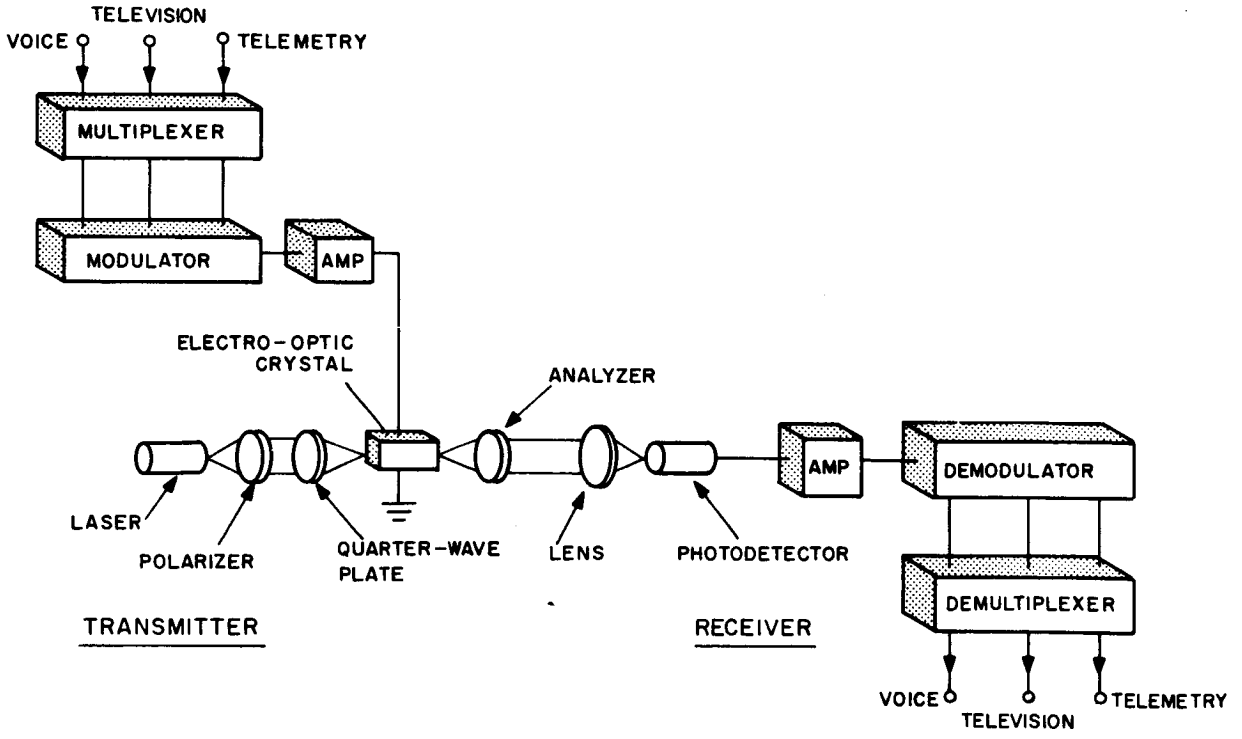


Fig 30. Major components of a wideband optical communications system.

In this system, amplitude modulation of the laser beam is accomplished as follows: The polarizer permits light polarized in only one plane to reach the crystal. As the plane polarized light passes through the crystal it becomes elliptically polarized, the amount of ellipticity depending on the voltage applied to the crystal.

Since the analyzer transmits light polarized in only one plane, the intensity of the light transmitted through the modulator depends on the amount of rotation of polarization introduced by the crystal. Hence, the light beam is amplitude modulated in accordance with the voltage applied to the crystal.

The configuration indicated in Fig. 30 is quite general. The multiplexer and demultiplexer blocks can be of any type, e.g., frequency division, time division, etc. Similarly, the modulator and demodulator blocks may be designed for the desired modulation technique, e.g., PCM, delta, analog-FM, etc. In such a system, there are three major design problems (in addition to design of the laser source itself): (1) the modulation technique to use, (2) the type of electro-optic crystal to use, and (3) the design of the electro-optic drive modulator. These factors are discussed below.

a. Comparison of Modulation Methods

A comparison given in Appendix J of modulation methods suitable for transmitting TV pictures reveals (as evidenced by Fig. 31) that:

(1) Wideband modulation methods such as FM, PCM and delta modulation permit an essentially noise-free TV picture (i.e., a TV picture having an SNR of 30 dB) to be transmitted over a much greater range than can be realized by narrow-band modulation methods such as AM.

(2) If a TV picture having an extremely high SNR is needed, say 50 dB or higher, then wideband PCM, which increases SNR exponentially with bandwidth expansion, is the best choice.

(3) If an SNR of only 30 dB is needed, other wideband modulation methods provide equal, or better, performance with less complex equipment.

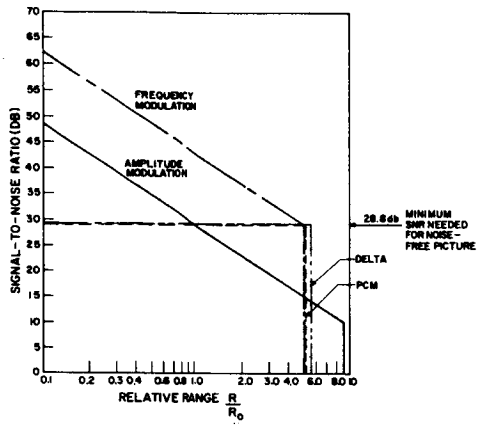


Fig. 31. Comparison of modulation methods. R_o is the range at which the SNR of the AM system is 30 dB.

(4) For an SNR of 30 dB, PCM, delta and FM modulation methods provide about the same communication range.

Thus, we conclude that wideband modulation should be used and that, as far as communication range is concerned, it makes little difference which wideband modulation method is used. Hence, the wideband modulation method which leads to the simplest most reliable equipment should be chosen. Accordingly, we chose FM for the experimental system. This experimental system is illustrated in Fig. 32, which is a specific member of the family of systems illustrated in Fig. 30.

Aside from being a simple, reliable modulation method, FM avoids a serious signal distortion problem introduced by piezoelectric resonances of the electro-optic crystal. Shown in Fig. 33 is the measured frequency response of a typical GaAs crystal, with frequency markers at 1, 3, 5 and 7 MHz. It can be seen that the piezoelectric resonances are prevalent up to about 700 kHz. Through the use of FM the modulation spectrum can be shifted away from the piezoelectric resonance range, thus avoiding signal distortion.

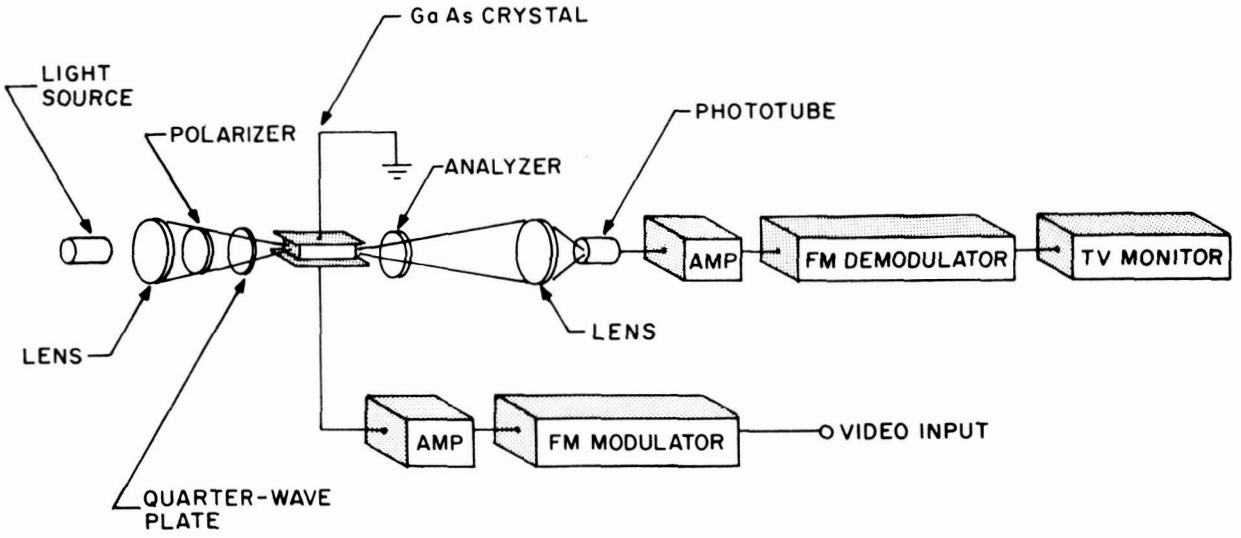


Fig. 32. Electro-optic TV communication system utilizing FM subcarrier modulation.

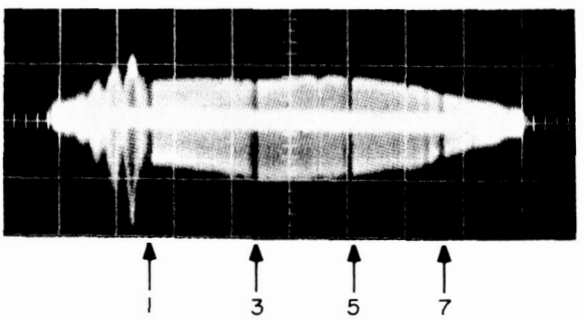


Fig. 33. Frequency response of GaAs crystal. (Markers are at 1, 3, 5 and 7 MHz.)

Under an RCA-sponsored program both baseband (i.e., directly modulating the light beam with the TV signal) and FM-subcarrier (i.e., frequency modulating a 5-MHz subcarrier with the TV signal) modulation methods were evaluated to determine the extent of the distortion introduced by piezoelectric resonances. The photographs shown in Figs. 34 and 35 illustrate the results. With baseband modulation (Fig. 34) the resultant TV picture is plagued by "herringbone" interference due to piezoelectric resonances; with FM (Fig. 35) no such interference appears.

b. Characteristics of Electro-Optic Crystals

To date the only practical light modulator that allows the realization of modulation bandwidths of 5 MHz or higher is the electro-optic crystal modulator illustrated in Fig. 36. As shown in this figure, the electro-optic crystal is placed between a polarizer and an analyzer. The polarizer permits only light polarized in one plane to reach the crystal. As the plane polarized light passes through the crystal, it becomes elliptically polarized, the amount of ellipticity depending on the voltage applied to the crystal. Since the analyzer transmits light polarized in only one plane, the intensity of the light transmitted through the modulator depends on the amount of rotation of polarization introduced by the crystal. Hence, the light beam is amplitude modulated in accordance with the voltage applied to the crystal.

1) Drive Voltage Requirements

As shown in Appendix L, the intensity of the light transmitted through the analyzer is given by

$$I = I_o \sin^2 \left(\frac{\sqrt{3}\pi r n^3 \ell V \sin \omega t}{2\lambda d} \right) \tag{76}$$

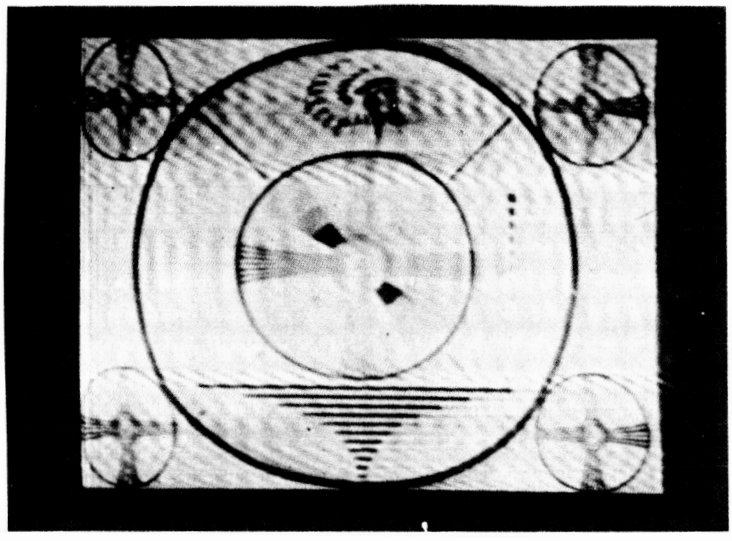


Fig. 34. Picture obtained with baseband modulation.

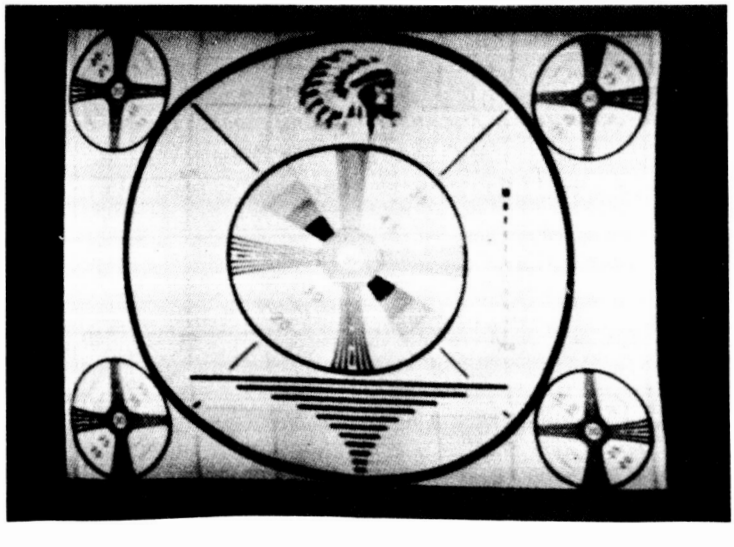


Fig. 35. Picture obtained with FM-subcarrier modulation.

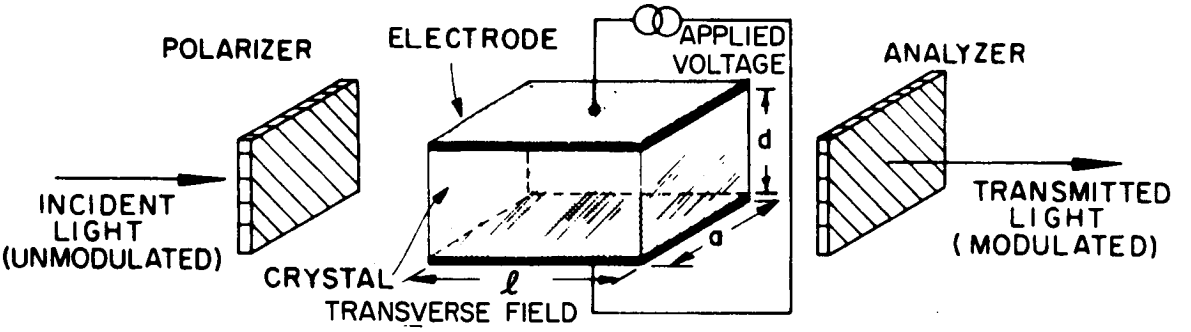


Fig. 36. Electro-optic modulator.

where

I_o = intensity of laser source (W)

λ = wavelength of incident laser light (cm)

r = electro-optic coefficient (specifically, r_{41}) at wavelength λ (cm/V)

n = index of refraction at wavelength λ

d = crystal thickness in direction of applied electric field (cm)

ℓ = crystal length in direction of incident laser light beam (cm)

ω = frequency of modulating signal (Hz)

V = peak amplitude of modulating voltage

Eq. 76 shows that the peak-to-peak drive voltage needed to achieve 100 per cent modulation is

$$V_{\lambda/2} = \frac{1}{\sqrt{3}} \frac{\lambda}{r_{41} n^3} \frac{d}{\ell} \tag{77}$$

This voltage is known as the half-wave voltage because it results in a 180-degree shift in the plane of polarization. Substituting Eq. 77 into Eq. 76 gives

$$I = I_o \sin^2 \frac{\pi V \sin \omega t}{2 V_{\lambda/2}} \tag{78}$$

A plot of this equation, Fig. 37, indicates that the modulated beam varies at twice the modulation frequency. This is a consequence of I reaching a maximum twice for every cycle of the modulating voltage; that is, I has a maximum when $\sin \omega t = 1$ and also when $\sin \omega t = -1$.

To eliminate this "double frequency" effect, the electro-optic modulator must be biased either electrically or optically. The latter is easier and can be accomplished by placing a quarter-wave plate, properly oriented to obtain circularly polarized light, between the polarizer and the crystal. For this case the transmitted intensity is given by (see Eq. L-17)

$$I = I_o \sin^2 \left(\frac{\pi V \sin \omega t}{2 V_{\lambda/2}} + \frac{\pi}{4} \right) \tag{79}$$

The plot of this equation, Fig. 38, shows that the transmitted intensity is $I_o/2$ when the modulation voltage is zero and that 100 per cent modulation is achieved when the peak-to-peak modulation voltage is equal to $V_{\lambda/2}$.

As shown in Appendix L (see Eq. L-18), a peak modulating voltage $V_m \left[= (V \sin \omega t)_{\text{peak}} \right]$ applied to the crystal will produce a percentage depth of modulation m given by

$$m = \left(\sin \frac{\pi V_m}{V_{\lambda/2}} \right) 100 \tag{80}$$

This equation, plotted in Fig. 39, shows that reasonably high per cent modulation can be achieved even when the modulation voltage is considerably less than $V_{\lambda/2}$. For example, 80 per cent modulation is achieved when $V_m/V_{\lambda/2} = 0.3$.

2) Crystal Constants

Since the solar-pumped lasers described in Sec. IV.A.1 above emit radiation at 1.06 μm, the electro-optic crystal modulator used with such lasers must be operable at this wavelength.

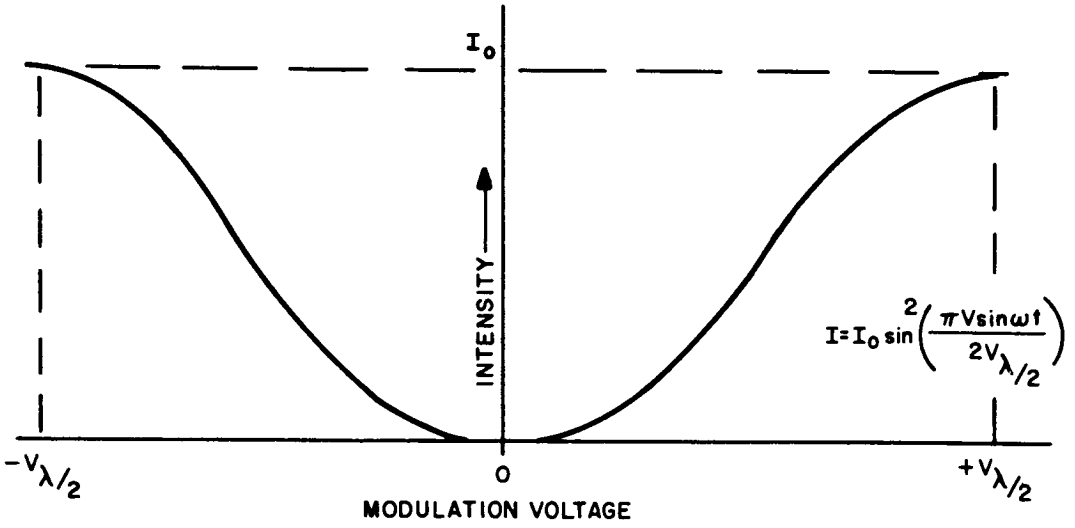


Fig. 37. Intensity vs modulation voltage.

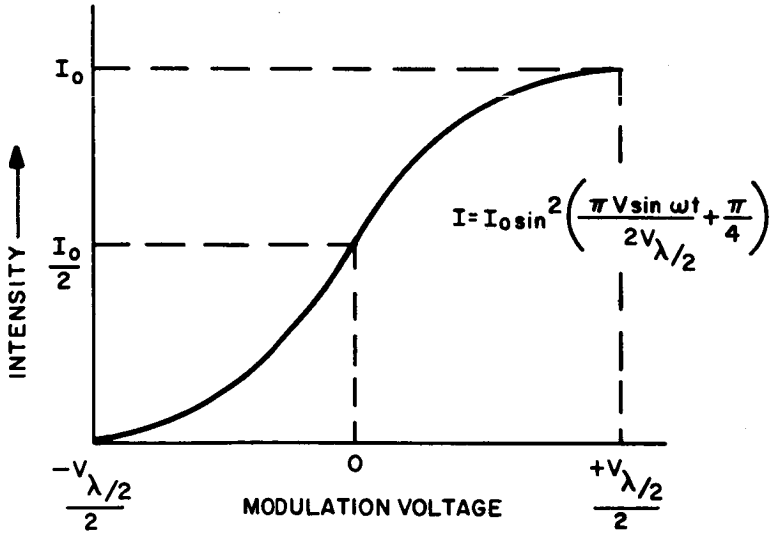


Fig. 38. Intensity vs modulation voltage with $\lambda/4$ plate.

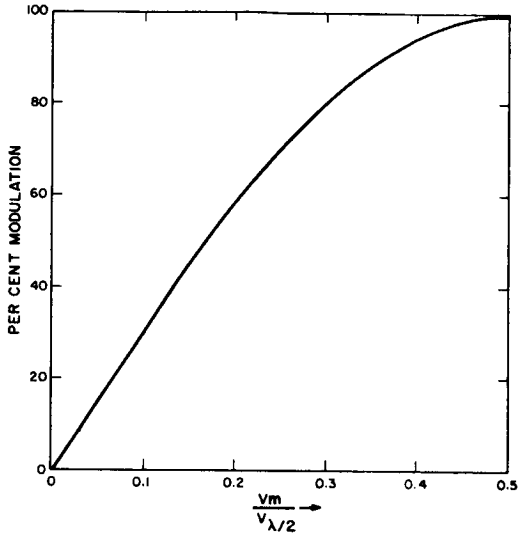


Fig. 39. Per cent modulation vs $V_m/V_{\lambda/2}$.

Constants of electro-optic crystals operating at 1.06 microns are given in Table IV. A summary of the advantages and disadvantages of each of these crystals given below leads to the conclusion that the GaAs crystal is the best choice at the present state of the art.

(a) ADP and KDP

Sterzer[1] has shown that because KDP and ADP are uniaxial in the absence of an electric field, the angular aperture is orders of magnitude smaller than that of cubic crystals such as CuCl and GaAs.

This means that the per cent modulation that can be achieved with ADP and KDP is critically dependent upon the divergence and the alignment of

[1] F. Sterzer, D.J. Blattner, and S.F. Miniter, "Cuprous Chloride Light Modulators," Jour. of Optical Society of America, January 1964.

Table IV. CONSTANTS OF ELECTRO-OPTIC CRYSTALS

Property	Crystal Type								
	ADP	KDP	CuCl	ZnS	Hexamine	KTN	$BaTiO_3$	$Bi_4(GeO_4)_3$	GaAs
Crystal Type	Ionic	Ionic	Ionic	Valence	Molecular	Ionic	Ionic	Valence	Valence
Crystal Structure	Hexagonal	Hexagonal	Cubic	Cubic	Cubic	Cubic	Cubic	Cubic	Cubic
Index of Refraction	1.53	1.51	1.93	2.4	1.59	2.29	2.4	2.07	3.34
Dielectric Constant	12	20	8	8	3	10,000	10,000	–	11.2
Loss Tangent	0.04	0.02	0.0015	<0.001	0.0003	0.004	0.01	–	<0.0003
Modulation	Longitudinal	Longitudinal	Linear Transverse	Linear Transverse	Linear Transverse	Quadratic Transverse	Quadratic Transverse	Linear Transverse	Linear Transverse
Electro-Optic Coefficient	6.3×10^{-10}	11×10^{-10}	6×10^{-10}	2×10^{-10}	4.2×10^{-10}	$(g_{11}-g_{12}) = 0.17\ m^4/C^2$	$(g_{11}-g_{12}) = 0.13\ m^4/C^2$	1×10^{-10}	1×10^{-10}
Half-Wave Voltage	22kV	13.5 kV	$11.5 \frac{d}{\ell}$ kV	$18 \frac{d}{\ell}$ kV	$30 \frac{d}{\ell}$ kV	–	–	$69 \frac{d}{\ell}$ kV	$13.4 \frac{d}{\ell}$ kV
Transparency Range	0.4-1.7μm	0.4-1.7μm	0.4-20.5μm	0.4-	0.4-2μm	0.5-5μm	0.5-6.9μm	0.4-6μm	0.9-15μm
Largest Available Length	15 cm	15 cm	1 cm	No Practical Length Available	1 cm	0.3 cm	1 cm	Greater Than 1 cm	5 cm
Environmental Considerations	Hygroscopic	Hygroscopic	Hygroscopic Easily Strained	–	Hygroscopic	Controlled Temp $20^o \pm 0.2^oC$	Controlled Temp $120^o \pm 0.2^oC$	–	–

the laser beam. Sterzer also has shown that because ADP and KDP exhibit only a longitudinal electro-optic effect,[1] they require either transparent electrodes (which are not efficient at high frequencies) or electrodes with holes (which operate only with fringing fields). Since the electric field must be applied parallel to the light path, the modulation voltage needed for a given per cent modulation is independent of crystal length and, therefore, large voltages are required for a reasonable per cent modulation. Hence, although ADP and KDP crystals are transparent in a convenient spectral range and relatively large crystals are available, they do not show much promise for the proposed application.

(b) Cuprous Chloride

From the standpoint of voltage requirements and the optical transmission band, cuprous chloride appears to offer considerable promise. However, large crystals of this material are extremely difficult to grow. Also, strains and metallic copper filaments commonly occur in the process of crystal growth, and they tend to make the breakdown voltage very low. After the crystal has broken down, it is no longer suitable for modulation purposes. Therefore, although CuCl shows great promise, its full potential will not be realized until better fabrication techniques are developed.

(c) Zinc Sulphide

ZnS may eventually find application in the visible region, but, thus far, this type crystal has not been grown in the size required for practical uses.

(d) Hexamine

Hexamine (hexatetramethylamine) is the best electro-optic modulator currently available for the visible region. In contrast to CuCl, it has

[1] ADP and KDP can be used in a transverse mode only with extremely monochromatic light.

an extremely high breakdown potential. It has the disadvantages, however, that it is extremely hygroscopic, $V_{\lambda/2}$ is relatively high, and it is unstable in contact with commonly used potting materials.

(e) KTN

Geusic[1] has shown that KTN crystals provide high per cent modulation with exceptionally low drive voltage. He has measured half-wave voltages as low as 15.3 volts in KTN crystals measuring 0.635 x 0.635 x 2.36 mm, operated with a DC bias of about 300 volts. Clearly, the drive voltage required for KTN is more than two orders of magnitude lower than that of any of the linear transverse crystals listed in Table IV.

However, KTN is not without problems of its own. One problem, due to the dielectric constant being three orders of magnitude higher than that of the linear transverse crystals, is that the capacitive load presented by KTN is rather large. Hence, although the drive voltage requirement is small, the drive current requirement is large. In other words, the problem of designing a wideband, high-voltage amplifier has been traded for the problem of designing a wideband, high-current amplifier. A second problem is that the operating temperature of KTN must be controlled to within 0.2°C, necessitating automatic temperature control. A third problem is due to KTN being an ionic crystal. To achieve high per cent modulation it is necessary to apply a DC bias across the crystal. The DC field forces the ions to drift toward the electrodes, causing the crystal to become polarized. Under normal operating conditions polarization can occur within one-half hour, preventing the crystal from functioning properly until the crystal is depolarized. A fourth problem is that the crystals are extremely difficult to grow. Crystals grown to date are very small and not entirely cubic.

---

[1] J. E. Geusic, F. S. Chen, S. K. Kurtz, J. G. Skinner, and S. H. Wemple, "The Use of Perovskite Paraelectrics in Beam Deflectors and Light Modulators," Proc. IEEE, Vol. 52, pp. 1258-1259; November 1964.

From the foregoing discussion it is evident that, although KTN shows great promise, there are formidable research problems to be overcome before this type crystal can be used in practical systems.

(f) $BaTiO_3$

This is a transverse-quadratic crystal having properties similar to those of KTN. Fig. 40 shows the potential gradient measured across a typical $BaTiO_3$ crystal after a DC field had been applied for one-half hour. This data clearly illustrates the polarization problem associated with this type crystal. Unless a uniform DC field can be maintained across the crystal, a modulation voltage comparable to that required for a linear transverse crystal is needed. Since this type crystal has the same problems which plague KTN, it is not a good choice for the proposed system.

(g) Bismuth Germanate Bi_4 (GeO_4)

This is a newcomer to the field of electro-optic modulators. It was recently developed by RCA's research group in Zurich, Switzerland. Although its half-wave voltage is relatively high, it is transparent in the visible range and preliminary results indicate that rather large crystals (greater than 5 cm) can be grown. With regard to the proposed program, however, it is not likely that suitable crystals will be available soon enough for the first experimental system.

(h) Gallium Arsenide (GaAs)

Until recently, GaAs showed little potential as an electro-optic modulator because the low resistivity of available crystals prevented application of the high electric fields required for a reasonable per cent modulation. However, research efforts by RCA's Semiconductor Materials Research Group in Somerville, N.J. resulted in the development of GaAs crystals having very high resistivity, on

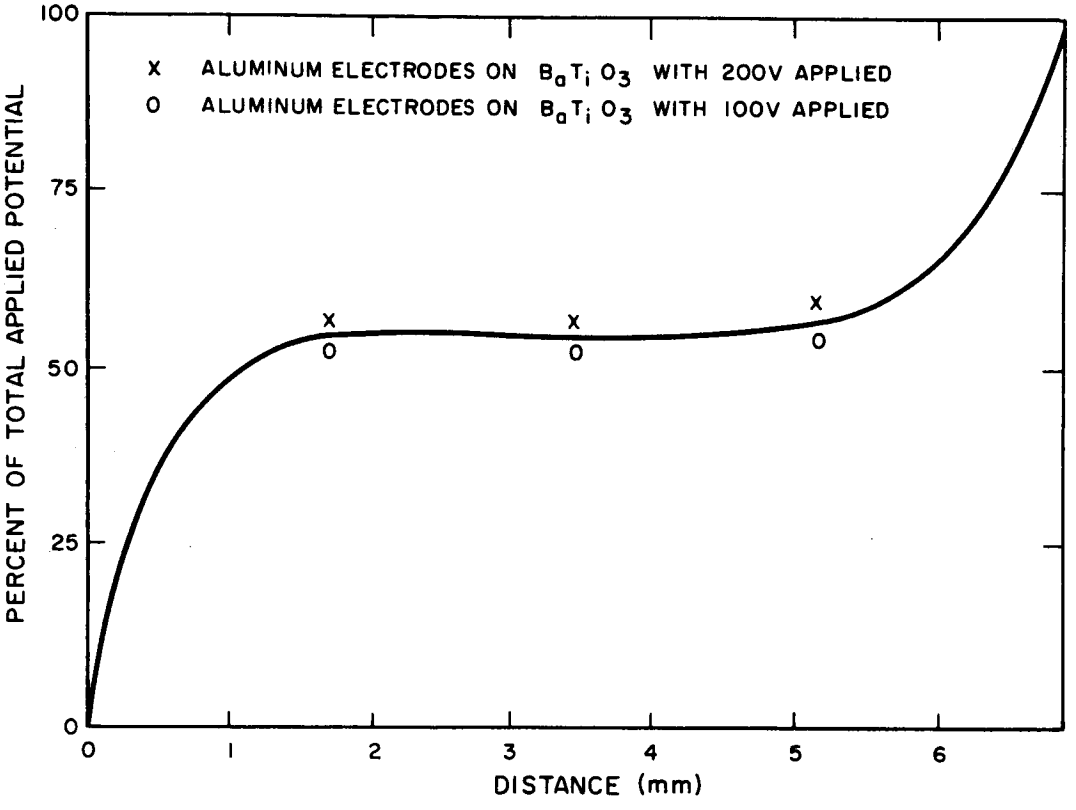


Fig. 40. Per cent of total applied potential vs distance across a 7-mm $BaTiO_3$ crystal.

the order of 1 megohm-cm. This technical breakthrough converted GaAs from a laboratory curiosity into a practical device for electro-optic modulation.

The half-wave voltage of GaAs is comparable to that of CuCl at near IR wavelengths. Also, in contrast to CuCl, long (5 cm and more) non-hygroscopic crystals of GaAs having high resistivity can be grown and cut and handled without introducing strains. Clearly, of all the crystals listed in Table IV, GaAs is the best choice at the present state of the art.

The transmission characteristics of a GaAs crystal are given in Fig. 41. In this graph the transmission loss in the spectral band from 1.4 to 16 microns is negligibly low, practically all the attenuation being due to reflection at the ends of the crystal. Therefore, with antireflection coatings on its ends, the transmission of the crystal can be increased by approximately 50 per cent.

To investigate the quality with which an RCA-GaAs crystal passes an image at IR wavelengths, the experiment shown in Fig. 42 was performed. Light impinging on the numerals "56" and on some resolution lines was ultimately imaged on photographic film after passing through a GaAs crystal. The image converter was used to convert the IR image into a visible one. The results of this experiment are given in Fig. 43, which shows a photograph of the image at the output face of the crystal. (The crystal used in this experiment was the same as that used in the TV experiments described in Secs. IV.D.2.a above and Sec. V.D below.) It is evident from Fig. 43 that GaAs is transparent at IR frequencies with good quality transmission. Note the clarity of the numerals and the sharpness and definition of the resolution lines.

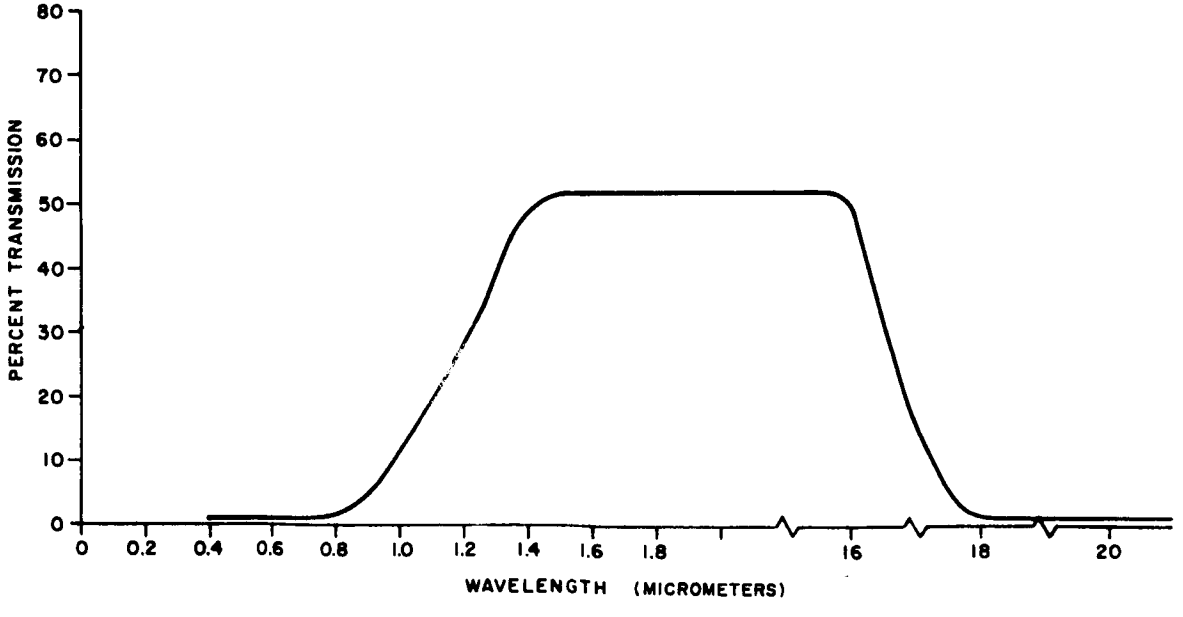


Fig. 41. Transmission characteristics of GaAs crystal (dimensions 1 cm x 3 mm x 3 mm).

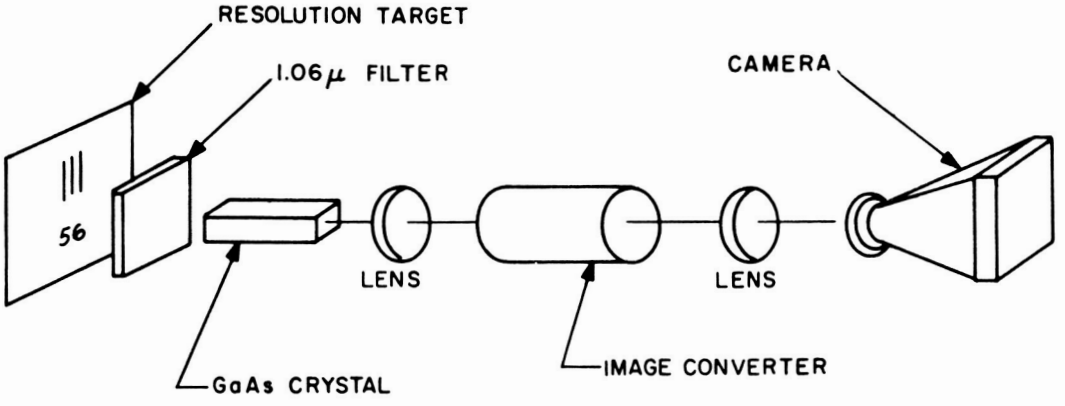


Fig. 42. Experimental setup for investigating transparency of GaAs.

Fig. 43. Image of numerals seen through a GaAs crystal.

In summary, a careful review of available electro-optic crystals has revealed that, compared to the other crystals, GaAs offers the following advantages:

It can be grown 5 cm long.

It has reasonably low drive requirements.

It is transparent in a spectral range where high power lasers are available.

It is nonhygroscopic.

It will not develop strains under normal handling.

It does not require temperature control.

A summary of its physical constants (from Table IV) follows:

Crystal type .....Valence

Crystal structure .....Cubic

Index of refraction.....3.34 (λ = 0.78μm to 8μm)

Dielectric constant.....11.2

Loss tangent .....Less than 0.0003

Electro-optic coefficient .....10^{-10} cm/volt

Spectral range of transparency...0.9 μ to 15 μm

3) Optimum Crystal Length

The analysis given in Appendix N shows that the optimum length of an electro-optic crystal is $\ell = -1/\ln \gamma$, where γ is the optical transmission of the crystal.[1] Increasing the crystal length beyond the optimum value will increase the per cent modulation (for a given drive voltage), but the increased attenuation through the crystal will more than offset the gain in per cent modulation.

[1] If shot noise caused by received signal power is greater than shot noise caused by received background power and detector dark current, then the optimum length is $\ell = -2/\ln \gamma$

4) Relationship between Deviation Ratio (or Modulation Bandwidth) and Signal-to-Noise Ratio

As shown in Appendix N, the SNR at the output of an FM receiver is proportional to both the square of the modulation bandwidth and the square of the per cent amplitude modulation of the subcarrier. If the available crystal drive voltage is not sufficient for 100 per cent amplitude modulation, then the per cent modulation will be inversely proportional to modulation bandwidth (or amplifier bandwidth). For this case the SNR improvement afforded by increasing the modulation bandwidth will be offset by the reduction of amplitude modulation of the subcarrier. In other words, if 100 per cent amplitude modulation cannot be realized, increasing the modulation bandwidth will not affect SNR.

At the present state of the art, it is difficult to achieve 100 per cent amplitude modulation with an information bandwidth of 5 MHz because the useful length of electro-optic crystals is restricted. As shown in the previous section, long GaAs crystals can be grown, but their transmission characteristics at the wavelengths where suitable wideband detectors and high-power lasers are available restrict their useful length to about 1 cm. Development of wideband photodetectors and high-power lasers that will operate efficiently at wavelengths of 2 micrometers or longer would change this situation, allowing the full benefits of wideband FM to be realized. Likewise, development of an electro-optic crystal similar to GaAs but having good transmission characteristics at wavelengths where multiplier phototubes can be used effectively will also allow the full benefits of wideband FM to be realized.

3. Wideband Electro-Optic Modulator

In the experiments described in Sec. V, the hybrid tube-transistor drive-amplifier shown in Fig. 44 was used to modulate the GaAs electro-optic crystal. The combination of tubes and transistors was used in order to minimize weight yet provide maximum possible drive voltage for the GaAs crystal. The schematic of this drive amplifier is shown in Fig. 45.

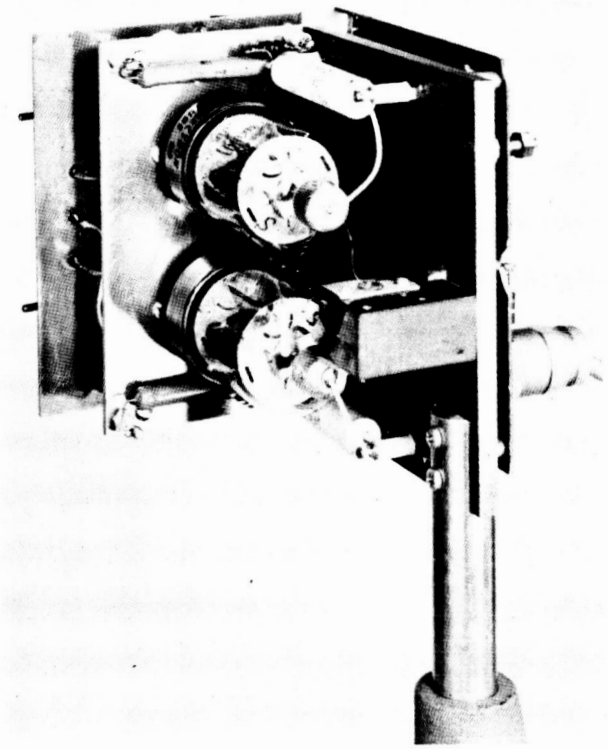


Fig. 44. Electro-optic modulator.

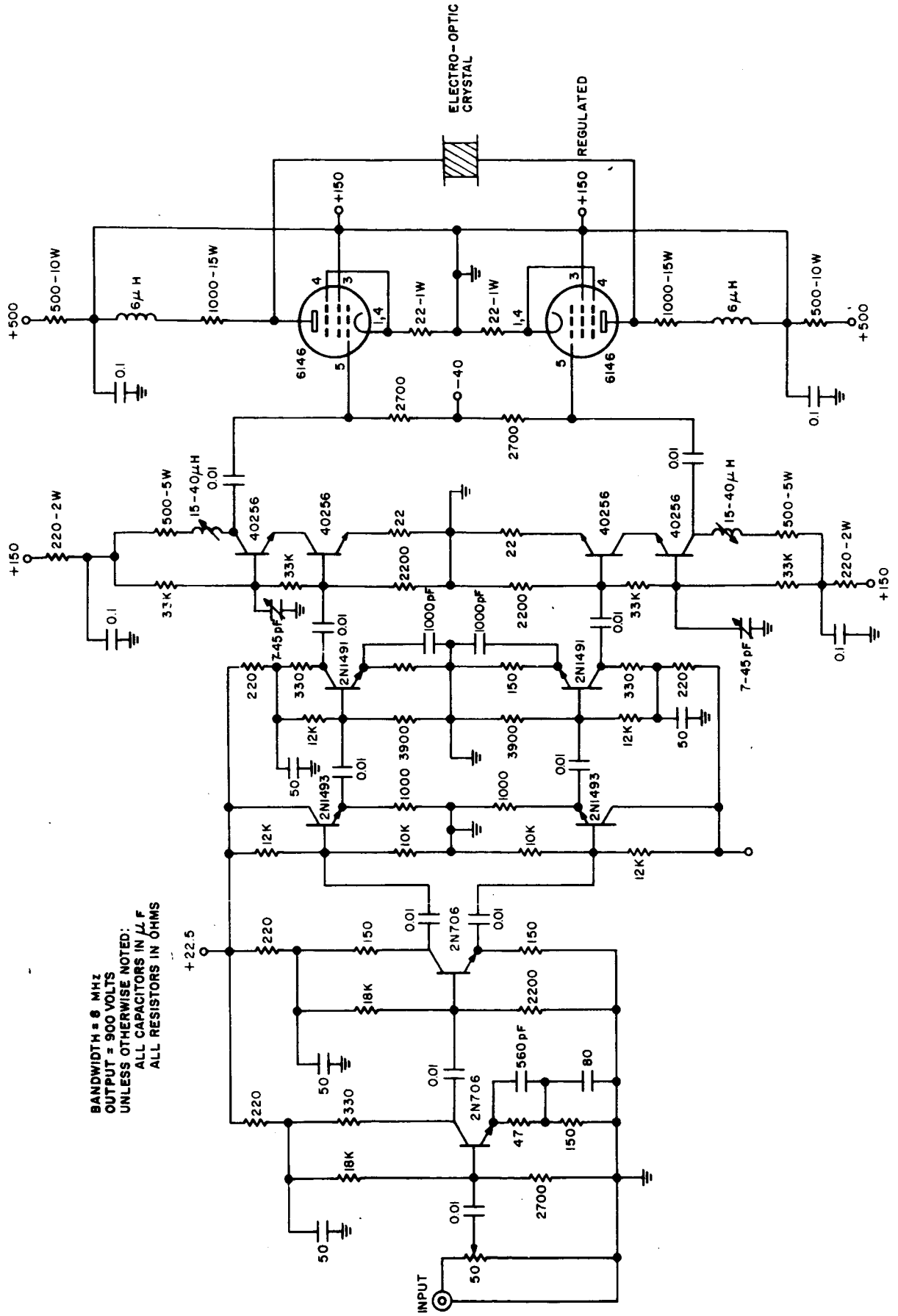


Fig. 45. Electro-optic drive amplifier.

E. BEAM EXTRACTION, COLLIMATION, DIRECTING, AND AIMING

When the beam of light emerges from the laser, it is unmodulated and travels along the axis of the solar energy collector. Several operations must be performed on the beam before it can be made useful as an information carrier, namely:

(1) The beam must be extracted from the mount supporting the solar energy collector and laser.

(2) The beam must be collimated and reduced in size so it can be directed through the electro-optic modulator crystal.

(3) After modulation, the beam must be recollimated to the narrow divergence required for the transmitted beam.

(4) The beam must be aimed toward the receiving station.

1. Beam Extraction

For a solar-pumped laser in which pump energy enters the end of the laser crystal, the beam emerges from the laser along the axis of the solar energy collector. Since the collector always points toward the Sun, the beam from the crystal is either directed at or away from the Sun. If the laser is to be used as part of a communication system, provision must be made for directing the beam to the receiving point. This involves extracting the beam from the mount holding the solar energy collector and laser crystal along some direction fixed with respect to the Earth or a space vehicle (depending on the application) and then directing the beam to the desired point.

The method of beam extraction to be employed depends upon the type of mount used. For an Earthbound system (such as System 1 of Table I in Sec. II. B) an equatorial mount is conveniently used to aim the energy collector at the Sun. Such a mount is similar to that used for the existing experimental system, as illustrated in Fig. 46. In this configuration, the beam from the laser follows the path indicated, being directed by reflections from six appropriately placed prisms.

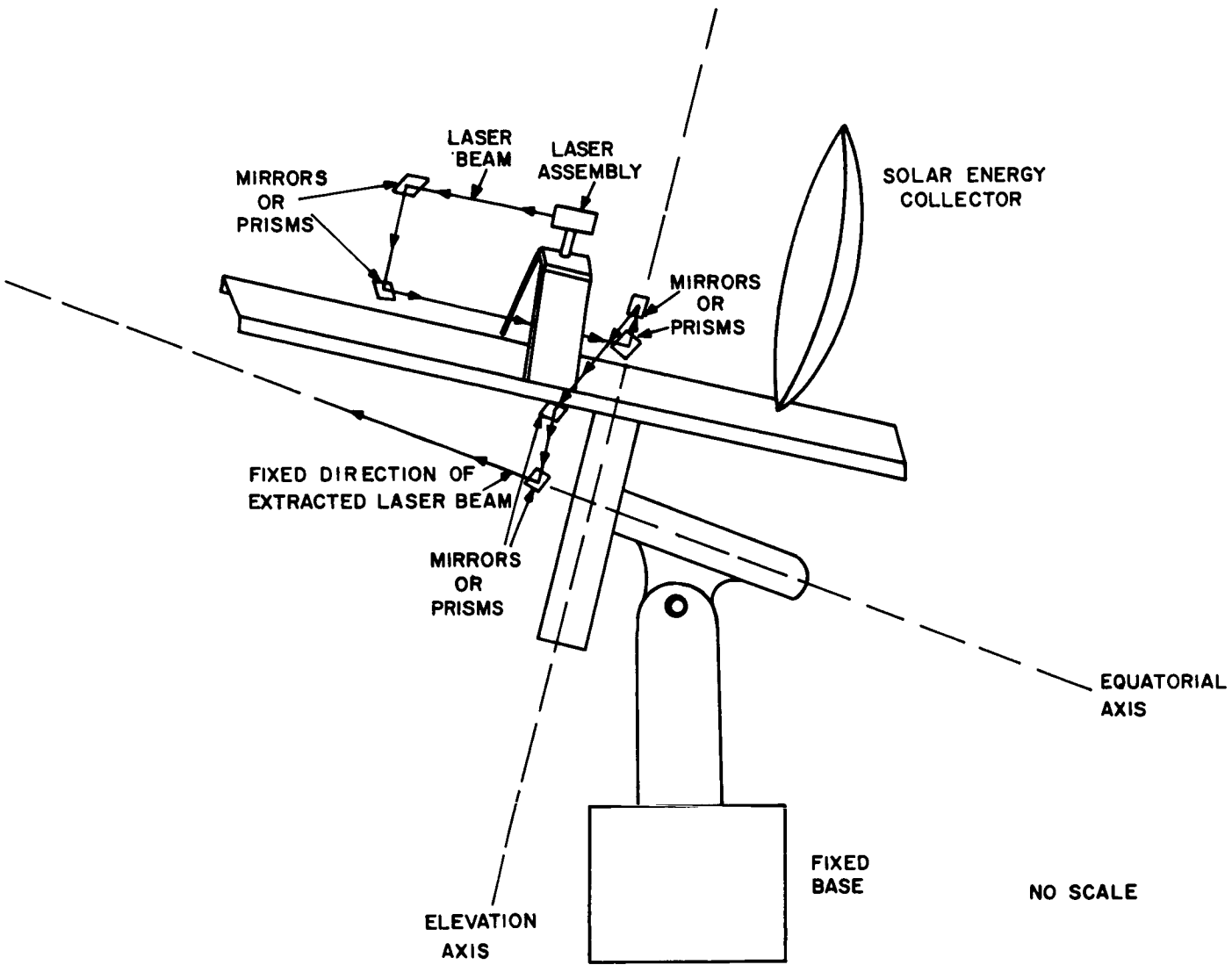


Fig. 46. Beam extraction from an equatorial mount.

For this particular arrangement, the beam emerges from the mount along the equatorial axis, the direction of which is fixed with respect to Earth. This extraction scheme is rather complicated due to the large number of reflections involved and was designed for use on a mount with no provisions for such extraction. A simplified extraction method for use on a mount which allows passage of the beam along the equatorial and elevation axes is shown in Fig. 47. Only two reflectors need be used in this case to extract the beam along the equatorial axis.

For spacecraft-mounted systems (such as Systems 2, 3, and 3′ of Table I in Sec. II. B), the energy collector and laser can be supported by an azimuth-elevation type mount. A simple scheme for beam extraction from this type mount is shown in Fig. 48. In this case, the beam emerges along the elevation axis which is fixed with respect to the spacecraft. Again, only two reflectors are used.

2. Beam Collimation

A scheme to pass the beam through an electro-optic modulator and to collimate it to a desired divergence is shown in Fig. 49. In order to specify the properties of the lenses and their relative locations in such a configuration, the following notation is needed, where the letters A, B, C, D, and E represent individual lenses and their positions:

δ_ℓ = intrinsic divergence of laser beam (rad)
δ_{AB} = divergence of beam in region between lenses A and B (rad)
δ_m = divergence of beam passing through modulator (rad)
δ_t = divergence of transmitted beam (rad)
d = diameter of laser crystal (m)
L_{xy} = distance between location x and location y (m)
F_x = focal length of lens x (m)

If the laser crystal has reflectors which are confocal, lens A is placed such that its focal point coincides with the center of the laser crystal at O. This

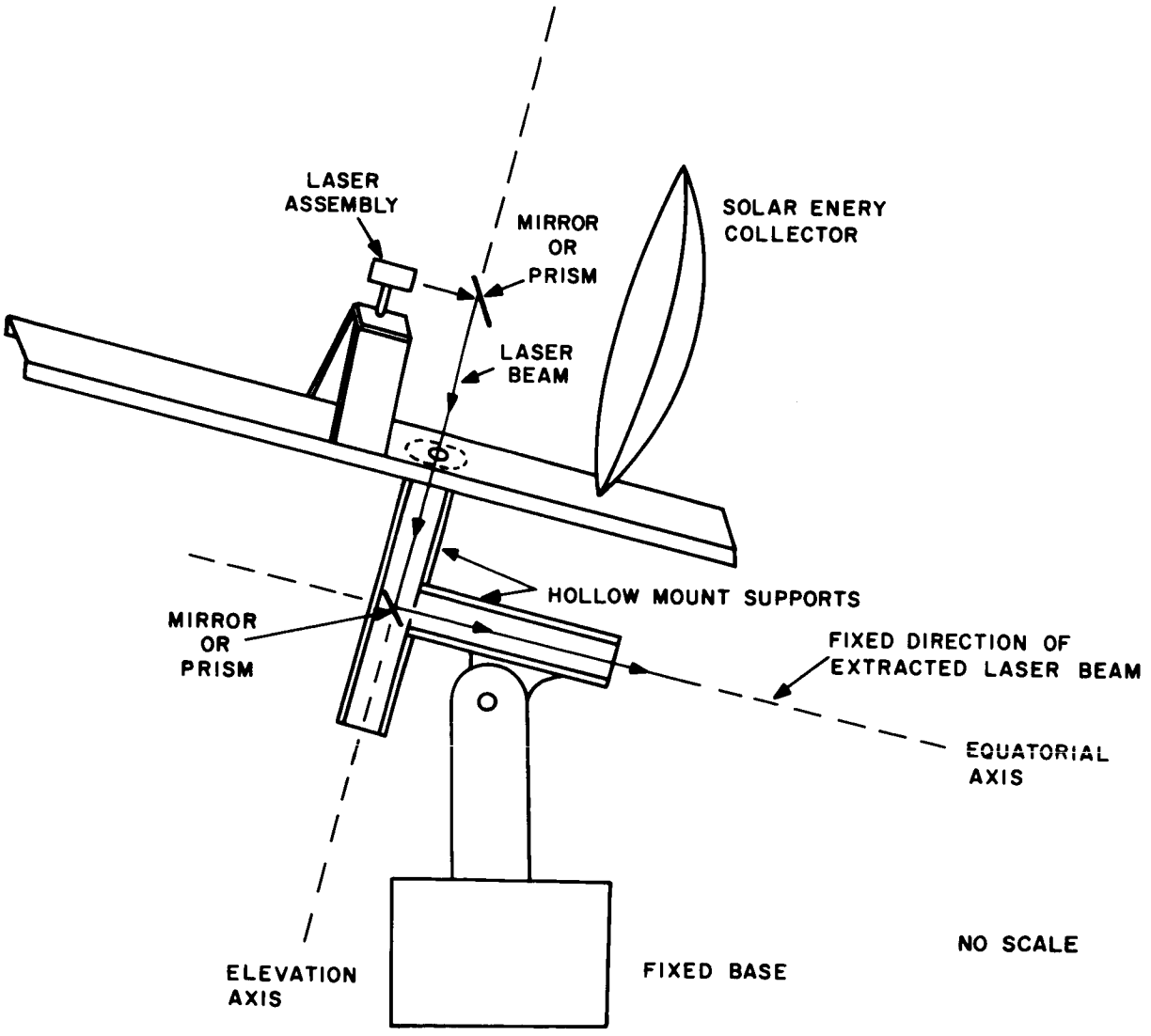


Fig. 47. Simplified beam extraction method for use on an equatorial mount.

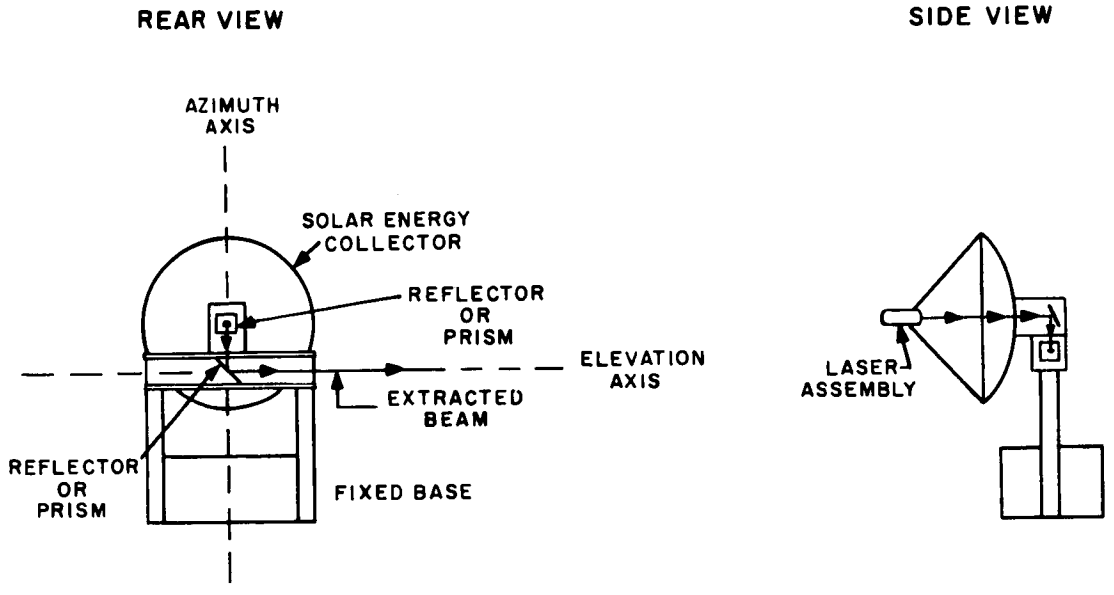


Fig. 48. Beam extraction from an azimuth-elevation mount.

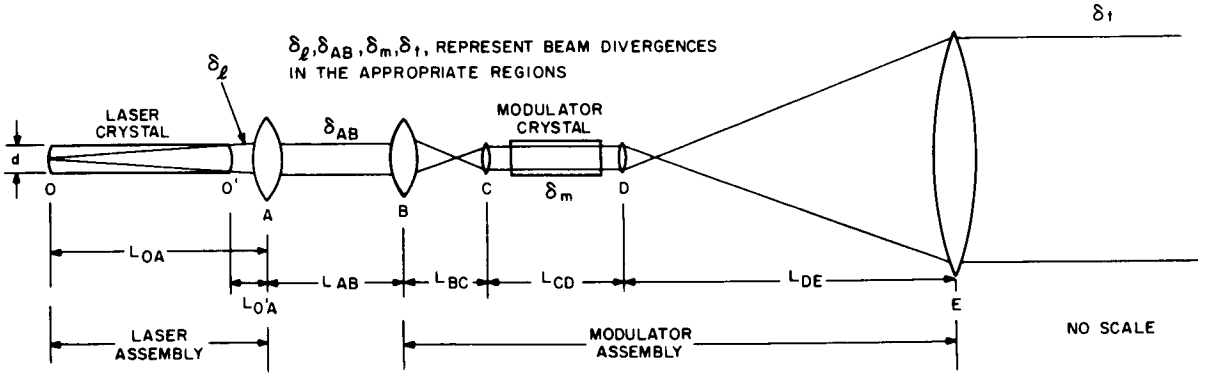


Fig. 49. Optics required to achieve modulation and collimation.

ensures that the divergence δ_{AB} is minimized and will approach the diffraction limited beam divergence as a limit. Lens A must have a minimum diameter of $d + L_{O'A}\,\delta_{\ell}$ in order to intercept the entire beam. (In general, the diameter of a beam with divergence δ increases by $L\delta$ in traveling over a distance L.) The beam then proceeds from the "laser assembly" (consisting of the laser crystal and lens A) over the distance L_{AB} to the "modulator assembly" (consisting of the modulator crystal and lenses B, C, D, and E). On arrival at lens B, the beam diameter is:

$$\text{Beam diameter at B} = d + L_{O'A}\,\delta_{\ell} + L_{AB}\,\delta_{AB}\,. \tag{81}$$

The minimum diameter of lens B is equal to the beam diameter at B.

It will now be assumed that the beam diameter must be reduced by a factor R before it passes through the modulator crystal to achieve efficient modulation. Thus,

$$R = \frac{\text{diameter of beam after lens C}}{\text{diameter of beam after lens B}}. \tag{82}$$

To accomplish this reduction in diameter, certain restrictions are placed on the focal length of lenses B and C, namely:

$$\frac{F_C}{F_B} = R \tag{83}$$

$$F_B + F_C = L_{BC} \tag{84}$$

If R and L_{BC} are specified, the focal lengths F_B and F_C and the minimum diameter of lens C are determined:

$$F_B = \frac{1}{1+R}\,L_{BC} \tag{85}$$

$$F_C = \frac{R}{1+R}\,L_{BC} \tag{86}$$

$$\text{Minimum diameter of lens C} = R\left(d + L_{O'A}\,\delta_{\ell} + L_{AB}\,\delta_{AB}\right). \tag{87}$$

The value of R is chosen to make the beam diameter less than the thickness of the modulator crystal. L_{BC} is chosen to satisfy the practical limitations on the size of the optical assembly and the ability to obtain optics with the required combinations of focal length and diameter.

After passing through the distance L_{CD}, wherein the modulator is contained, the beam arrives at lens D with a diameter given by:

$$\text{Beam diameter at lens D} = \text{Minimum diameter of lens D}$$

$$= R\,(d + L_{O'A}\,\delta_{\ell} + L_{AB}\,\delta_{AB}) + L_{CD}\,\delta_{m} \tag{88}$$

The modulated beam must now be recollimated to the divergence δ_t. This requirement demands that the focal lengths of lenses D and E be related as follows:

$$\frac{F_D}{F_E} = \frac{\delta_t}{\delta_m} \tag{89}$$

$$F_D + F_E = L_{DE} \tag{90}$$

Since δ_t is assumed known, only L_{DE} must be determined. L_{DE} is chosen so that the optical assembly is not excessively long. Finally, the minimum diameter of lens E can be found:

$$\text{Minimum diameter of lens E} = \frac{\delta_m}{\delta_t}\left[R(d + L_{O'A}\,\delta_{\ell} + L_{AB}\,\delta_{AB}) + L_{CD}\,\delta_{m}\right] \tag{91}$$

The above information may be summarized as shown in Table V. In this table, the results of the analysis of a typical example are listed along with the general system parameters defined above.

Table V. COLLIMATION SYSTEM PARAMETERS

General System Parameters	Typical System
L_{OA}	5 cm
$L_{O'A}$	1 cm
L_{AB}	100 cm
L_{BC}	2 cm
L_{CD}	10 cm
L_{DE}	131.04 cm
δ_ℓ	0.1 rad
δ_{AB}	3 mrad
δ_m	9 mrad
δ_t	100 urad
F_A (= L_{OA})	5 cm
F_B	1.6 cm
F_C	0.4 cm
F_D	1.44 cm
F_E	129.6 cm
R	1/4
d	0.4 cm
Diameter A (= $d + L_{O'A}\,\delta_\ell$)	0.5 cm
Diameter B (= $d + L_{O'A}\,\delta_\ell + L_{AB}\,\delta_{AB}$)	0.8 cm
Diameter C (= $R\,[d + L_{O'A}\,\delta_\ell + L_{AB}\,\delta_{AB}]$)	0.2 cm
Diameter D (= $R[d + L_{O'A}\,\delta_\ell + L_{AB}\,\delta_{AB}] + L_{CD}\,\delta_m$)	0.29 cm
Diameter E $\left(= \frac{\delta_m}{\delta_t} R\,\left[d + L_{O'A}\,\delta_\ell + L_{AB}\,\ell_{AB}\right] + \frac{\delta_m}{\delta_t} L_{cd}\,\delta_m\right)$	26.1 cm

3. Directing and Aiming the Beam

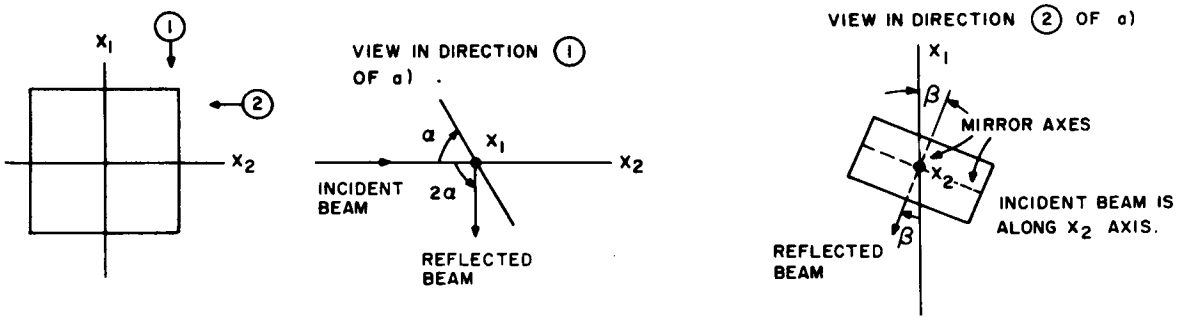
After the beam has been extracted from the mount supporting the solar energy collector and laser, the operations of aiming and directing the laser beam must be performed. These operations, while not independent, may be analyzed separately.

a. Directing the Beam

This operation involves the deflection of the beam from its extraction path to the target. Two suitable deflecting devices are the plane mirror and the Dove prism. Their use as deflectors may be described in the following manner:

1) The Reflecting Mirror

Consider the plane mirror shown in Fig. 50.



a) $\alpha = \beta = 0$ MIRROR AXES COINCIDENT WITH SPACE AXES x_1, x_2

b) ROTATION BY α ABOUT x_1 AXIS

c) ROTATION BY β ABOUT x_2 AXIS AFTER ROTATION ABOUT x_1 AXIS BY ANGLE α

Fig. 50. Plane mirror used as a beam deflector.

Assume that a beam of light parallel to the x_2 axis impinges on the mirror surface. According to the law of reflection, if the mirror is inclined to the x_2 axis by an angle α, and then rotated about axis x_2 by an angle β, the reflected beam is deflected by an angle 2α from its original path in the direction of the inclination about axis x_2 and by an angle β from its original path in the direction of the rotation about x_2. If the incident beam of light corresponds to the extracted laser beam, it is possible to direct it into "almost any point" in the 4π

steradian solid angle about the intersection of the axes x_1 and x_2 by allowing α and β to assume values in the ranges $0 \leq \alpha \leq \frac{\pi}{2}$, $0 \leq \beta < 2\pi$. The phrase "almost any point" must be used since it is not possible to deflect an entire beam of finite diameter at certain angles with a mirror of finite size. The relationship between mirror size and beam diameter is illustrated in Fig. 51. It is seen that if the diameter of the incident beam is 2ρ, and if the entire beam is to be reflected when the mirror is inclined by angle α with respect to x_2, then the mirror must have a minimum length $L(\rho, \alpha)_{min}$ of $2\rho \csc \alpha$ along its inclined axis and a length of 2ρ along its other axis (which lies along x_1). Fig. 52 is a plot of the ratio $L(\rho, \alpha)_{min}/2\rho$ vs angle of inclination and may be used either to determine a limiting angle α_o for a mirror with given dimensions $L (\geq 2\rho)$ and 2ρ, or to determine the value $L(\rho, \alpha_o)$ if 2ρ and α_o are specified. Thus,

$$\alpha_o = \csc^{-1}\left(\frac{L}{2\rho}\right) \tag{92}$$

$$L(\rho, \alpha_o) = 2\rho \csc \alpha_o \tag{93}$$

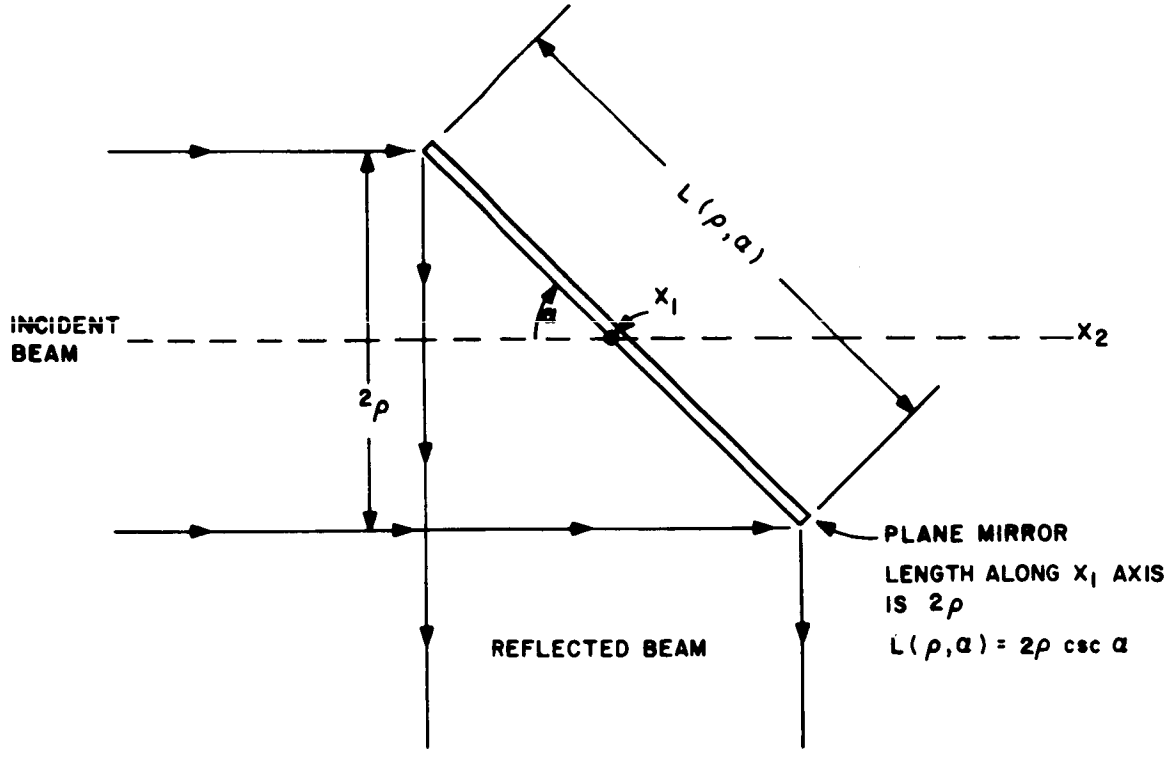


Fig. 51. Relationship between mirror size and beam diameter.

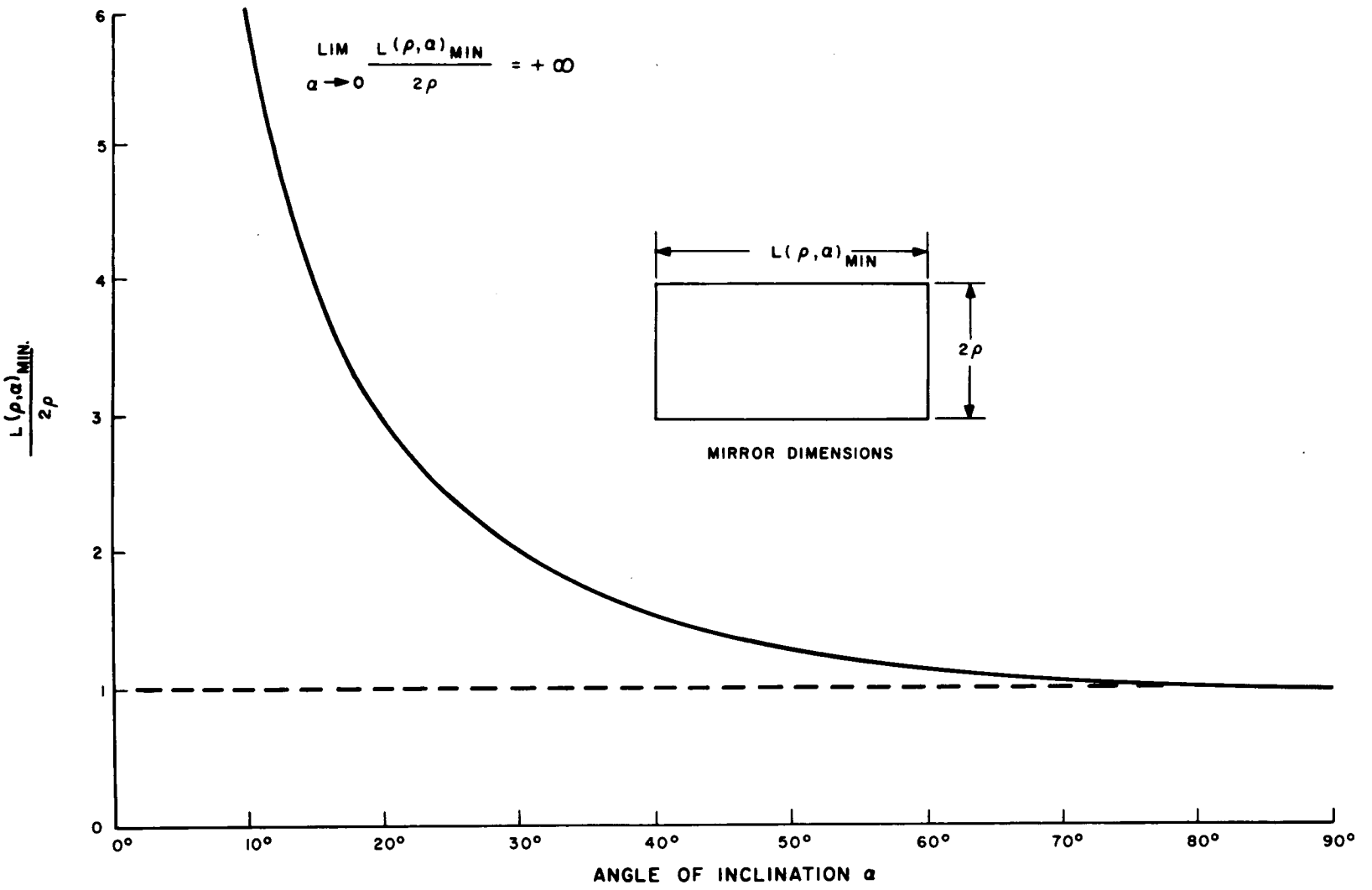


Fig. 52. $L(\rho, \alpha)_{min}/2\rho$ vs angle of inclination.

It is important to notice from Fig. 52 that as $\alpha \rightarrow 0$, the length of the reflector along its inclined axis must increase without limit if the entire beam is to be reflected, regardless of the value of 2ρ.

2) The Dove Prism

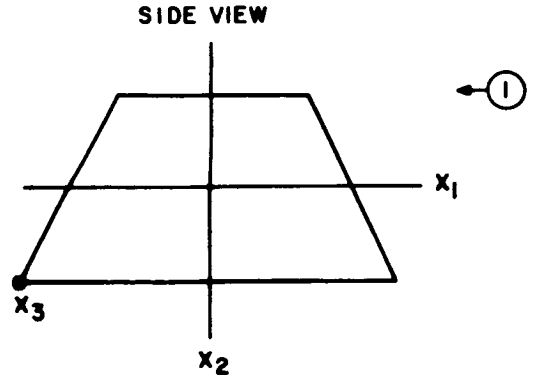
A Dove prism used as a beam-deflecting device is shown in Fig. 53. In this case, the laser beam (centered along space axis x_1) enters the prism through one end. If the prism is rotated about prism axis x_3 by an angle α and about axis x_1 by an angle β, the beam emerges from the opposite end of the prism along a line inclined by an angle β in the direction of the rotation about x_1 and an angle 2α in the direction of the rotation about x_3. It is possible to direct the beam into any point in the 2π steradians forward of the plane which contains the x_2 axis and is normal to the x_1 axis. In this case, straight through transmission $\sim \alpha = 0$ may be accomplished.

b. Aiming the Beam

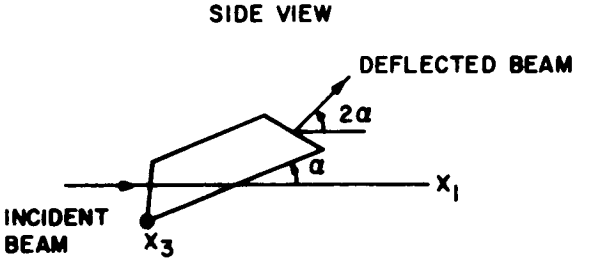
Thus far, explicit mention was not made of the method of controlling the position of the deflector used. For an Earthbound experimental system, the beam direction can be accomplished by manual positioning. On a space vehicle in motion, a continuously variable servo control will be necessary. In either case, it would be desirable to have a method of directly (visually) observing the target point of the deflected beam. The collimation system and deflectors described allow the introduction of an aiming device which will allow simultaneous visual observation of the target and the position of the transmitted laser beam with respect to the target. Such a sighting device is shown in Fig. 54. This sighting device functions as follows:

(1) The laser beam is directed onto the aperture of the beam splitter such that almost 100 per cent of the incident beam is passed.

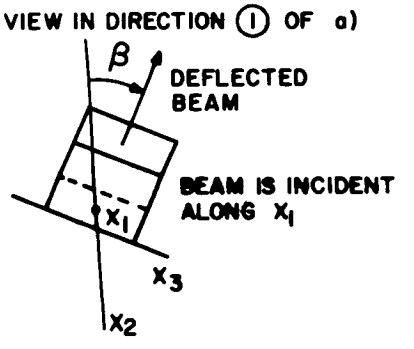
(2) The small fraction not transmitted is reflected onto M and then back through the beam splitter into the telescope in such a way that this reflected beam is perpendicular to the direction of the transmitted beam. (This requires an accurate alignment here.)



a) DOVE PRISM AND SPACE AXES X_1 AND X_2 AND PRISM AXIS X_3.



b) ROTATION BY α ABOUT PRISM AXIS X_3.



c) ROTATION BY β ABOUT X_1 SPACE AXIS AFTER ROTATION ABOUT PRISM AXIS X_3 BY α.

Fig. 53. Dove prism used as a beam deflector.

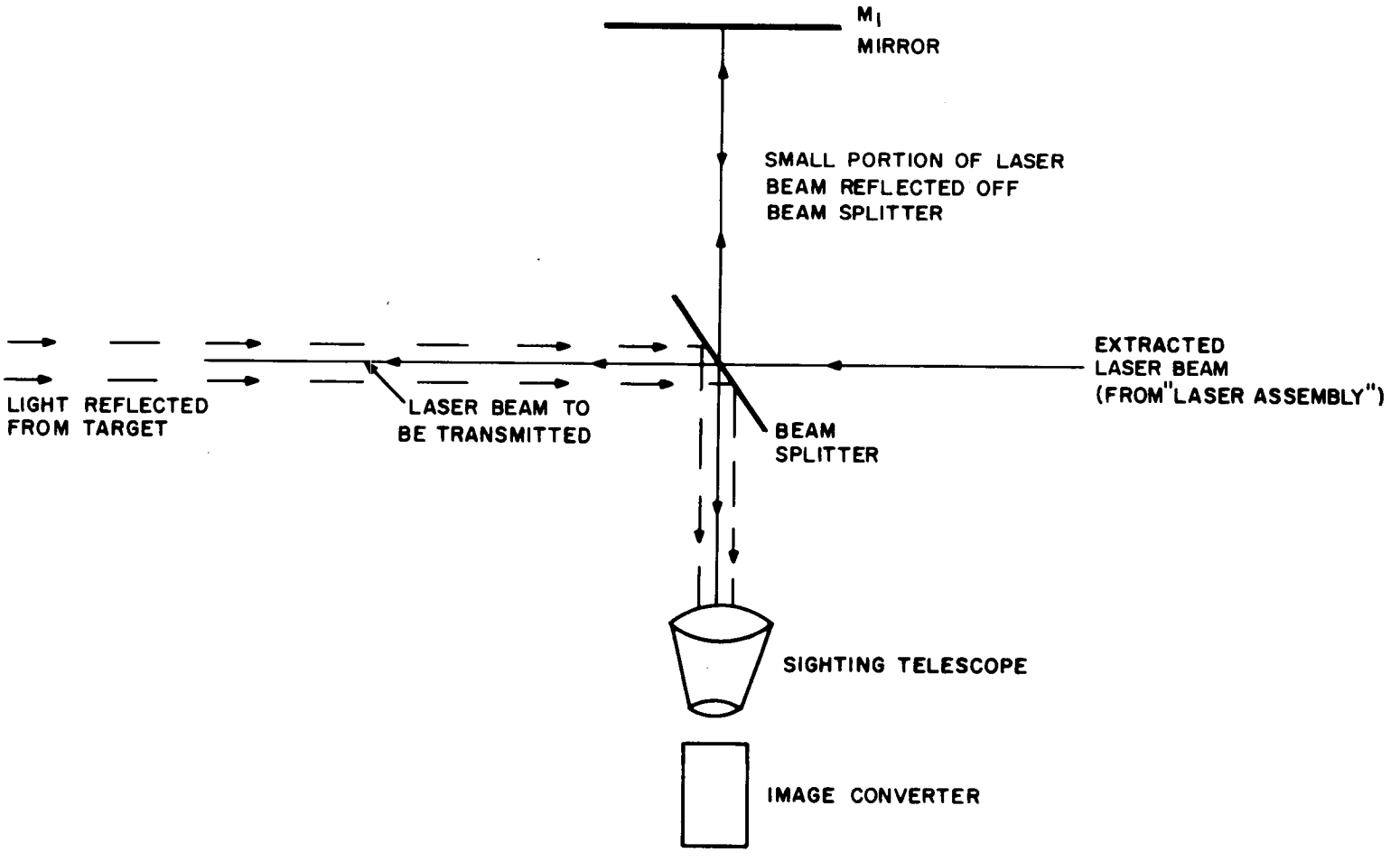


Fig. 54. Sight for observing coincidence of transmitted laser beam and target.

(3) The target, assumed illuminated, sends light to the transmitter site which is directed onto the beam splitter and then into the telescope.

(4) The superimposed reflected portion of the laser beam and target are viewed through the alignment telescope and image converter. This view shows the target of the transmitted beam.

Another possibility for a sighting device is an accurately positioned separate astronomical telescope which is boresighted with the transmitted laser beam. This method of sighting offers no built-in target-beam coincidence indication and the convenience of alignment solely at the transmitting station. These "disadvantages" do not, however, preclude the use of the boresighted telescope.

c. Some Practical Considerations

Either of the beam deflectors described can be incorporated into the collimation optics in either of the two accessible regions; i.e.,

(1) In the path L_{AB} — In this position the deflector size is minimal as a result of the small beam diameter (2ρ) here. Thus the weight of the deflection device is minimized. However, insertion of the deflector here requires that the massive "modulator assembly" portion of the collimation optics be movable. It is conceivable that this may not be desirable since it imposes rather severe requirements on the servo system which would control the position of the "modulator assembly."

(2) After the collimation optics — Here, the size of the deflector is determined by the size of the final collimation lens. The size of this lens is in turn determined by the divergence of the transmitted beam, large lenses being required for small values of beam divergence. Thus, for the case of a small divergence beam being considered, a relatively large deflector would be necessary. If a plane mirror is used, this size requirement can be satisfied rather easily (for limited range of inclination angle α). The mirror need not be excessively massive. On the other hand, if a large Dove prism is used, the mass involved becomes difficult to maneuver in a manner similar to that of a large movable "modulator assembly" as mentioned in (1) above.

These considerations indicate that a plane mirror inserted after the collimation optics is the best compromise between size and movability. The size limitations on the mirror due to its angular position can be overcome by appropriate positioning of the space vehicle.

The aiming system described above is most conveniently placed in the path L_{AB}, since in this position the sizes of the elements are minimal, and the magnification by the telescope formed by lenses B, C, D, and E can be used for furnishing an enlarged target image to the aiming system.

Section V

EXPERIMENTAL RESULTS

A. LABORATORY PUMPING

The preliminary tests of the lasers used in the solar-pumped experiments were performed in the laboratory, where ideal conditions (freedom from stray radiation, rigidity of apparatus) could be simulated. These tests were conducted to evaluate the best materials and techniques which would later be used in solar-pumped operation.

The lasers were tested in the pumping ellipse schematically shown in Fig. 55. Operation of this device followed the standard procedure of placing the laser at one focal axis and the proper pump lamp at the other focal axis.

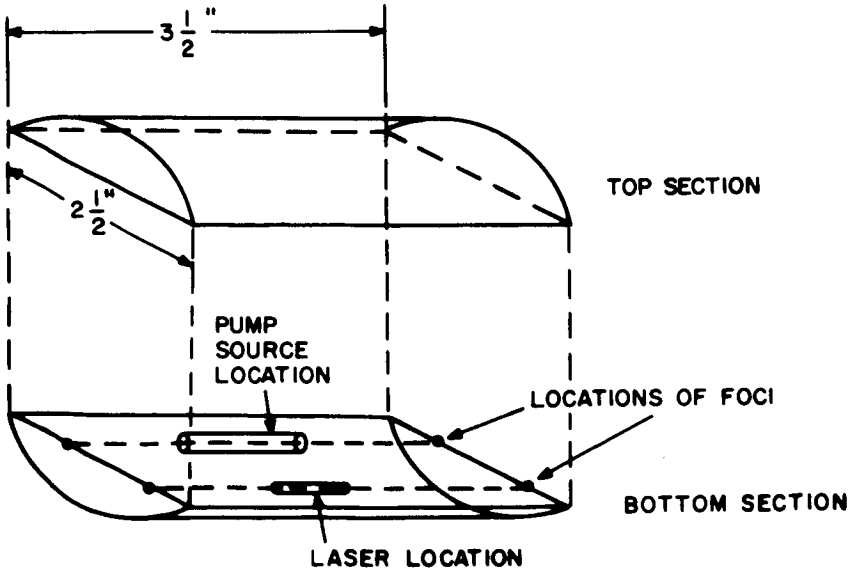


Fig. 55. Geometry of the laboratory pumping ellipse.

The types of lasers used in this test are conveniently divided into two classes: those which operate at temperatures ≤ 77°K ($CaF_2:Dy^{2+}$) and those which operate at 300°K (YAG: Nd^{3+} and YAG:Nd^{3+}-Cr^{3+}).

1. $CaF_2:Dy^{2+}$ Laser

Several lasers of this type were operated inside a linear Dewar inserted in the pumping ellipse. The experimental arrangement is shown in Fig. 56. In this setup, the $CaF_2:Dy^{2+}$ crystal is mounted in the center of the Dewar with supercooled (68°K) liquid nitrogen flowing over it. Sodium nitrite is pumped through a glass jacket surrounding the crystal to filter the ultraviolet radiation from the pump lamp. The Dewar is mounted at one focus of the elliptical cavity (shown with cover removed) and a 1 kW Sylvania DXN lamp (tungsten filament, quartz envelope, iodine vapor atmosphere) is mounted at the other focus. Both the position of the lamp within the cavity and the position of the cavity with respect to the Dewar are adjustable by means of control knobs located on the sides of the cavity assembly.

The laser output was detected with a fixed indium arsenide photodiode (Philco L-4530), the output of which was amplified and displayed on an oscilloscope. In this way, information about the temporal characteristics of the laser emission was obtained. The spatial distribution of the laser output was measured by displaying the amplified output of an indium arsenide photodetector, which was mounted on an X-Y micropositioner on the Y-axis of an X-Y recorder, the X-axis of which was calibrated in a position coordinate. The power output of the laser was measured by directing the laser beam into a Westinghouse laser radiometer (Type RN-1).

The above measurements were performed on several $CaF_2:Dy^{2+}$ lasers, some of which were reduced by a gamma-radiation process and others by the electrolytic process. Of those tested, the best was selected for use in a magnetic modulation experiment in the laboratory and for later use in the solar-pumped experiments. The crystal selected had the following characteristics:

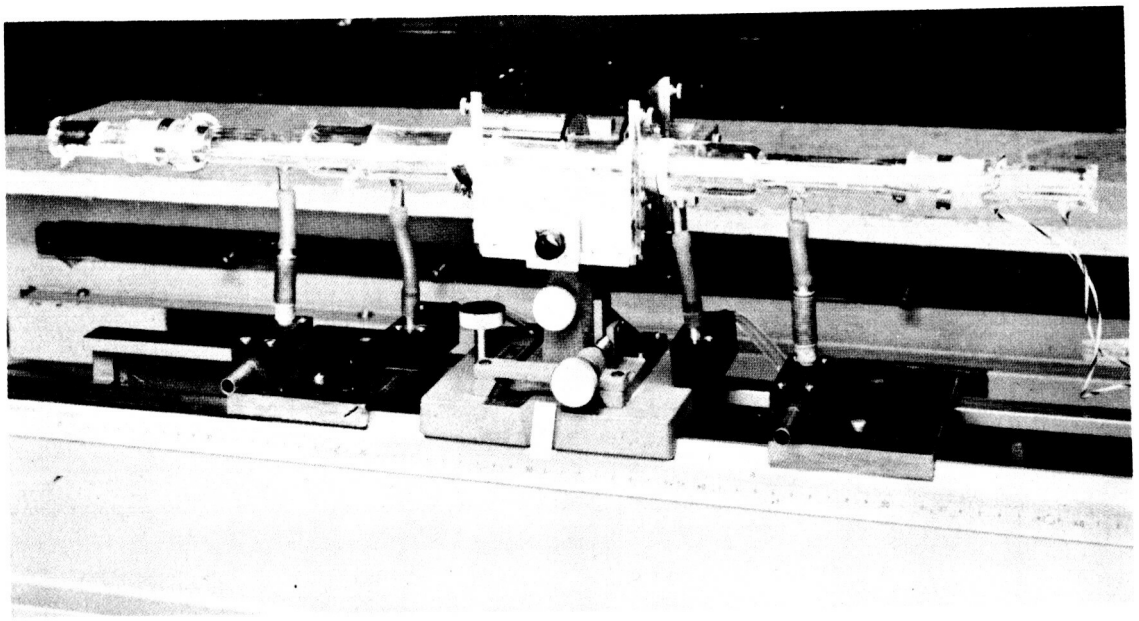


Fig. 56. Laboratory laser setup.

Dy^{2+} concentration	0.03 molar per cent
Length	1.0 inch
Diameter	1/8 inch
Power output	$\approx$ 40 mW
Threshold pump power (tungsten lamp pumping)	$\approx$ 300 W
Laser reflector type	spherical, 2.0 inch radius of curvature
Frequency of ripple on output	$\approx 10^4$ Hz
Linewidth ($\Delta\nu$)	3×10^9 Hz
Reduction method	electrolytic
Laser wavelength	2.36 μm

The laboratory magnetic modulation experiment consisted of placing the magnetic modulation coil, (a 20-turn three-layer solenoid, wound concentric with the crystal and centrally located along its length) and driving it with the pulse modulator electronics. The solenoid was 1/8 inch long, 1/4 inch in diameter, and had a measured inductance of 6 μH. It was found that 100 per cent modulation could be achieved with a current of one ampere in the coil. Modulation at an audio frequency rate (voice) was achieved. Detection of the modulation was accomplished by using the indium arsenide detector as the input element of a PFM receiver.

2. YAG:Nd^{3+} and YAG:Nd^{3+}-Cr^{3+} Lasers

These lasers were operated in the same ellipse used for the CaF_2:Dy^{2+} experiments. The ellipse was used in essentially the same manner as described in Fig. 55. Since these experiments could be conducted at a temperature of 300°K, the Dewar system was replaced by a "flow tube" which allowed the laser crystal to be force-cooled by pure water. Fig. 57 illustrates the arrangement of the apparatus used in the 300°K experiments.

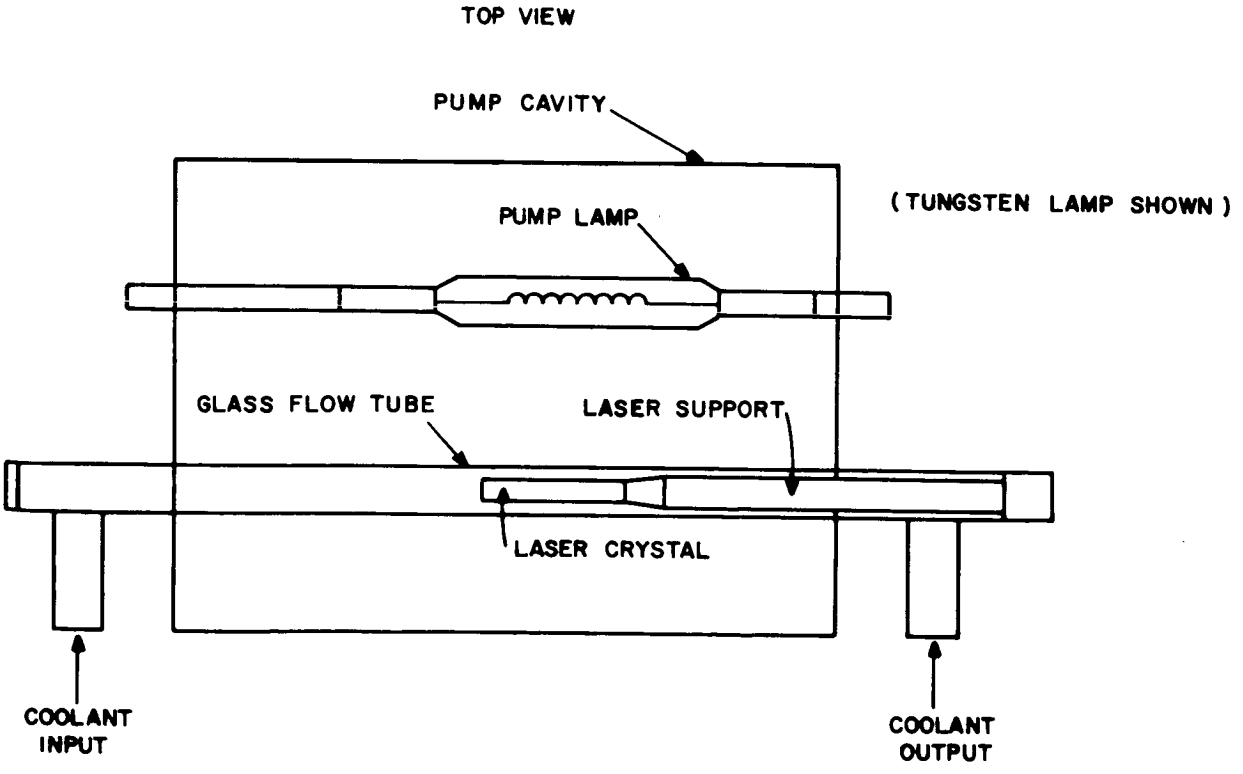


Fig. 57. Arrangement of apparatus for 300°K experiments.

The pump source for the YAG:Nd^{3+} lasers was the 1-kW Sylvania DXN tungsten-iodine vapor lamp used for pumping CaF_2:Dy^{2+}. The source for the YAG:Nd^{3+}-Cr^{3+} lasers was a 1-kW General Electric A-H6 mercury vapor lamp. This selection of lamps was based on knowledge of the laser absorption spectra and the emission spectra of available lamps.

The detector which was used in the 300°K laboratory experiments was an RCA 7102 multiplier phototube. The output of this detector was displayed on an oscilloscope.

Successful laboratory operation of the YAG:Nd^{3+}-Cr^{3+} laser was achieved, and the temporal characteristics of the outputs were studied. Results of these experiments indicated that the double-doped crystals would probably be unsuitable as CW carrier generators in modulation experiments. (This was later found true in solar-pumped experiments described in Sec. B below.) Much spiking was observed in the output.

It was not possible to operate the YAG:Nd^{3+} laser in the laboratory. This was not unexpected since the narrow absorption spectrum of the laser did not match the broad emission spectrum of the tungsten lamp well, and the 1-kW power output of the lamp did not permit the laser to absorb enough effective pump power. No further attempt was made to operate single-doped YAG lasers in the laboratory. It was decided that these tests could best be performed during solar-pumping operation where the spectral match between laser and pump source was more favorable.

Power output of the YAG:Nd^{3+}-Cr^{3+} lasers was to be measured with the Westinghouse RN-1 laser radiometer, but an explosion of an A-H6 lamp irreparably damaged the pump cavity and such measurements could not be performed in the laboratory.

Two laser crystals were selected for use in solar-pumped operation. One was the YAG:Nd^{3+}-Cr^{3+} laser which exhibited the least spiking and the other was a

YAG:Nd^{3+} laser which was known to be of good quality. These crystals had the following characteristics:

YAG:Nd^{3+}-Cr^{3+} Laser

Nd^{3+} concentration	1.3 atomic per cent
Cr^{3+} concentration	1.3 atomic per cent
Length	1.0 inch
Diameter	0.094 inch
Threshold pump power (Hg-lamp pumping)	600 watts
Laser reflector type	spherical
Frequency of ripple on output	5×10^4 Hz (deep spiking, which is pump power dependent)

YAG:Nd^{3+} Laser

Nd^{3+} concentration	1.3 atomic per cent
Length	1.25 inches
Diameter	0.094 inch
Threshold pump power	not attained
Laser reflector type	spherical

B. OPERATION OF THE SOLAR-PUMPED LASER MODEL

1. The CaF_2:Dy^{2+} Laser

In these experiments, the Dewar assembly used in the laboratory experiments was mounted on the optical bench affixed to the equatorial mount in the manner shown in Fig. 58.

This laser is side pumped. A rectangular solar image formed by the cylindrical lens is directed onto the crystal side. The same type of cooling (supercooled liquid nitrogen) used in the laboratory experiments was used in solar-pumped operation. Also, the same Philco L-4530 indium arsenide photodetector

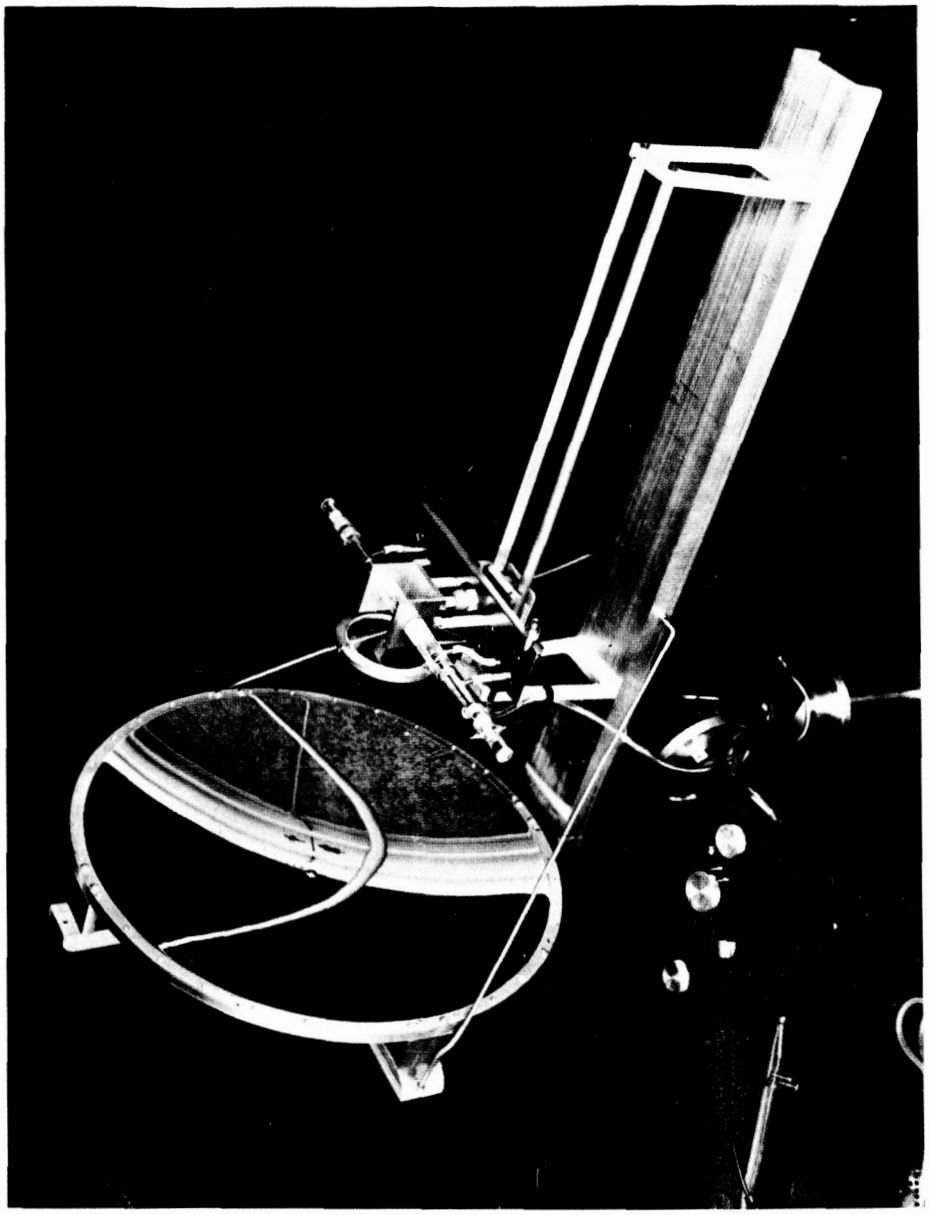


Fig. 58. Arrangement of apparatus for solar-pumped operation of $CaF_2:Dy^{2+}$.

was used to observe the laser characteristics. Reliable operation of the laser was obtained and it was found that the results of solar-pumped operation were very similar to those obtained in the laboratory. Approximately 80 to 100 W of solar power were needed to reach threshold. Ripple frequency was 6×10^4 Hz.

The magnetic modulation coil and electronics were incorporated into the apparatus. A PFM receiver was connected to the output of the photodetector. The first attempt at magnetically modulated solar-pumped operation was successful, and quality of a voice transmission was good. This experiment was repeated several times and was found to give repeatedly good results.

2. YAG:Nd^{3+} and YAG:Nd^{3+}-Cr^{3+} Lasers

In these experiments, there was one major departure from the type of operation used in all previous experiments. These lasers were end pumped. It had not been possible to achieve end-pumped operation in any of the laboratory experiments due to the lack of a pump source of sufficient intensity and an efficient energy coupling technique. The mechanical arrangement of the Dewar assembly precluded end pumping in the CaF_2:Dy^{2+} solar-pumped experiments. End pumping was achieved by using one of two specially designed water-cooled pump cavities. (These "flow tubes" are described in detail in Sec. V. C below.) Use of these flow tubes allowed a better optical coupling between the laser and solar image than had previously been obtained.

In the first experiments, the YAG:Nd^{3+}-Cr^{3+} laser described above (Sec. V. A. 2) was used. Operation was successful and results similar to those obtained in the laboratory were obtained. The threshold of this laser was approximately 100 watts[1] of solar power. As expected, the output showed severe spiking behavior and did not offer much promise for success in the anticipated electro-optic modulation experiments. For this reason, further tests on this laser were not made.

The YAG:Nd^{3+} laser was operated in the flow tube used in the YAG:Nd^{3+}-Cr^{3+} experiment. An RCA phototube 7102 (with a 1.06 μm filter on its face) was needed to detect the laser output. Output of this laser showed a good DC level with a ripple (frequency $\approx 5 \times 10^4$ Hz) of approximately 20 per cent. The threshold of

---

[1] These and other pump power figures refer to collected solar power. Actual pump power coupled to the lasers is approximately 25% of these values, since the crystal end area was approximately one-fourth the size of the solar image. Pump power measurements were made by measuring the solar constant with an optical calorimeter (Optics Technology, Inc. model 600) and the actual exposed area of the solar energy collector. Pump power was varied by partially blocking the mirror aperture with circular rings.

this laser was found to be approximately 100 watts. Measured power output of this laser under good pumping conditions (clear, sunny day) was on the order of 90 milliwatts, making it the most powerful, most efficient laser tested. The properties measured indicated that this laser could be used as a source for electro-optic modulation experiments.

The modulating element which was to be used in an attempt at wideband modulation was a GaAs crystal. This material exhibits an electro-optic effect under the influence of relatively high electric fields. Typical dimensions of the crystals available are 1 inch x 3 mm x 3 mm. The crystal and the associated electrical and optical components required for electro-optic modulation experiments were placed in the laser system as shown in Fig. 59.

The purpose of the lenses was to direct a major portion of the diverging laser beam (due to spherical crystal ends) into the GaAs crystal, and then onto a small portion of the photocathode of the detector. The first of these was to ensure adequate depth of modulation and the second was to reduce the effects of background light on the photomultiplier operation.

A television signal (bandwidth = 5 MHz) was transmitted over the laser beam. Fig. 60 shows the result of a transmission on a nearly perfect day with a well adjusted experimental apparatus.

It was found that alignment of the elements of the optical system was critical. It was found somewhat difficult to obtain optimum alignment with the particular elements used and that a more rigid arrangement would be desirable. Despite this, reliable operation was achieved even after a complete disassembly and reassembly of the apparatus.

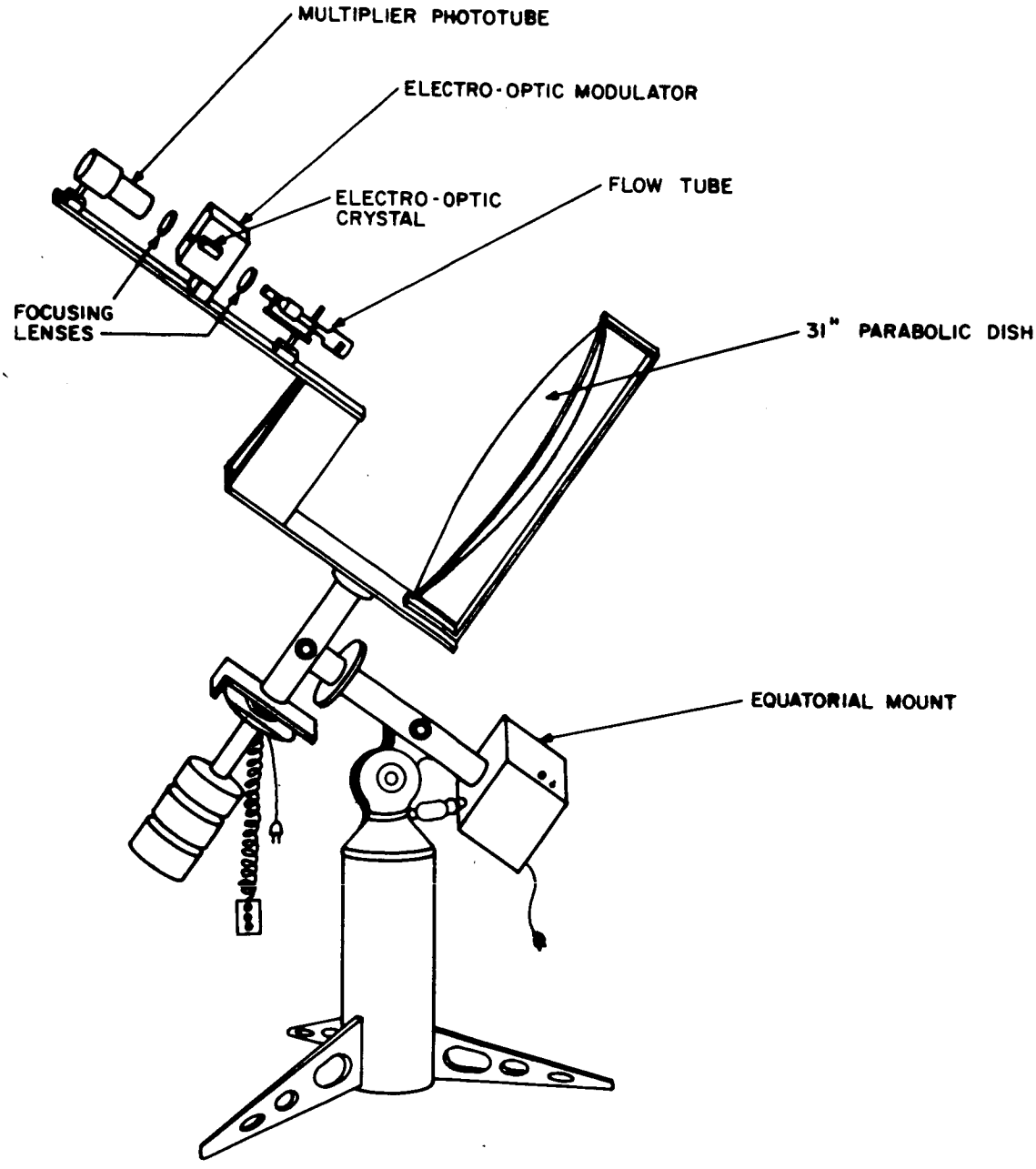


Fig. 59. Arrangement of apparatus for solar-pumped electro-optic modulation experiments.

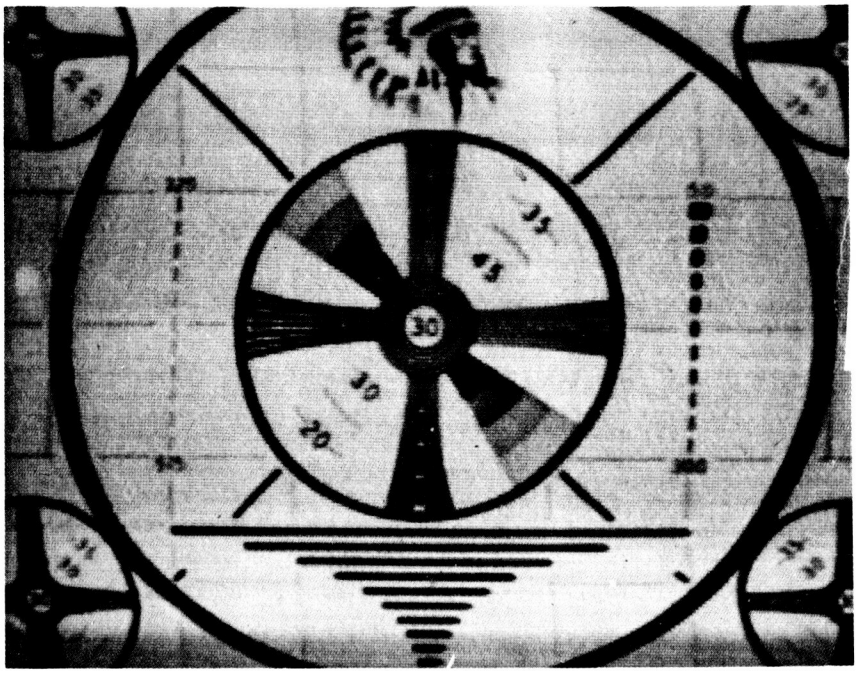


Fig. 60. Television picture transmitted via a solar-pumped laser.

C. FLOW TUBE CONSTRUCTION

The flow tube is used to accomplish two functions:

(1) Maintaining the YAG lasers at the proper operating temperature (300°K) under conditions of solar pumping.

(2) Serving as a pump cavity for the laser crystal.

The first flow tube which was used during the YAG laser tests described above in Sec. V. B is shown in Fig. 61.

The flow tube is made in essentially three sections:

(1) The crystal holder section

(2) The coolant section

(3) The crystal holder support section.

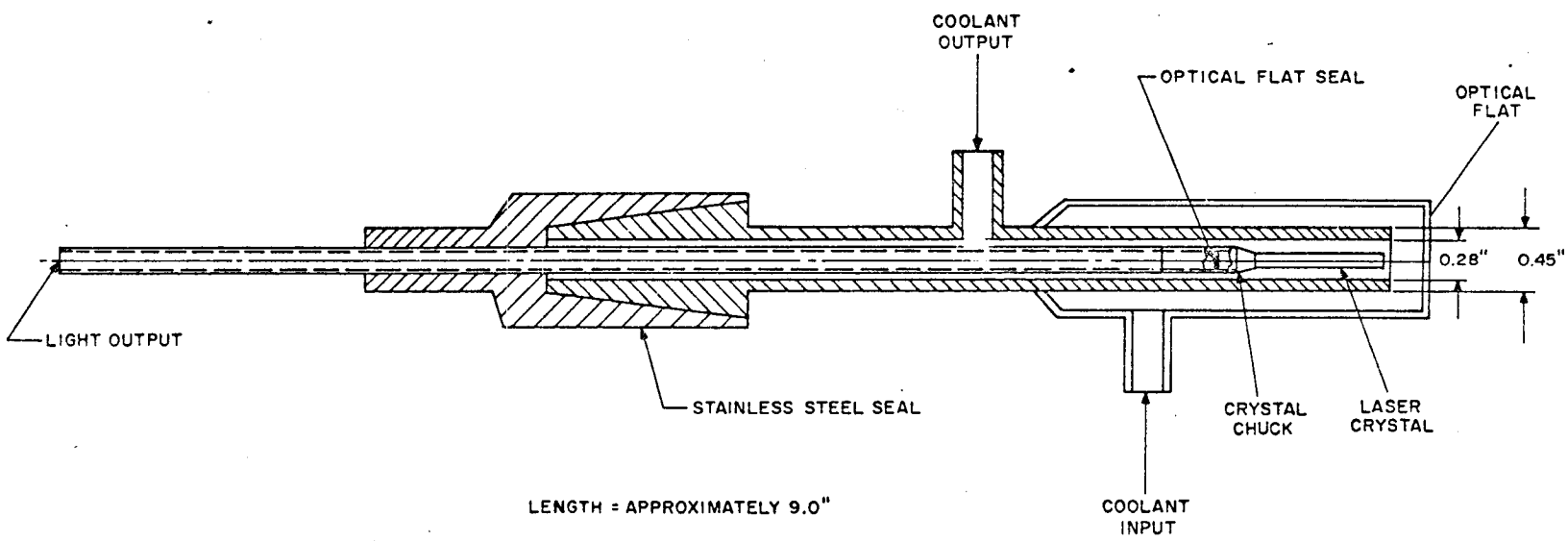


Fig. 61. First experimental flow tube.

The first of these has a small four-jaw chuck which holds the laser crystal. This section is in turn held by the third, and its position (adjustable) with respect to the third section determines the position of the laser in the coolant section. Cooling of the laser is accomplished by forcing cool water through the flow tube under pressure. The system used in the experiments performed was the closed system shown schematically in Fig. 62. Pure water was used to prevent damage to crystal end reflectors from particles contained in ordinary tap water. The water in the closed system was cooled by passing it through the copper-coil heat exchanger.

The coolant section also serves as the pump cavity. The inside walls are silver plated and diffusely reflect pump light entering the flow tube at other than normal incidence so that efficient use can be made of the pump power available. This reflection serves to effectively increase the path length for the solar pump radiation in the crystal.

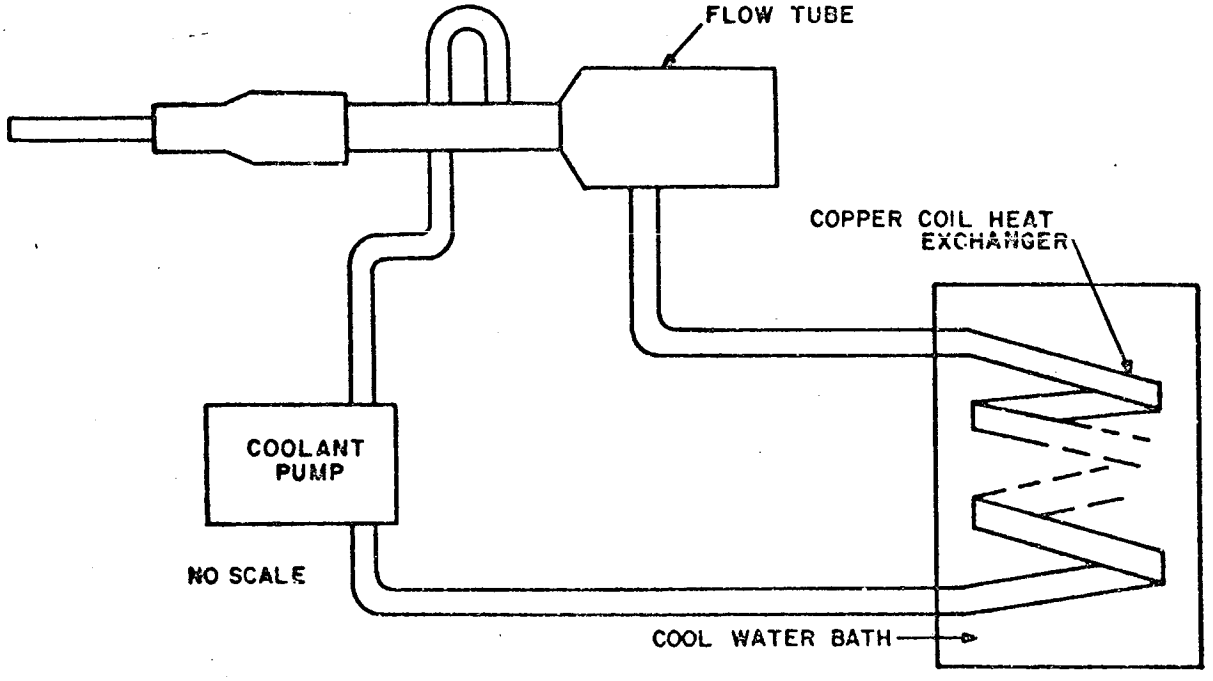


Fig. 62. Closed-system laser cooling.

The chief disadvantage of this flow tube is the long path which the diverging laser beam must follow (approximately 6 inches in a 1/8 inch diameter tube) before it can be utilized. In traversing this path, reflections from the side walls of the tube cause interferences among the various portions of the beam. The beam should be collimated as soon as possible after its emergence from the laser to ensure that a well-shaped beam is available for use, particularly when electro-optic modulation experiments are conducted.

In an attempt to overcome this difficulty, as well as test the pumping efficiency of two different-size pump cavities, a second flow tube illustrated schematically in Fig. 63 was constructed.

This tube allows the laser beam to be utilized after it travels approximately one-half inch from the laser crystal. Also, for certain values of laser reflector centers of curvature and focal lengths of the collimating lens, a collimated laser beam can be extracted directly from the flow tube.

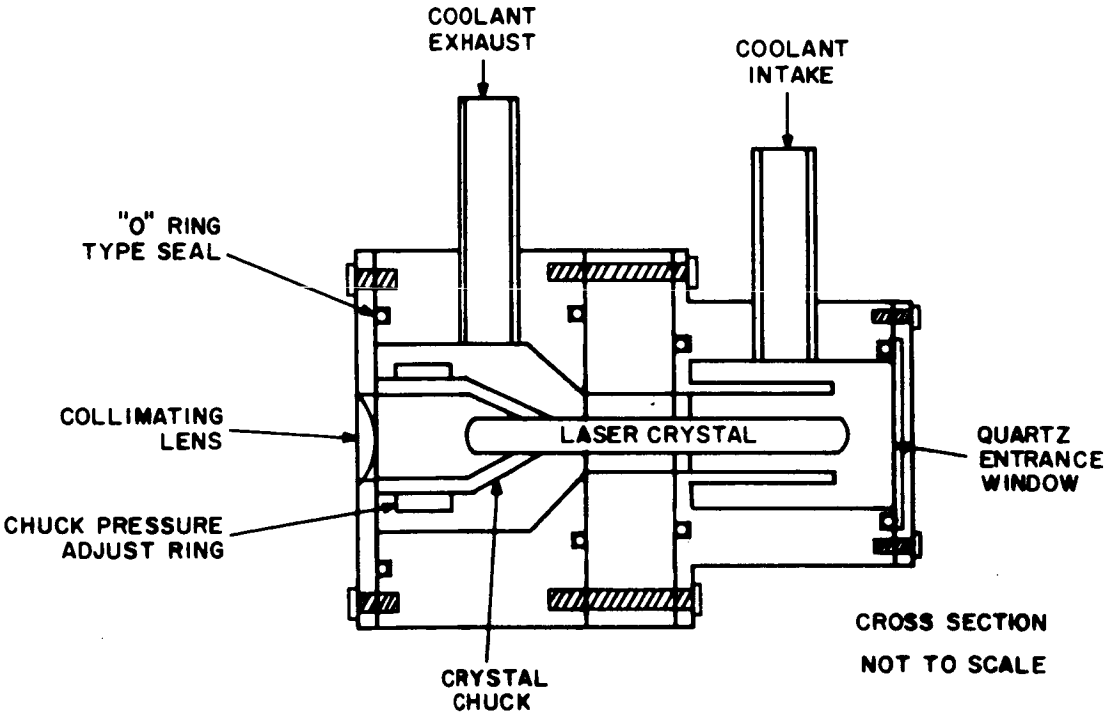


Fig. 63. Second experimental flow tube.

The YAG:Nd^{3+} laser was mounted in the second flow tube and was operated under solar-pumped conditions. It was found that the threshold of the laser was approximately the same as that previously obtained with solar pumping in the first flow tube. This indicated that the large diameter pump cavity was not necessary and that efficient pumping could be achieved in a cavity 3/16 inch in diameter. (The diameter of the first tube was 7/16 inch.)

D. ELECTRO-OPTIC CRYSTAL MODULATOR

Using the modulator circuit shown in Fig. 45 to drive a GaAs crystal measuring 3 mm x 3 mm x 2 cm with a 1000-volt (peak-to-peak) 5-MHz sine wave, 30 per cent modulation was achieved. Substituting $n = 3.34$, $r_{41} = 10^{-10}$, $d = 0.3$ cm, and $\ell = 2$ cm into Eq. 77 indicates that the half-wave voltage should have been $V_{\lambda/2} = 2140$ volts; and substituting $V_m = 500$ volts into Eq. 80 indicates that the per cent modulation should have been $m = 66$ per cent. The difference between the experimental and theoretical results can probably be attributed to strains within the crystal.

The modulating performance of the electro-optic modulator of Fig. 45 is illustrated in Fig. 64 where a light beam passing through it is modulated with a 5-MHz sine wave and then detected with a multiplier phototube.

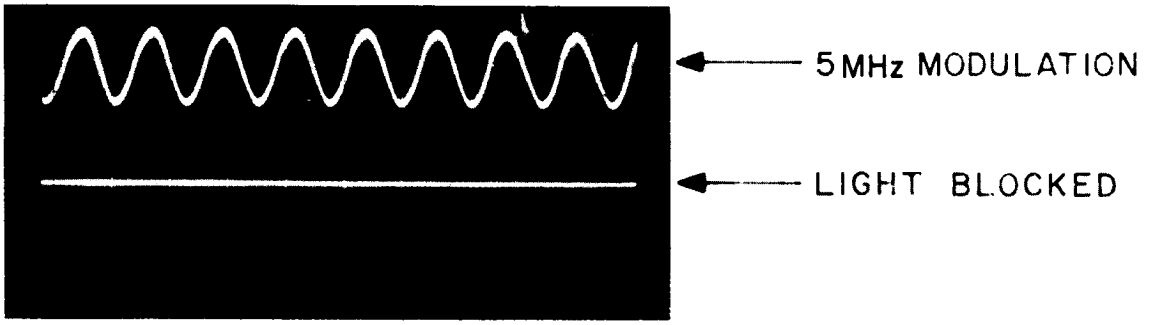


Fig. 64. Performance of electro-optic modulator.

Section VI

CONCLUSIONS AND RECOMMENDATIONS

1. Compared to other laser pumping methods, solar pumping offers a reduction of more than 6:1 in both the size and weight of the laser transmitter.
2. A television picture can be transmitted via a solar-pumped laser using an electro-optic modulator. A 5-MHz information bandwidth was demonstrated using a GaAs electro-optic modulator. FM-subcarrier modulation was used to avoid the effects of piezoelectric resonances in the GaAs crystal.
3. The relative merit of different wideband modulation methods depends upon the desired output SNR. If an extremely high SNR is needed, say 50 dB or higher, then a wideband modulation method such as PCM, which increases SNR exponentially with bandwidth expansion, is the best choice. But, if an SNR of only 30 dB is needed (as in the TV system considered here), other wideband modulation methods, such as wideband FM or delta modulation, provide equal, or better, performance with less complex circuits.
4. Extrapolating the state of the art to 1970 indicates that a solar-pumped laser operating at 1.06 μm could be developed having the following characteristics:

Power output	5 - 10 W
Transmitter beamwidth	10 μrad
Receiver beamwidth	10 μrad
Collector diameter	45 inches
Receiver area	1.81 m^2
Accuracy of aiming of transmitter and receiver	5 μrad
Information bandwidth	6 MHz

5. Work that should be pursued under a follow-up program includes the development of

   a. An automatic solar tracking system, employing a four-quadrant solar cell and servo controls.
   b. An optical system for collimating and pointing the laser beam.
   c. More efficient electro-optic modulator.
   d. A solar power supply for modulator, coolant pump and electronics.
   e. An RF-biased photoconductive detector.
   f. A CW beacon.

Appendix A
RELATIONSHIP BETWEEN PHOTODETECTOR OUTPUT CURRENT AND INCIDENT LIGHT POWER AND SIGNAL-TO-NOISE RATIO OF AN OPTICAL RECEIVER

The rms noise current from a photodetector is

$$I_n = \sqrt{2eIB}, \tag{A-1}$$

where

I = average diode current (amperes)

B = detector bandwidth (Hz)

e = charge on an electron (1.6×10^{-19} coulomb).

Received signal power, received background noise power, and leakage current contribute to the average current, I. The contribution due to received light (either signal or background) is

$$I = \eta pe, \tag{A-2}$$

where

η = quantum efficiency of photoemissive surface

p = average number of quanta received per second.

Since the energy per photon is

$$E = h\nu, \tag{A-3}$$

where

h = Planck's constant (6.62×10^{-34} joule-sec)

ν = frequency of received light (Hz)

then the average number of quanta received per second is

$$p = \frac{P}{h\nu}, \tag{A-4}$$

where P is the average received power. Substituting Eq. A-4 into Eq. A-2 gives

$$I = \frac{\eta e}{h\nu} P. \tag{A-5}$$

For a given photodetector, the term $\eta e/h\nu$ is a constant called the responsivity ρ. Hence, the output current can be expressed as

$$I = \rho P \tag{A-6}$$

This equation indicates clearly that a photodetector (e.g., a photodiode or multiplier phototube) is an ideal square law device; i.e., its output current is directly proportional to its input power.

Fig. A-1 shows the spectral responsivity of photosensitive surfaces. From this information, it is evident that S-1 surfaces have sensitivity of 0.3×10^{-3} A/W at the 1.06-micrometer wavelength of the YAG laser.

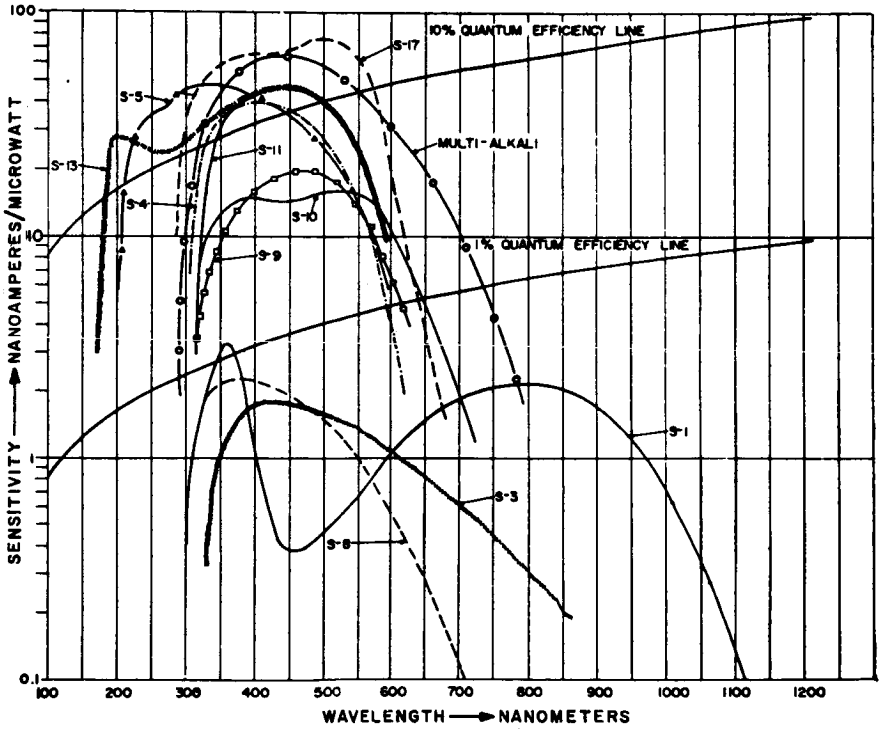


Fig. A-1. Spectral sensitivity of photosensitive surfaces.

Referring to the conceptual diagram of an optical receiver shown in Fig. A-2, we see that the average signal power and the average background power reaching the photodetector are P_s and P_b, respectively. Thus, according to Eq. A-6, the resultant signal and background noise currents are

$$I_s = \rho P_s \tag{A-7}$$

and

$$I_b = \rho P_b \tag{A-8}$$

The total average output current from the photodetector is equal to the sum of these two currents plus the dark current I_d, viz.,

$$I = I_d + I_s + I_b \tag{A-9}$$

Substituting Eq. A-9 into Eq. A-1 yields the rms shot noise current

$$I_n = \sqrt{2\, e\, (I_d + I_s + I_b)\, B}$$

where e is the charge on an electron and B is the bandwidth of the filter shown in Fig. A-2.

Thus, shot noise power N_s delivered by the photodetector to a load R_ℓ is

$$N_s = I_n^2\, G^2\, R_\ell = 2\, e\, (I_d + I_s + I_b)\, BG^2 R_\ell \tag{A-10}$$

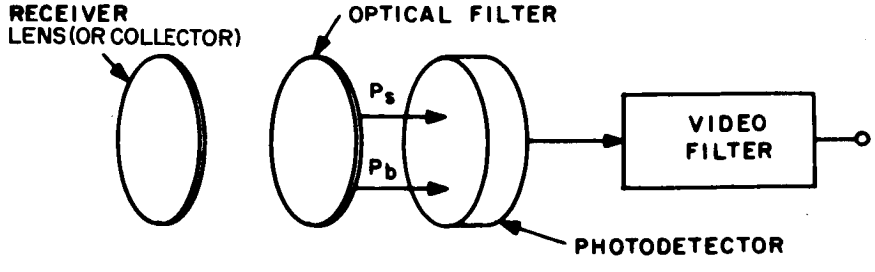


Fig. A-2. Conceptual diagram of an optical receiver.

and the signal power is

$$S = I_s^2 G^2 R_\ell = \rho^2 P_s^2 G^2 R_\ell \tag{A-11}$$

upon using Eq. A-6. In these equations, G is the internal gain (if any) of the photodetector.

Other sources of noise in an optical receiver are the preamplifier (if any) and the photodetector load resistor R_ℓ. The noise power introduced by the amplifier, expressed in terms of equivalent thermal noise at its input port, is

$$N_a = (F - 1)\, kTB \tag{A-12}$$

and the noise power introduced by the load resistor is

$$N_r = kTB \tag{A-13}$$

where

F = noise factor of amplifier

k = Boltzmann's constant (1.38×10^{-23} joule/°K)

T = temperature (°K)

Combining Eqs. A-10, A-11, A-12 and A-13, the signal-to-noise ratio at the output of an optical receiver (before demodulation) is

$$\frac{S}{N} = \frac{S}{N_s + N_a + N_r}$$

$$= \frac{\rho^2 P_s^2 G^2 R_\ell}{2\, e\, B\, (\rho P_s + \rho P_b + I_d)\, G^2 R_\ell + FkTB} \tag{A-14}$$

Appendix B

RELATIONSHIP BETWEEN TRANSMITTED AND RECEIVED POWER IN AN OPTICAL COMMUNICATION SYSTEM

After passing through focusing optics, the energy released from a laser transmitter is radiated in a cone-shaped beam of angular width equal to α_t radians. For small angles (i.e., $\alpha_t < 10$ degrees), this corresponds to a solid angle of $(\pi/4)\ \alpha_t^2$ steradians. Thus, if the total transmitted power is P_t, the power transmitted per steradian is $P_t/(\pi/4)\ \alpha_t^2$. As shown in Fig. B-1, an optical receiver will intercept some of this power over a solid angle of A_r/R^2 steradians, where A_r is the area of the receiver collector and R is the transmission range. In addition, as this power passes through the atmosphere and receiving optics (including optical filter), it will be attenuated by the factors T_a and T_o, respectively. Hence, the received signal power is given by

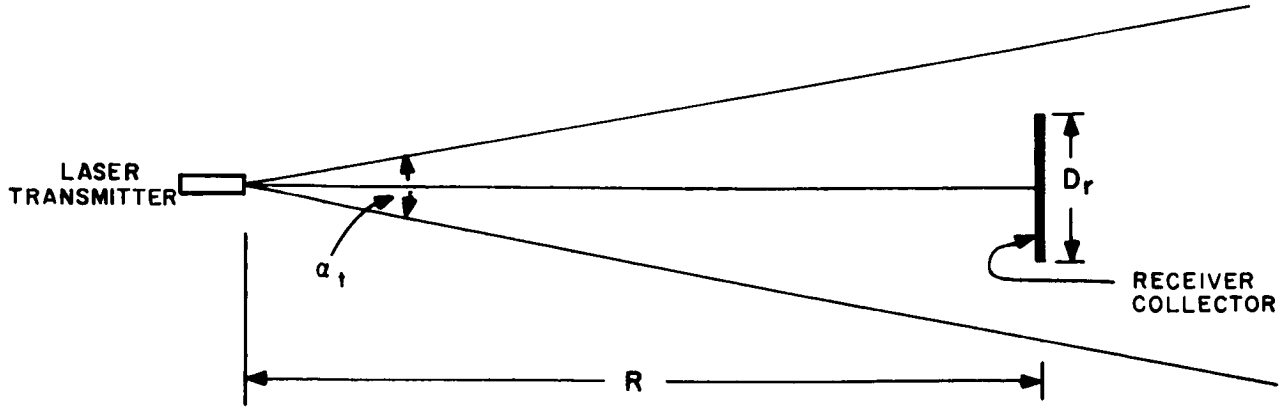


Fig. B-1. Geometry of transmission between laser transmitter and optical receiver.

$$P_s = T_a T_o \frac{P_t}{(\pi/4)\,\alpha_t^2} \frac{A_r}{R^2} \tag{B-1}$$

If the receiver collector is circular with diameter D_r, then Eq. B-1 becomes

$$P_s = T_a T_o \frac{P_t}{(\pi/4)\,\alpha_t^2} \frac{(\pi/4)\,D_r^2}{R^2} \tag{B-2}$$

or

$$P_s = T_a T_o \frac{D_r^2 P_t}{\alpha_t^2 R^2} \tag{B-3}$$

Solving for P_t here yields an expression for the transmitter power required for a given received signal power in an optical communication system:

$$P_t = \frac{\alpha_t^2 R^2}{T_a T_o D_r^2} P_s \tag{B-4}$$

Appendix C

BACKGROUND NOISE

The major source of background noise for a spaceborne receiver pointed at Earth is reflected sunlight. A reasonable approximation for the radiant emittance of the sun is obtained by assuming that the sun radiates like a blackbody at a temperature of 5600° K. From Planck's law, the radiant emittance of a blackbody (radiant flux emitted per unit area) within a range of wavelength $d\lambda$ is

$$L = \left\{ C_1 \lambda^{-5} \Big/ \left[\exp\left(C_2/\lambda T\right) - 1 \right] \right\} \, d\lambda, \qquad \text{(C-1)}$$

where

$C_1 = 2\pi C^2 h = (3.74 \times 10^{-12}$ cm^2-joule/s)

$C_2 = hC/k = (1.438$ cm-$^{\circ}$K)

T = temperature ($^{\circ}$K)

λ = wavelength (cm)

$d\lambda$ = bandwidth of optical filter (cm).

For example, the radiant emittance within a 10-angstrom bandwidth centered at a wavelength of 1 micrometer is $L = 3.3$ watts/cm^2/Å (or 3.3×10^4 watts/meter^2/Å). The Sun, being a diffuse radiator having a radiance of L/π watts/meter^2/steradian, irradiates the Earth with

$$M = \pi \left(\frac{L}{\pi}\right) \sin^2 \alpha$$

$$= L \sin^2 \alpha \text{ watts/meter}^2/\text{Å}, \qquad \text{(C-2)}$$

where α is the angle subtended by the Sun (as seen from the Earth). Since the Sun subtends 0.0046 radian of arc, the irradiance, M, is about 0.1 watt/meter^2/Å.

The reflected sunlight that is detected is termed background noise. The Earth, being a diffuse reflector, has a radiance of (ξ M/π) watts/meter^2/steradian/Å, where ξ is the reflection coefficient (or reflectivity). Since the ground area "seen" by the receiver is $\pi \alpha_r^2 R^2/4$ (where $\pi \alpha_r^2/4$ is the receiver beamwidth in steradians and R is the slant range), the effective radiant emittance of the Earth is (ξ M/π) ($\pi \alpha_r^2/4$) watts/meter^2/Å. Hence, because the receiver collects radiation over a surface A_r, where A_r is the collection area of the receiving antenna, the received background power is

$$P_b = \frac{\xi T_a T_o A_r B_{opt} M \alpha_r^2}{4} , \qquad \text{(C-3)}$$

where T_a and T_o are the transmission of the atmosphere and target receiver optics, respectively, and B_{opt} is the bandwidth of the optical filter which precedes the photodetector.

Appendix D

THRESHOLD DETECTION LEVEL

Communication range is limited by the threshold detection level of the received signal. In particular, the threshold detection level marks a received signal level below which output signal-to-noise ratio deteriorates rapidly. For example, consider the threshold detection level of a system employing binary modulation. Fig. D-1 illustrates the relationship between probability of error (i. e., the probability of mistaking one binary state for another) and signal-to-noise ratio. This relationship has been verified both theoretically and experimentally.[1]

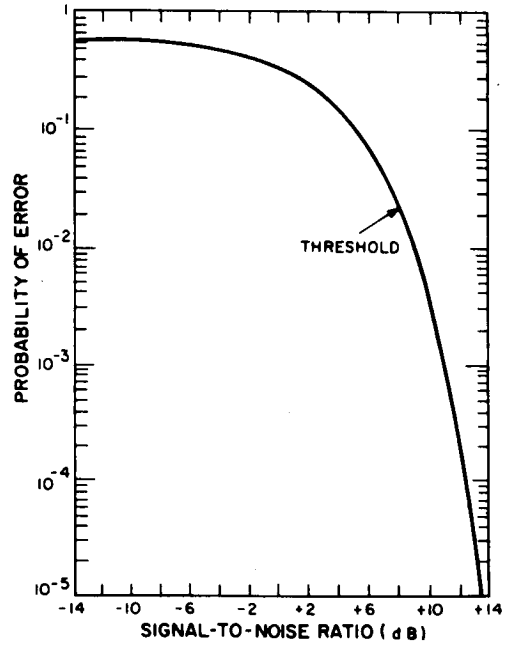


Fig. D-1. Probability of error vs signal-to-noise ratio for a communication system. Note that the threshold detection level occurs at about 9 dB.

[1] J. G. Lawton, "Comparison of Binary Data Transmission Systems," Proc. Nat'l Convention on Military Electronics, pp. 54-61, 1958.

Note in this figure that the threshold detection level occurs at a signal-to-noise ratio of about 9 dB. At 9 dB the probability of error is 0.01, whereas at 12 dB (just a 3-dB increase in signal-to-noise ratio) the probability of error is only 0.0001, a 100:1 reduction. From this result it is concluded that as long as the received signal-to-noise ratio (at the input to the receiver demodulator) is greater than 9 dB, the performance of a digital communication system will be essentially error free. This means that the maximum communication range is that range at which the signal-to-noise ratio has dropped to 9 dB (or, approximately, a power ratio of 10:1). Stated another way: above this signal-to-noise power ratio of 10, such a system provides a great improvement over the performance of narrowband systems (e.g., AM), while below this threshold, the wideband system performance deteriorates rapidly; i.e., it does not perform as well as an analog AM system.

Although shown specifically for a digital modulation technique, this threshold effect is inherent in all wideband modulation schemes, including wideband FM (see Appendix J). In the case of FM, the improvement over an AM system can be understood easily by means of the rotating phasor diagram in Fig. D-2. In this diagram, the quantities used for the calculation of signal-to-noise ratio for an FM system are represented as phasors referenced to the phase angle $\omega_c t$. Specifically, the ratio generally calculated is that of mean signal power with noise absent to mean noise power in the presence of an unmodulated carrier. The former is represented in the diagram by the phasor $A_c \cos \omega_c t$ and the latter by $A_n \cos (\omega_c + \omega)t$, where the subscript c defines the carrier and the subscript n defines noise, with the noise being calculated at the frequency components $\omega_c + \omega$.

For the case of signal-to-noise ratios > 10 (the situation illustrated in the diagram), then $A_n << A_c$ and the angle θ in the diagram is given by

$$\theta \approx \frac{A_n}{A_c} \sin \omega t \qquad \text{(D-1)}$$

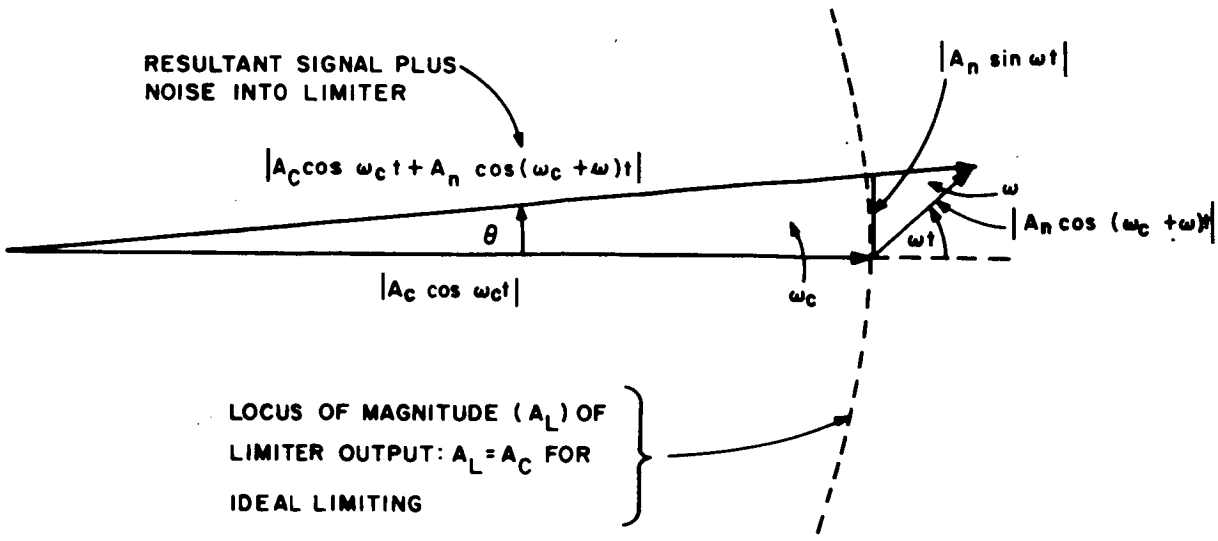


Fig. D-2. Carrier plus noise in an FM system for situation where noise amplitude is much less than carrier amplitude (i.e., $A_n << A_c$).

Thus, the output of the FM limiter in this case is

$$A_L \cos(\omega_c t + \theta) = A_L \cos\left(\omega_c t + \frac{A_n}{A_c} \sin \omega t\right) \qquad \text{(D-2)}$$

The locus of the magnitude of this equation is indicated in Fig. D-2. Note from this locus how the action of the limiter serves to "chop off" much of the effect of the noise disturbance, limiting this disturbance to a very small variation in phase θ about the reference phase $\omega_c t$. Clearly, this is a great improvement over the amplitude noise disturbances of AM.

Analytically, Eq. D-2 is exactly in the form of an FM signal which is narrowband, because A_n/A_c, the modulation index, is much less than 1. Thus, since a desired signal will be designed to produce wideband FM, the noise effect is reduced. In addition, as the signal deviation (equivalently, the FM bandwidth) is

increased, the effect of the noise will become smaller and smaller. Thus, we again have an improvement in performance over the AM system (a narrowband system) as shown graphically above. (Note that although our analysis here is for one frequency component ω only, superposition applies because narrowband FM systems are linear.)

However, when A_c/A_n drops below 10, the phasor diagram above indicates that the phase variation θ becomes larger and, therefore, the effect of the noise becomes greater. Analytically speaking, Eq. D-2 tells us that with A_n/A_c no longer much less than 1, the noise equation becomes one of wideband FM; thus, more noise comes through the limiter to the demodulator and the performance decreases over that of a narrowband system such as AM. In the literature,[1] it is shown analytically and experimentally that this threshold improvement really disappears rapidly below a signal-to-noise ratio of 9 dB.

---

[1] D. Middleton, "On the Theoretical Signal-to-Noise Ratios in FM Receivers: A Comparison with Amplitude Modulation," J. Appl. Phys., vol. 20, pp. 334-351, April 1949.

Appendix E

ATMOSPHERIC TRANSMISSION

The problem of describing what happens when information is carried via light through Earth's atmosphere is not simple, for the atmosphere is an everchanging heterogeneous medium, and it is difficult to represent even its major properties by a mathematical expression. This circumstance dictates an empirical solution based on a large amount of observed statistics.

The heterogeneous nature of the atmosphere is caused by the variety of materials it contains, e.g., gas molecules (water vapor, carbon dioxide, oxygen, ozone, etc.), sea salt, combustion nuclei, soot, dust, and so on. Because of wind and atmospheric currents, some of these particles are found at any given location and they act in two ways to attenuate radiation incident upon them: (1) by absorption and (2) by scattering. In either case, if a beam of light of initial intensity I_o traverses a distance x through an absorbing or scattering medium, its intensity is reduced according to a law of the form

$$I(x) = I_o e^{-\sigma x} \qquad \text{(E-1)}$$

where σ_o is the attenuation coefficient (absorption and/or scattering, depending on the mechanism involved).

For Earth's atmosphere, σ is the sum of four terms: (1) the attenuation coefficient describing air molecule scattering, called the Rayleigh (or molecular) scattering coefficient, (2) the attenuation coefficient describing scattering caused by spherical hygroscopic particles (haze and fog), called Mie (or large particle) scattering, (3) the attenuation coefficient describing scattering caused by large non-hygroscopic particles such as soot, and (4) the attenuation coefficient describing

band absorption by gas molecules. Not all these coefficients are significant for the application intended. For example, absorption can take place only where the constituents of the transmission medium have absorption bands. Compared to the scattering coefficients, absorption by atmospheric gas molecules is insignificant in the visible and near infrared spectrum and may be neglected.[1] Thus, the major source of atmospheric attenuation for the application at hand (i.e., communication at optical frequencies) is scattering. Values of the scattering coefficient at sea level (total including the sum of the three types described above) are given in Table E-1 for various meteorological conditions.[2]

Table E-1. SCATTERING COEFFICIENTS AT SEA LEVEL FOR VARIOUS METEOROLOGICAL CONDITIONS

Scattering Coefficient, σ_o (km^{-1})	Meteorological Description
80	Dense fog
20	Thick fog
8	Moderate fog
4	Light fog
2	Thin fog
1	Haze
0.4	Light haze
0.2	Clear
0.08	Very clear
0.014	Exceptionally clear (pure air)

1. Handbook of Geophysics, United States Air Force, pp. 16-20, The Macmillan Co., New York, 1960.

2. J.A. Curcio and K.A. Durbin, "Atmospheric Transmission in the Visible Region," ASTIA Document 227 798, pp. 52-53, October 6, 1959.

Now from Eq. E-1, it follows that the transmission of a light beam through the atmosphere at sea level is given by

$$T_a = \frac{I(R)}{I_o} = e^{-\sigma_o R} \qquad \text{(E-2)}$$

where T_a = atmospheric transmission

σ_o = atmospheric scattering coefficient at sea level (km^{-1})

R = transmission range (km)

I_o = intensity of light beam at its source (watts)

$I(R)$ = intensity of light beam at range R (watts)

Using Table E-1 and Eq. E-2, the effect of atmospheric attenuation on light transmission at sea level was calculated and plotted in Fig. E-1. From this plot it is observed, for example, that on a very clear day, a transmitted light beam is

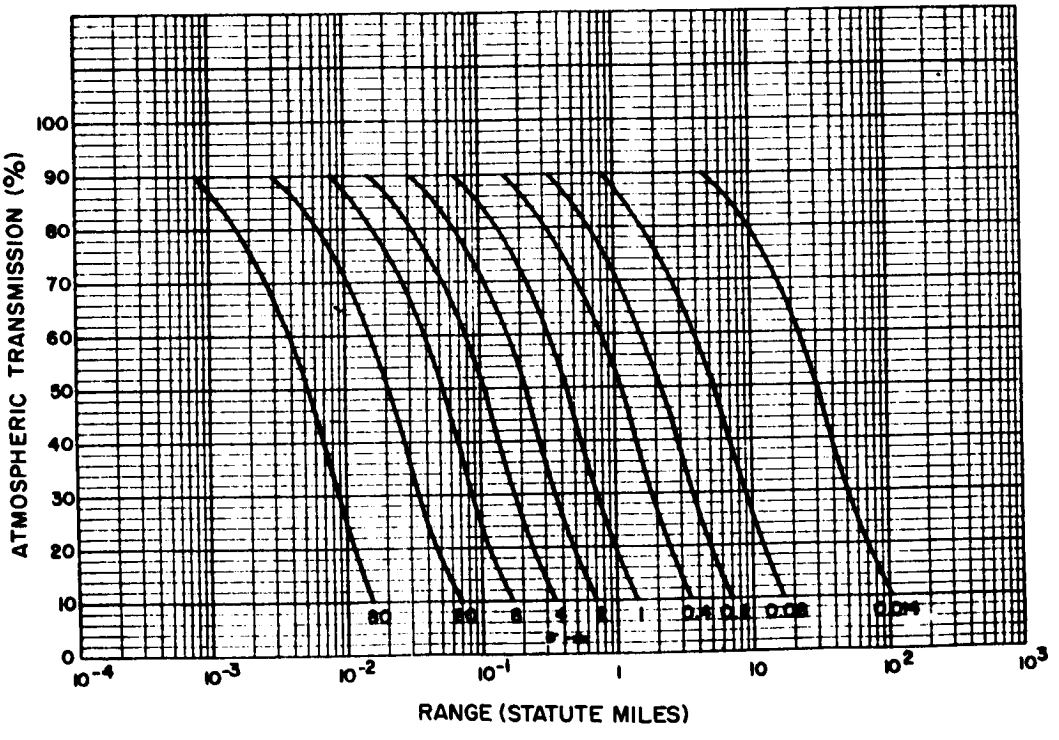


Fig. E-1. Atmospheric transmission vs range as a function of scattering coefficient at sea level.

attenuated about 75 per cent over a distance of 10 miles at sea level. Using Table E-1 and Fig. E-1, a plot of atmospheric transmission for three practically useful meteorological conditions (haze, clear, very clear) is presented in Fig. E-2. This plot illustrates clearly, for example, that on a clear day, $T_a \approx 0.2$, 0.1, and 0.04 at sea level for ranges of 5, 7, and 10 miles, respectively.

For a space-oriented system, the sea level atmospheric transmission information detailed above must be modified to describe transmission through the atmosphere in a vertical or oblique direction with respect to Earth's surface. Since σ is a function of altitude (i.e., height above Earth's surface), then from Eq. E-1, if a beam of light is transmitted perpendicularly to the surface of Earth, the intensity at height h is

$$I(h) = I_o e^{-y} \qquad \text{(E-3)}$$

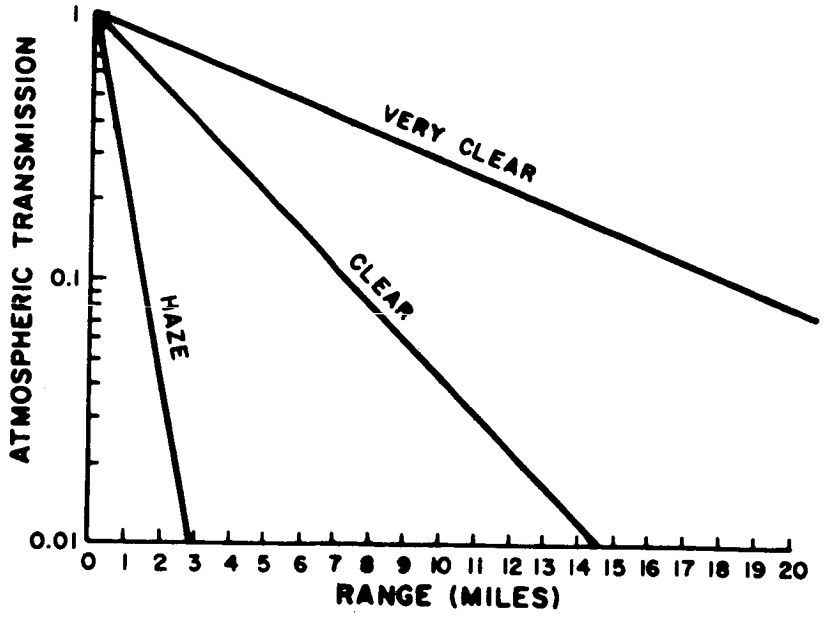


Fig. E-2. Atmospheric attenuation vs range at sea level as a function of meteorological condition.

where[1]

$$y = \int_0^h \sigma(h)\, dh \qquad \text{(E-4)}$$

It follows that the atmospheric transmission through the vertical atmosphere is

$$T_{av} = \frac{I(h)}{I_o} = e^{-y} \qquad \text{(E-5)}$$

To a good approximation,[2] the relation between the scattering coefficient and altitude is

$$\sigma(h) = \sigma_o e^{-h/H} \qquad \text{(E-6)}$$

where $\sigma(h)$ = scattering coefficient at altitude h (km^{-1})

σ_o = scattering coefficient at sea level (km^{-1})

h = altitude above Earth's surface (km)

H = altitude scale factor known as "scale height"

($H \approx 8$ km for Rayleigh scattering and 1.2 km for Mie scattering)

Substituting Eq. E-6 into Eq. E-4 gives

$$y = \sigma_o H (1 - e^{-h/H}) \qquad \text{(E-7)}$$

Substituting this expression for optical thickness into Eq. E-3 gives

$$T_{av} = \exp\left[\sigma_o H (e^{-h/H} - 1)\right] \qquad \text{(E-8)}$$

---

1. Since the factor y is related to the height h of atmosphere under consideration, it is commonly referred to as the "optical thickness."

2. Handbook of Geophysics, United States Air Force, pp. 16-20, The Macmillan Co., New York, 1960.

For oblique paths, the transmittance is

$$T_a = \exp\left[M\sigma_o H\,(e^{-h/H}-1)\right] \qquad \text{(E-9)}$$

where M is a factor known as the "optical air mass". For elevation angles greater than 30 degrees, meteorological tables[1] show that the optical air mass is

$$M \approx \frac{1}{\sin\gamma} \qquad \text{(E-10)}$$

where γ is the elevation angle of the light source.

As noted above, in the visible and near infrared regions of the electromagnetic wave spectrum, atmospheric attenuation is caused mainly by scattering. It is also true[2] that most of this scattering attenuation is a result of Mie scattering (i.e., scattering by water vapor, haze, and smog). Thus, with the value of H equal to 1.2 km and using $\sigma_o = 0.2\ \text{km}^{-1}$ (clear-day meteorological conditions), Eq. E-9 becomes

$$T_a = \exp\left[0.24M\,(e^{-h/1.2}-1)\right] \qquad \text{(E-11)}$$

This equation is plotted in Fig. E-3 indicating the attenuation to be expected as a function of altitude h and elevation angle γ.

Note from Fig. E-3 that for $\gamma = 90$ degrees (transmitting straight up through the atmosphere) attenuation of the light beam is only about 20 per cent. Of course, this represents only an average value over the frequency range of interest (visible

---

1. R. Penndorf, "Tables of the Refractive Index for Standard Air and the Rayleigh Scattering Coefficient for the Spectral Region between 0.2 and 20.0 Microns and Their Application to Atmospheric Optics," Jour. of the Optical Society of America, vol. 47, p. 176, February 1957.

2. W. J. Hannan et al., "Feasibility of Techniques for a Doppler Optical Navigator," Technical Documentary Report No. AL TDR 64-209, Contract No. AF 33(657)-11458, pp. 8-10, September 1964.

and near infrared). A more accurate plot for the cases γ = 30 and 90 degrees is given in Fig. E-4 where atmospheric transmission is given as a function of total scattering[1]. Note in this plot that for γ = 90 degrees, T_a varies between 62 and 93 per cent,depending on the wavelength of the laser source.

1. Ibid, p. 12.

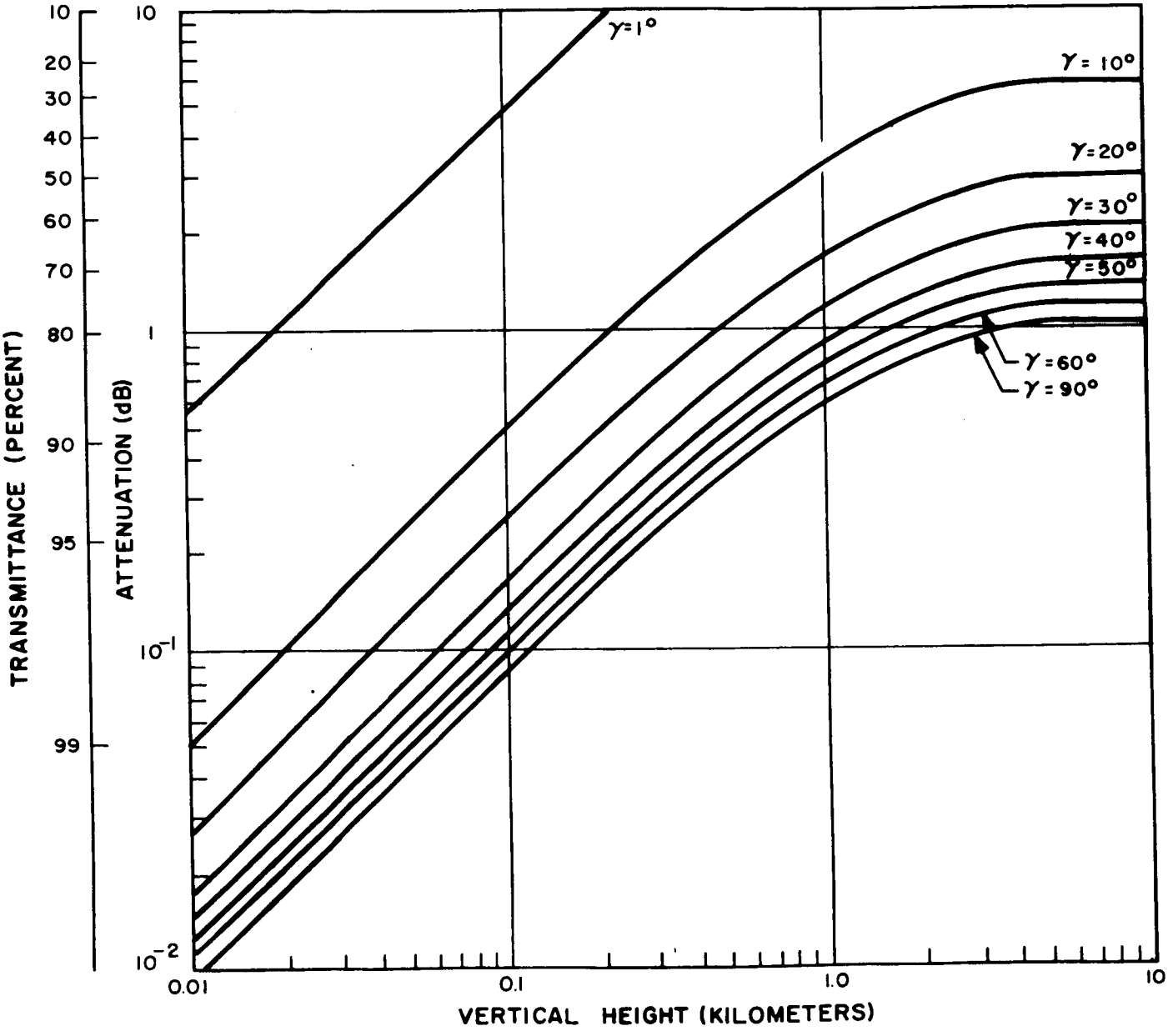


Fig. E-3. Atmospheric transmission as a function of altitude h and elevation angle γ.

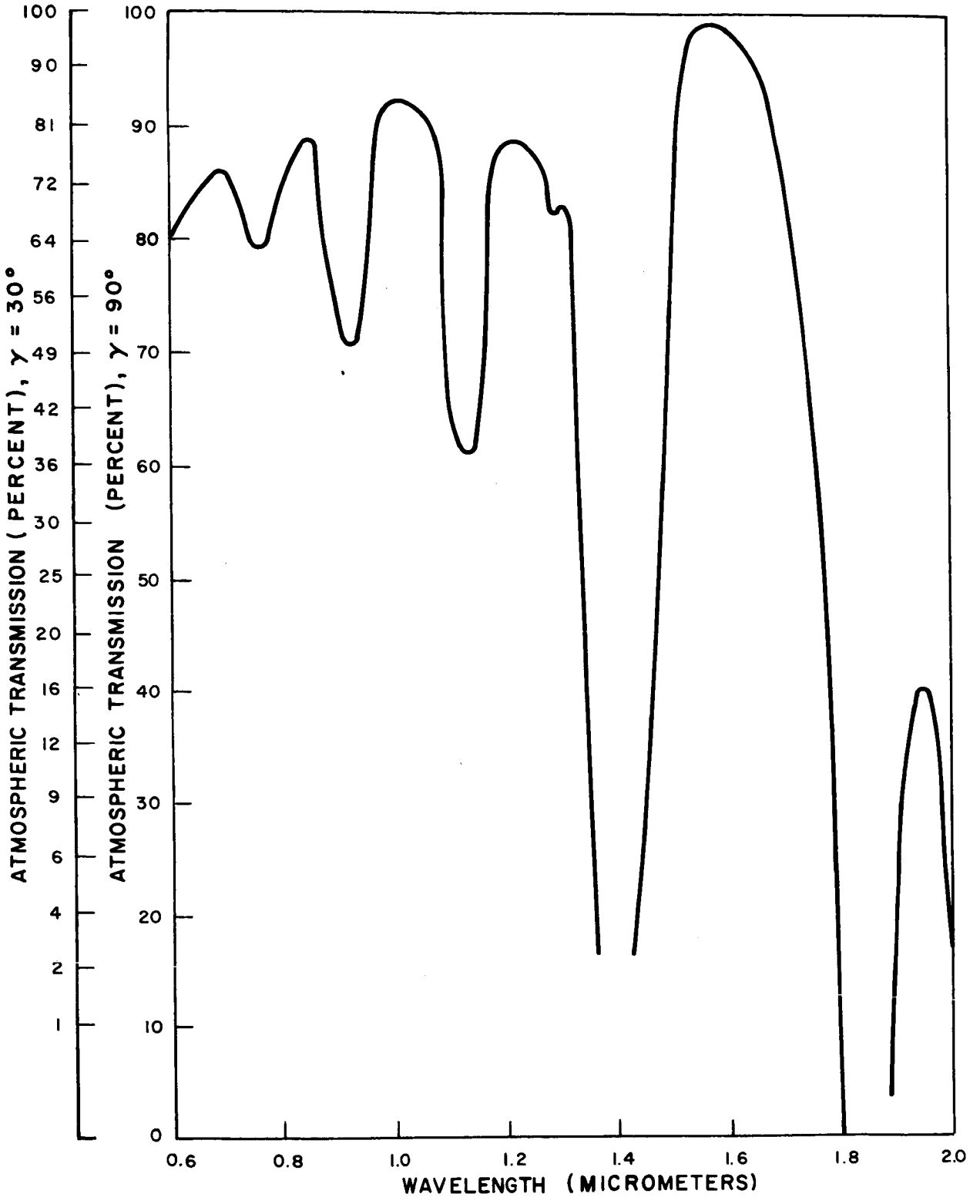


Fig. E-4. Atmospheric transmission as a function of total scattering.

Appendix F

SOLAR CELL ARRAY WEIGHT VS USEFUL OUTPUT POWER

The process of converting solar power to electrical power is illustrated in Fig. F-1. Incident solar power P_s is converted to low voltage direct current electric power P_{dc} by means of solar cells. This power, in turn, is converted to more useful, higher voltage direct current power P_o by means of an electronic DC-to-DC converter.

The weight of the solar cell array can be specified in terms of:

1) P_i, the solar power incident on a square meter of area A, equivalent to the area of the solar cell array normal to the axis between the Sun and the array.
2) η_1, the efficiency by which solar power is converted to electric power by the solar cell array.
3) η_2, the efficiency by which electric power from the solar cell array is converted to useful electric power ($\eta_2 = 1$ if this converter is not used, as would be the case if low voltage power is sufficient for the application intended).
4) P_O, the required power output at the end of the conversion process.
5) w_S, the weight per unit area of the solar cell array.

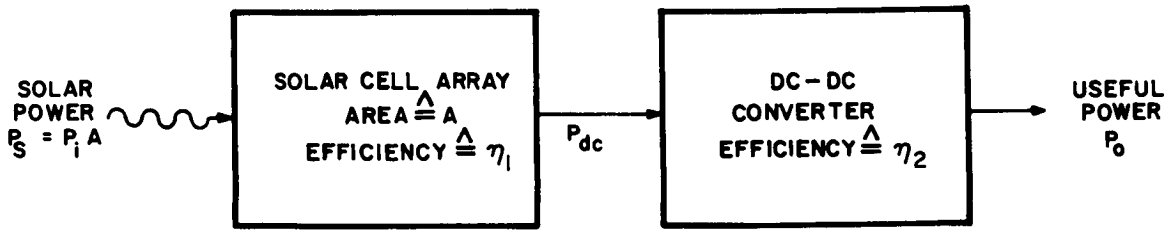


Fig. F-1. Solar to electrical power conversion.

Thus, the solar power P_s collected by the solar cell array is

$$P_s = P_i A \tag{F-1}$$

and the DC power P_{dc} available at the output of the array is

$$P_{dc} = \eta_1 P_s = \eta_1 P_i A \tag{F-2}$$

The output power P_o is, therefore,

$$P_o = \eta_2 \eta_1 P_i A \tag{F-3}$$

This equation can now be inverted to yield the area of the solar cell array necessary to furnish the desired power output P_o. Thus,

$$A = \frac{P_o}{\eta_1 \eta_2 P_i} \tag{F-4}$$

Hence, the weight w_a of a solar cell array for a given power output P_o is

$$w_a = w_s A = \frac{w_s P_o}{\eta_1 \eta_2 P_i} \tag{F-5}$$

where w_s is the weight per unit area of the solar cell array. When the DC-to-DC converter weight w_c is known, the total weight w_t of the power conversion system can be found from

$$w_t = w_c + \frac{w_s P_o}{\eta_1 \eta_2 P_i} \tag{F-6}$$

Equation F-5 can be used to obtain graphs relating the weight of a solar cell array to the power it can supply. Typical parameter values for this relationship[1,2] are

$\eta_1 = 0.14$

$\eta_2 = 0.70$

w_s = 1 to 3 lb/ft^2

P_i = 1322 W/m^2 at 93 x 10^6 mi range from Sun (average range of Earth from Sun)

P_i = 555 W/m^2 at 140 x 10^6 mi range from Sun (average range of Mars from Sun)

Substituting these parameters in Eq. F-5 yields

$$w_a = \frac{w_s P_o}{(0.14)(0.70)(1322)} = 0.0078\, w_s P_o \qquad \text{(F-7)}$$

and

$$w_a = \frac{w_s P_o}{(0.14)(0.70)(555)} = 0.0184\, w_s P_o \qquad \text{(F-8)}$$

Equation F-7 relates solar array weight to required electric power output for near-Earth operation (e.g., Earth satellite operation), and Eq. F-8 gives similar information for operation near Mars (e.g., manned space probe). These equations are plotted in Figs. F-2 and F-3, respectively.

---

1. P. A. Iles, "The Present State of Silicon Solar Cells," p. 13, IRE Trans. Military Electronics, vol. MIL-6, no. 1, January 1962.

2. N. W. Snyder, "Solar-Cell Power Systems for Space Vehicles," p. 89, IRE Trans. Military Electronics, vol. MIL-6, no. 1, January 1962.

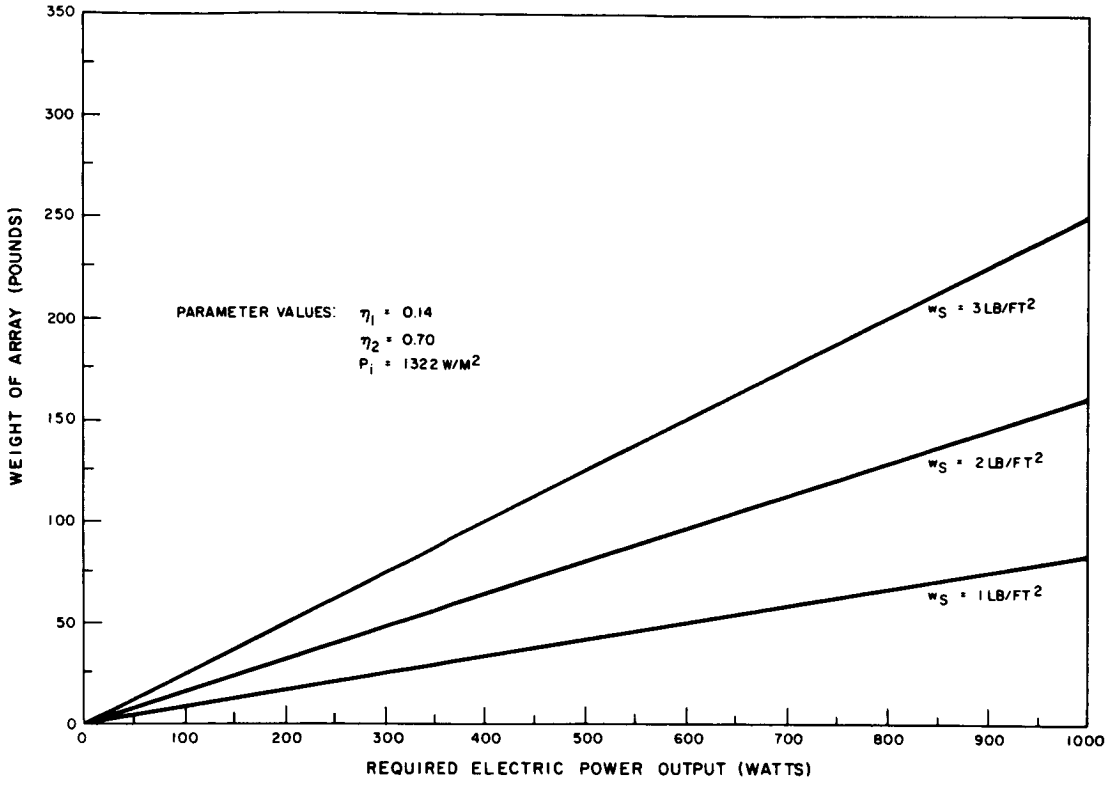


Fig. F-2. Solar cell array weight vs required electric power output for operation near Earth.

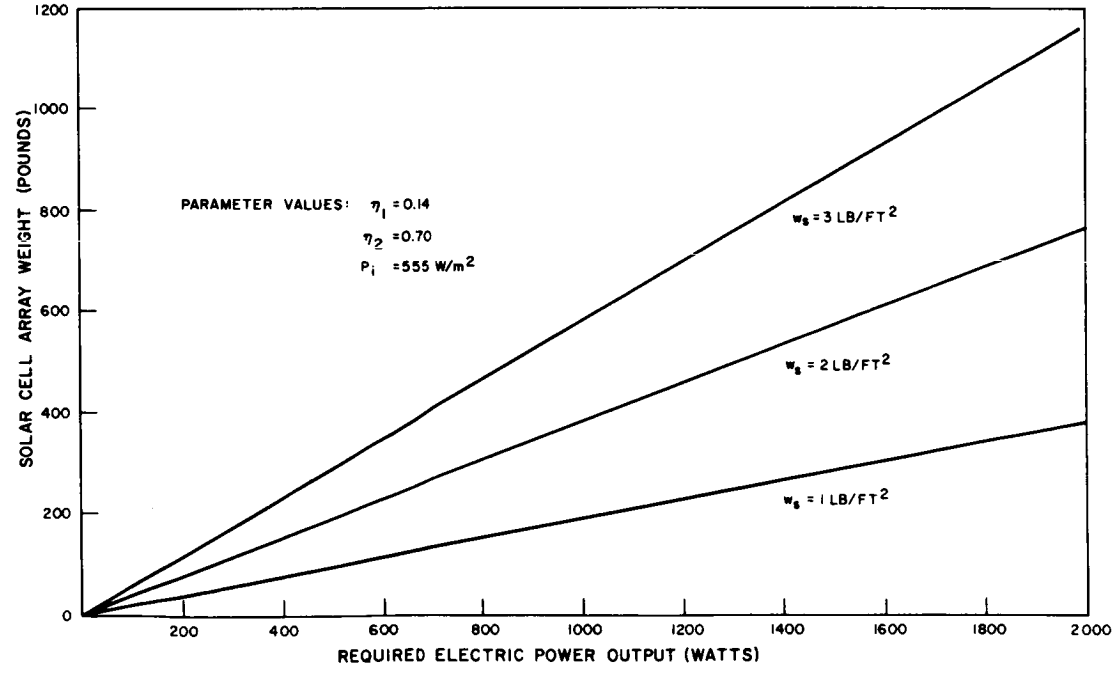


Fig. F-3. Solar cell array weight vs required electric power output for operation near Mars.

Appendix G

OPTIMIZATION OF CRYSTAL SIZE FOR A CYLINDRICAL LASER IN A SOLAR-PUMPED LASER SYSTEM

It is possible to define an optimum laser system as one which furnishes the required power output while operating at maximum efficiency. It is the purpose of this analysis to demonstrate that there exists an optimum size for the laser crystal used and to discuss the parameters involved in the choice of the crystal for a specified system.

The problem of optimizing a system with respect to laser crystal size has two aspects:

1) Coupling of solar energy to the laser crystal.
2) Design of an optical system which functions at maximum efficiency.

The first of these considerations can be examined qualitatively rather easily. It does not lend itself to any simple quantitative analysis. Consider the elements of an end-pumped laser system shown in Fig. G-1.

The solar image must be focused on the end of the laser crystal. There are three possible cases for a crystal with perfectly reflecting side walls (the only type to be considered):

1) $i > 2r_0$
2) $i < 2r_0$
3) $i = 2r_0$

where

i = size of solar image

r_0 = laser crystal radius.

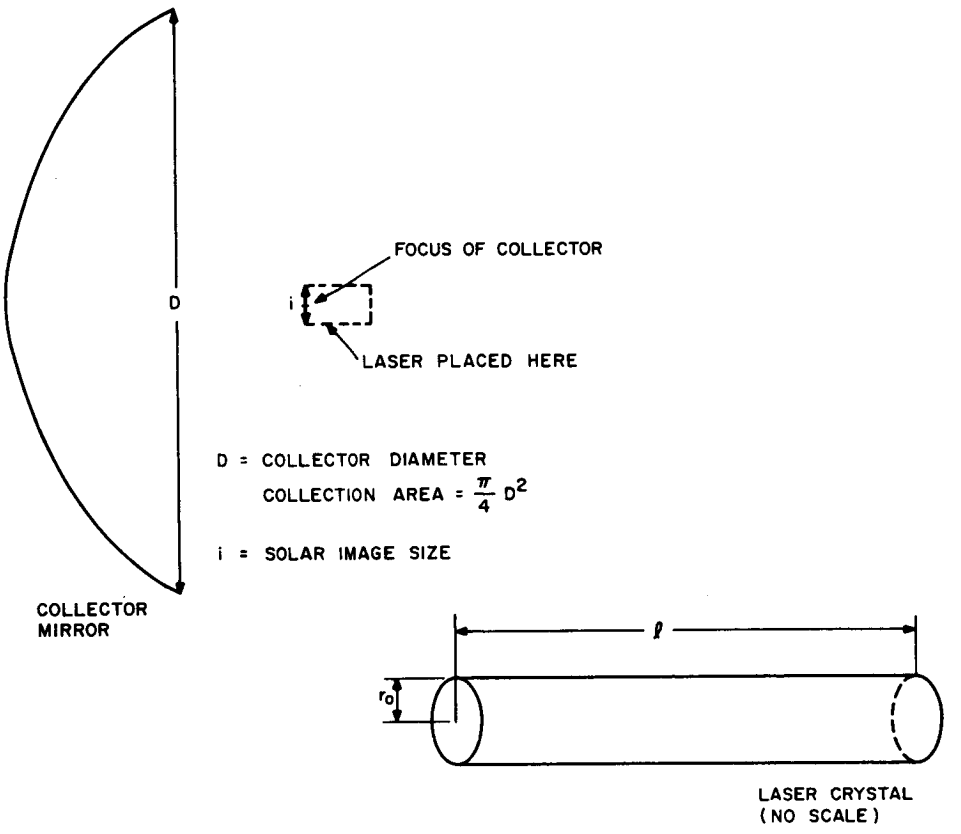


Fig. G–1. End-pumped laser system.

For case 1 it is evident that all available solar energy cannot be coupled to the laser crystal. Since loss of part of the solar image corresponds to a considerable loss of available pump energy, it follows that coupling of this type is inefficient and cannot be considered an optimum arrangement.

For case 2 the entire image falls on the crystal, but the pumping here is highly nonuniform. There are areas of the crystal face strongly pumped, while other appreciable areas are not pumped at all. Regions which are not pumped cannot contribute to the power output of the laser and may add losses. Also, it is possible that thermal problems might arise if an extremely high energy density is present at only some portion of the area of the crystal end. This may result in

the presence of a high-temperature region near the center of the crystal end which might prohibit laser action in that region or actual crystal fracture due to non-uniform crystal thermal expansion. It may also result in damage to the crystal end reflectors. Since none of the above effects can be tolerated, it must be concluded that case 2 is not an optimum coupling arrangement.

For case 3 the entire end of the crystal is pumped. Thus, inefficiencies caused by loss of part of the solar image, as in case 1, cannot be present, and the crystal can accept the maximum amount of radiation of which it is naturally capable. (The limit is again imposed by thermal considerations.)

It may be concluded that optimum pumping efficiency is achieved when the crystal diameter equals the size of the solar image.

Having defined an optimum energy coupling technique, it is now possible to design the energy collector which utilizes this technique to furnish a desired amount of power, P_{out}.

It can be shown that the output power of the cylindrical laser[1] is given by[2]:

$$P_{out} = P_{fc} (\xi - 1) \qquad \text{(G-1)}$$

where

$$P_{fc} = \frac{4\pi^3 h \nu^3 r_0^2 \alpha \Delta \nu}{c^2} = \text{Threshold power}$$

$$\xi = \frac{\text{Pump Power}}{P_{fc}}$$

α = loss per pass

1. A four-level, uniformly pumped laser system is being considered.
2. A. Yariv, J. P. Gordon, "The Laser," Proc. IEEE, vol. 51, no. 1, Jan. 1963.

$\Delta\nu$ = linewidth

ν = laser frequency

c = velocity of light in laser crystal.

Eq. G-1 can be written:

$$P_{out} = P - \frac{4\pi^3 h \nu^3 \alpha \Delta\nu}{c^2} r_0^2 \qquad \text{(G-2)}$$

where P = pumping power.

The power P enters the crystal through the end face being pumped. At that end there exists an areal power density $\mathcal{G}$ defined as:

$$\mathcal{G} = \frac{P}{\pi r_0^2} + \mathcal{G}' \qquad \text{(G-3)}$$

or

$$P = \pi r_0^2 (\mathcal{G} - \mathcal{G}') \qquad \text{(G-4)}$$

where

$\frac{P}{\pi r_0^2}$ = areal density of pump power

$\mathcal{G}'$ = areal density of power unsuitable for pumping (e.g., portions of solar spectrum not in laser pump bands).

It will be assumed that there exists a limiting areal power density, $\mathcal{G}_{max}$, on the crystal end beyond which thermal effects in the crystal degrade performance as previously described. It will be further assumed that the laser is operated under the condition $\mathcal{G} \approx \mathcal{G}_{max}$. This implies operation well above threshold at a point where maximum possible output from the laser is being achieved.

Under the condition of operation at $\mathcal{P}=\mathcal{P}_{max}$, Eq. G-4 becomes

$$P_{max} = \pi r_0^2 (\mathcal{P}_{max} - \mathcal{P}'_{max}). \qquad \text{(G-5)}$$

This equation gives the pump power available in a crystal of radius r_0 when the maximum areal power density at the crystal end is specified.

Use of Eq. G-2 now allows specification of the radius r_0 when a desired power output is required. Substitution of Eq. G-5 into Eq. G-2 gives:

$$P_{out} = \pi r_0^2 (\mathcal{P}_{max} - \mathcal{P}'_{max}) - \frac{4\pi^3 h\nu^3 \alpha \Delta\nu}{c^2} r_0^2 \qquad \text{(G-6)}$$

Solving for r_0,

$$r_0 = \left| \left\{ \frac{P_{out}}{\pi \left[(\mathcal{P}_{max} - \mathcal{P}'_{max}) - \frac{4\pi^2 h\nu^3 \alpha \Delta\nu}{c^2} \right]} \right\}^{1/2} \right| \qquad \text{(G-7)}$$

This is the optimum value of r_0 since the laser furnishes the required power P_{out} and operates under the highest pump level possible.

To this point no attention was given to the choice of a length dimension. Eq. G-2 above is suitable only for obtaining r_0. The choice of a length is made according to the following criteria:

1) The power absorbed, and hence output power, depends on the length. A length should be chosen such that an appreciable amount of available pump power is absorbed.

2) Scattering losses increase with length. Thus the crystal must not be so long that losses offset the effects of greater absorption with increasing length.

Experiments have been performed with crystals ($CaF_2:Dy^{2+}$, $YAG:Nd^{3+}$, $YAG:Nd^{3+}$ $-Cr^{3+}$) for which length $\approx$ 20 - 30 r_0. These crystals had low thresholds and operated efficiently at the pump levels available.

After the optimum dimensions have been determined for the crystal, knowledge of the magnitude of the solar constant completely defines the energy collection system. The solar constant is given by

$$\epsilon = \frac{P + \pi r_0^2 \mathcal{G}'_{max}}{A} = \frac{\text{Power Collected}}{\text{Collector Area}} \qquad \text{(G-8)}$$

where A = collector area. The condition[1] that $2r_0$ equals the size of the solar image implies:

$$r_0 = \frac{f\delta}{2} \qquad \text{(G-9a)}$$

or

$$f = \frac{2r_0}{\delta} \qquad \text{(G-9b)}$$

where

f = focal length of collector

δ = angle of solar subtense at laser system location.

The collector area is found from Eq. G-8:

$$A = \frac{P + \pi r_0^2 \mathcal{G}'_{max}}{\epsilon} \qquad \text{(G-10a)}$$

or

$$A = \frac{\pi r_0^2 \mathcal{G}_{max}}{\epsilon} \qquad \text{(G-10b)}$$

[1]This condition should be imposed at the maximum operating distance corresponding to $\delta = \delta_{min}$. A constant power output will then be obtained for all values $\delta > \delta_{min}$. This follows since $P = [(A\epsilon) / (\pi/4)(f\delta)^2]$ is constant. ϵ and δ^2 vary inversely as the square of the distance from the Sun.

The collector diameter is:

$$D = \sqrt{\frac{4A}{\pi}} = 2r_0 \sqrt{\frac{\mathcal{G}_{max}}{\epsilon}} \quad . \tag{G-11}$$

The "F-number" of the optical system is then found as:

$$F = \frac{f}{D} = \frac{2r_0}{\delta} \Big/ 2r_0 \sqrt{\frac{\mathcal{G}_{max}}{\epsilon}}$$

$$F = \frac{1}{\delta} \sqrt{\frac{\epsilon}{\mathcal{G}_{max}}} . \tag{G-12}$$

Consideration should now be given to the determination of the as yet unspecified parameter, $\mathcal{G}_{max}$. It was assumed above that there exists a maximum temperature above which local heating effects seriously diminish efficient laser performance. It was assumed that specification of a maximum temperature, T_{max}, would place an upper bound on $\mathcal{G}$ since the crystal temperature, T(r, Z), is a function of $\mathcal{G}$. It remains to be shown that this is so and that $\mathcal{G}_{max}$ can be determined from knowledge of T_{max}.

Consider the case of a laser of end area A pumped uniformly at an end with power density $\mathcal{G}$. The power absorbed by this crystal is given by:

$$\left| P_{abs} \right| = \left| A\mathcal{G}\beta \int_0^{\ell} e^{-\beta Z} \, dZ \right|$$

or

$$\left| P_{abs} \right| = \left| A\mathcal{G} (e^{-\beta \ell} - 1) \right| \tag{G-13}$$

where

β = absorption coefficient

ℓ = crystal length.

Under the assumption that most of the energy absorbed is transformed into thermal energy, and not useful output, it follows that the heat production rate A_o per unit time per unit volume (constant due to small absorption in length involved) produced by the incident radiation is given by:

$$A_o = \frac{|P_{abs}|}{V} = \left| \frac{\mathcal{P}}{\ell} \left(e^{-\beta \ell} - 1\right) \right|. \tag{G-14}$$

It is possible to determine the temperature distribution in the crystal in terms of A_o and conversely. Fig. G-2 illustrates a case where a solution can be obtained.

The steady-state diffusion equation for the case where there exists a source function A_o is:

$$\frac{\partial^2 T}{\partial r^2} + \frac{1}{r} \frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial Z^2} = -\frac{A_o}{K} \tag{G-15}$$

where K is the thermal conductivity of the material. With the boundary conditions

$$T = 0 \quad \begin{cases} r > r_0 \\ Z < 0, \; Z > \ell \end{cases} \tag{G-16}$$

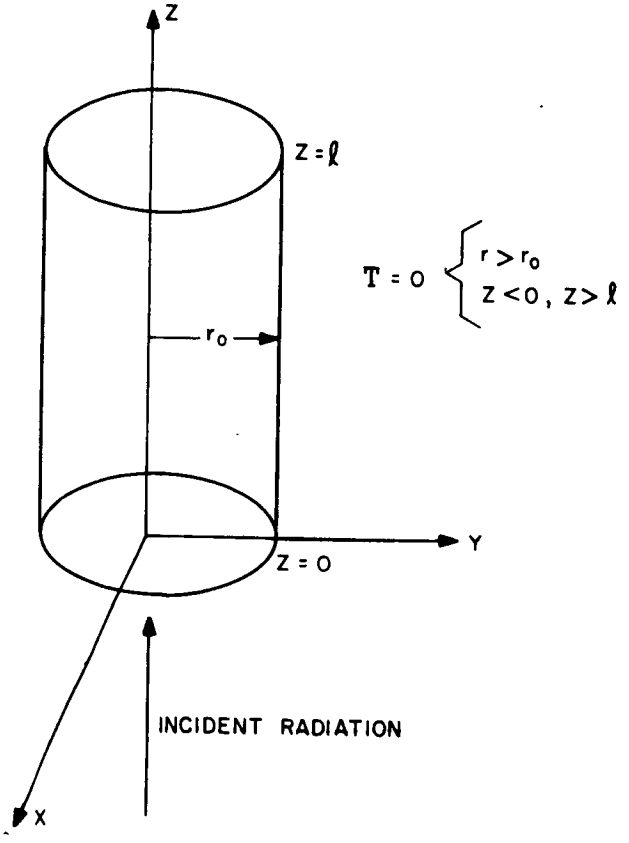


Fig. G-2. Conditions on uniformly end-pumped laser.

this equation has the solution:

$$T(r, Z) = A_0 \left\{ \frac{Z(\ell - Z)}{2K} - \frac{4\ell^2 A_0}{K\pi^3} \sum_{n=0}^{\infty} \frac{I_0 \left[(2n+1) \frac{\pi r}{\ell} \right]}{(2n+1)^3 I_0 \left[(2n+1) \frac{\pi r_0}{\ell} \right]} \sin (2n+1) \frac{\pi Z}{\ell} \right\} \quad \text{(G-17)}$$

or

$$T(r, Z) = A_o \, \Phi \, (r, Z) \quad \text{(G-18)}$$

In Eq. G-17, I_0 is the modified Bessel function of the first kind of order zero.

For the case considered, Eq. G-17 predicts a maximum temperature at the crystal center, as would be expected. The qualitative graphs of T vs r and Z shown in Fig. G-3 illustrate this.

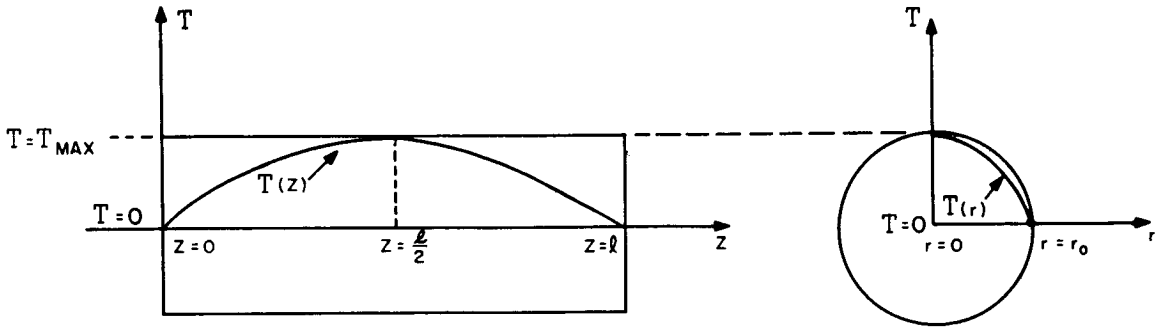


Fig. G-3. T vs r and Z for cylindrical laser uniformly end-pumped.

Now Eq. G-18 shows that if the maximum allowable value of T is specified, the maximum allowable value of A_o can be found from

$$A_o = \frac{T\left(0, \frac{\ell}{2}\right)}{\Phi\left(0, \frac{\ell}{2}\right)} \qquad \text{(G-19)}$$

since T_{max} must occur at $r = 0$, $Z = \ell/2$. Now since $A_{o_{max}}$ is known, Eq. G-14 furnishes the corresponding maximum areal power density which can pass through the crystal end:

$$\mathcal{Q}_{max} = \frac{\ell A_{o_{max}}}{\left|(e^{-\beta\ell} - 1)\right|} \qquad \text{(G-20)}$$

The above solution is only an approximation to the problem of determining $\mathcal{P}_{max}$ for a practical laser. It must be modified if thermal conductivities are temperature dependent and if the crystal boundaries are not at T = 0, or if pumping or crystal characteristics rule out the approximation A_o = constant. It does, however, demonstrate a method for determining a value for $\mathcal{P}_{max}$.

Information regarding the value of T_{max} must come from

1) strength of the cavity reflectors
2) strength of the crystal with regard to thermal expansion
3) behavior of laser action with temperature.

The first and second of these can be considered only for particular crystals and will not be discussed further. The third may be discussed for the general case in a limited sense to discover the general properties of laser output power as a function of temperature.

Consider the laser system shown in Fig. G-4 below. The aim of this discussion is to obtain an approximation for the change in pump power required to maintain a constant power output as the temperature of the laser (assumed uniform at least locally) changes from a value T to T′. The case $T' > T$ is of particular interest since, to maintain a desired output level, the input power must be increased. For the solar-pumped system being considered, this places a limit on the maximum allowable value of the laser temperature locally or in the entire volume since the amount of input power is determined by the geometry and will be constant. This maximum temperature is T_{max} used above.

According to Eq. G-1,

$$P_{out} = P_{fc\,T}\,(\xi - 1) \qquad \text{(G-21)}$$

at a temperature T.

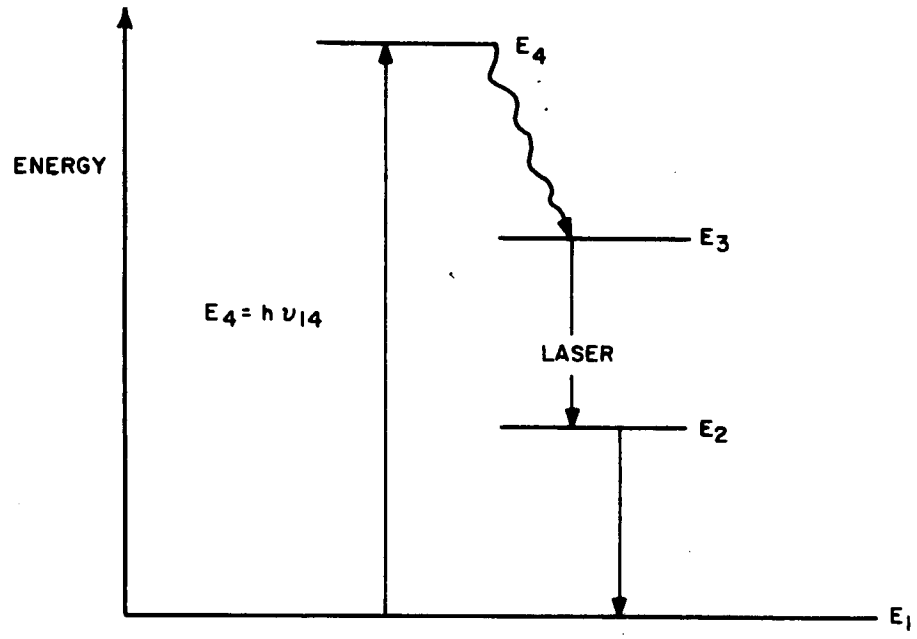


Fig. G-4. Four-level laser system.

Also, at temperature T':

$$P_{out} = P_{fcT'} (\xi' - 1). \tag{G-22}$$

If power output is to remain constant after a change in temperature,

$$\frac{\xi' - 1}{\xi - 1} = \frac{P_{fcT}}{P_{fcT'}} \tag{G-23}$$

must be satisfied. Since

$$P_{fcT} = \frac{4\pi^2 h \nu^3 V \alpha \Delta\nu}{c^2 \ell} \quad \text{and} \quad P_{fcT'} = \frac{4\pi^2 h \nu^3 V \alpha \Delta\nu'}{c^2 \ell}, \tag{G-24}$$

Eq. G-23 can be written

$$\frac{\xi' - 1}{\xi - 1} = \frac{\Delta\nu}{\Delta\nu'} . \tag{G-25}$$

This implies

$$\xi' = 1 + \frac{\Delta\nu}{\Delta\nu'} (\xi - 1). \tag{G-26}$$

Since

$$\xi = \frac{P_{inT}}{P_{fcT}} \text{ and } \xi' = \frac{P_{inT'}}{P_{fcT'}} , \tag{G-27}$$

this implies

$$P_{inT'} = P_{inT} + P_{fcT'} \left(1 - \frac{\Delta\nu}{\Delta\nu'}\right) . \tag{G-28}$$

The change in input power required to maintain a constant output power when the linewidth changes is

$$\Delta P_{in} = P_{inT'} - P_{inT} = P_{fcT'} \left(1 - \frac{\Delta\nu}{\Delta\nu'}\right) . \tag{G-29}$$

If temperature increases appreciably, power must also be supplied to excite more atoms into the upper state to maintain the necessary population difference between the laser levels, due to the fact that the lower-level population is not negligible. The population of this state is given by:

$$n_2 = n_0 e^{-E_2/kT} \tag{G-30}$$

where

n_0 = total number of active atoms in the crystal.

To excite this number of atoms to the upper laser level at temperature T', a power

$$P = \frac{h\nu_{14}}{t_s'} n_0 e^{-E_2/kT'} \tag{G-31}$$

is required (where t_s' = lifetime of upper laser level at temperature T'). This power P must be added to ΔP_{in} in Eq. G-29:

$$\Delta P_{in} = P_{fcT'} \left(1 - \frac{\Delta\nu}{\Delta\nu'}\right) + \frac{h\nu_{14}}{t_s'} n_0 e^{-E_2/kT'} . \tag{G-32}$$

From Eq. G-32,

$$\frac{\Delta P_{in}}{P_{fcT'}} = \left(1 - \frac{\Delta\nu}{\Delta\nu'}\right) + \frac{1}{P_{fcT'}} \frac{h\nu_{14}}{t_s'} n_0 e^{-E_2/kT'} . \tag{G-33}$$

But

$$P_{fcT'} = \frac{\Delta n_c' h\nu_{14}}{t_s'}$$

where $\Delta n_c'$ = population difference at temperature T for power output P_{out} and $P_{fcT'} = \Delta\nu'/\Delta\nu \; P_{fcT}$. When this information is used in Eq. G-33, it follows that

$$\frac{\Delta P_{in}}{P_{fcT}} = \frac{\Delta\nu'}{\Delta\nu} \left(1 + \frac{n_0}{\Delta n_c'} e^{-E_2/kT'}\right) - 1 . \tag{G-34}$$

Thus the pump power change required as $T \rightarrow T'$ to maintain constant output power is

$$\Delta P_{in} = K P_{fcT} \qquad \text{(G-35)}$$

where

$$K = \frac{\Delta\nu'}{\Delta\nu} \left(1 + \frac{n_0}{\Delta n'_c} e^{-E_2/kT'} \right) - 1 .$$

If $T' > T$, $\Delta\nu' > \Delta\nu$ and ΔP_{in} represents an increase in power input. The graph in Fig. G-5 shows K vs E_2/kT' with $\Delta\nu'/\Delta\nu$ as a parameter. Typical values for n_0 and $\Delta n'_c$ have been assumed:

$$n_0 \approx 10^{18}$$
$$\Delta n_c \approx 10^{16}$$

For example, consider YAG. At T = 300°K, YAG is an ideal four-level laser since $E_2/kT = 9.6$. If temperature is increased to $T' = 400°K$, $E_2/kT' = 5.8$. Thus, even if $\Delta\nu' = \Delta\nu$, K = 0.3, and thus input power must increase by $0.3\ P_{fc_T}$ to maintain a constant output for the crystal with $n_0 = 10^{18}$ atoms and $\Delta n_c = 10^{16}$ atoms. This indicates that, if this laser were solar pumped and if a temperature of 400°K could exist in the crystal, a maximum power output no greater than $P - 1.3\ P_{fc}$ could be obtained, where P is the pump power. At T = 300°K, the same laser could furnish a maximum output power = $P - P_{fc}$. If the power output required were $P - P_{fc}$, this particular laser could not furnish it if the crystal temperature were greater than 300°K. In this case, the system parameter T_{max} would be specified as 300°K.

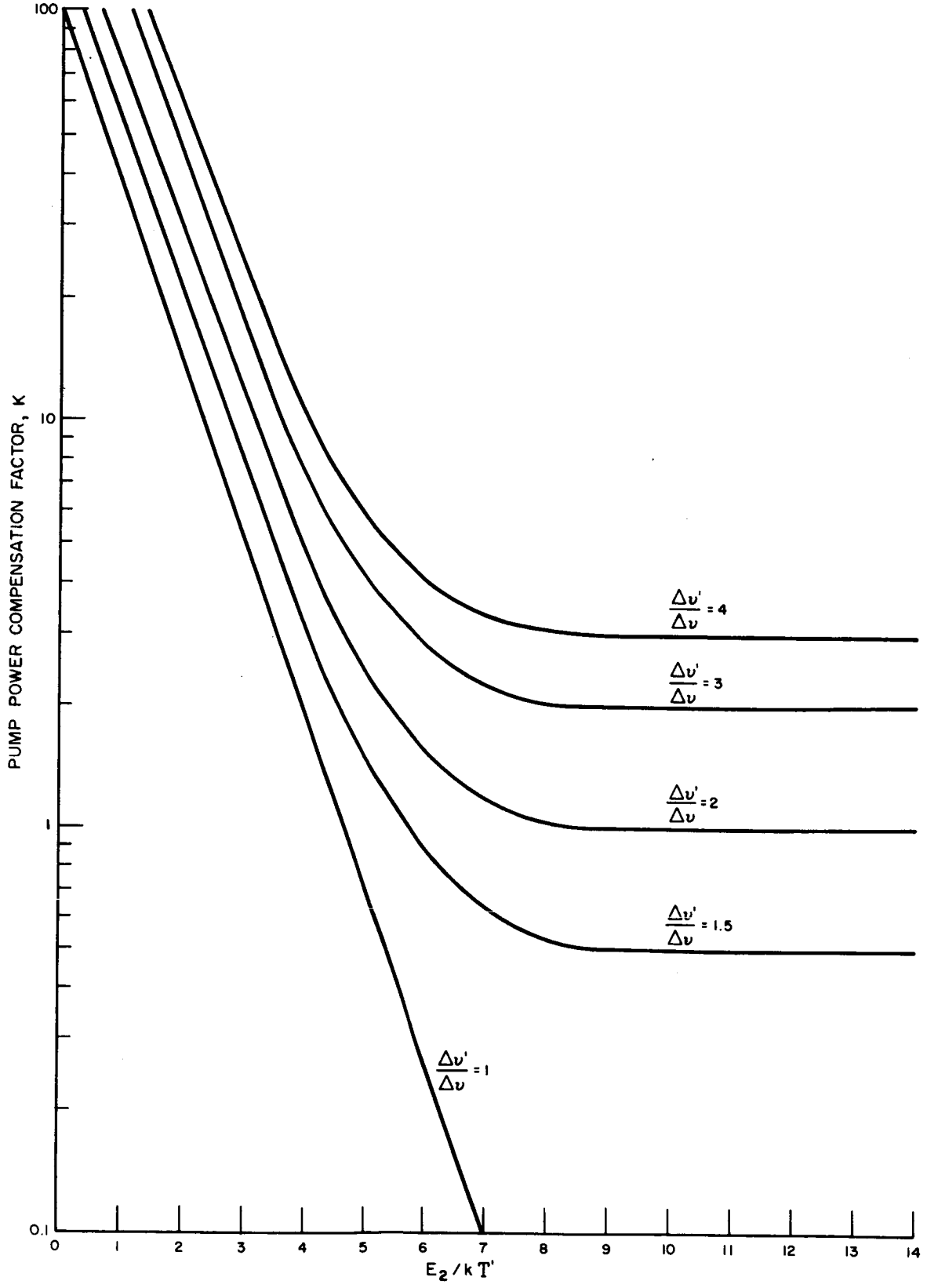


Fig. G-5. Pump power compensation factor K vs E_2/kT'.

Appendix H

DRIVE CURRENT REQUIRED FOR MAGNETIC MODULATION

Laser action in $CaF_2:Dy^{2+}$ occurs between two triply degenerate states $^5I_7 \rightarrow {}^5I_8$. In a magnetic field, the degeneracy of the states is removed, and each of these levels splits into three separate levels as shown in Fig. H-1. Also shown in Fig. H-1 are the allowed magnetic dipole transitions among the split levels, which correspond to the laser output, and the gyromagnetic ratios for the upper and lower states.

Use may be made of the normal Zeeman effect splittings to accomplish modulation of the laser output. Two approaches may be taken:

1) Frequency modulation of the center of the fluorescent line using the Zeeman splitting caused by application of a uniform field. (Amplitude modulation results because a narrow-band Fabry-Perot cavity is used.)
2) Broadening of the effective width of the fluorescent line by using an inhomogeneous field to shift the fluorescent line in varying amounts throughout the crystal.

Modulation can be accomplished if the power in the wave propagating through the laser crystal can be varied in a controllable fashion. Yariv and Gordon[1] define a gain coefficient as:

$$\gamma = \frac{dE_w}{d\ell} \Big/ E_w \qquad \text{(H-1)}$$

where E_w is the energy of the wave propagating in the laser, and ℓ is the path length over which the wave travels. Energy added to the wave must come from energy stored in the upper laser level, so that, if loss mechanisms are neglected, the change in the wave energy, dE_w, must equal the change in the energy of this upper state, dE_u:

1. A. Yariv and J.P. Gordon, "The Laser," Proc. IEEE, 51, 4, Jan. 1963.

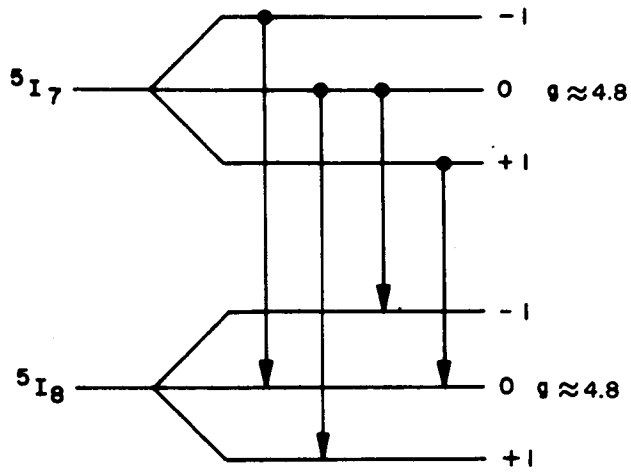


Fig. H-1. Zeeman splitting of levels $^5I_7$ and $^5I_8$ and CaF_2 and allowed laser transitions.

$$dE_W = dE_u \tag{H-2}$$

Inserting this in Eq. H-1 and dividing by dt, we get

$$\frac{dE_u}{dt} = \gamma E_W \frac{d\ell}{dt} . \tag{H-3}$$

Loss mechanisms such as scattering and fluorescence, which are linear, can be taken into account by writing the rate of energy loss from the wave as

$$P_\ell = \frac{dE_\ell}{dt} = \delta E_W \frac{d\ell}{dt} , \tag{H-4}$$

where δ is the loss coefficient. The rate of energy increase of the wave as a result of increases in energy from the upper laser level energy and decreases due to the loss mechanisms, is given by

$$\frac{dE_W}{dt} = \frac{dE_u}{dt} - \frac{dE_\ell}{dt} \tag{H-5}$$

The efficiency η of the total process may be defined as the ratio of the power addition to the wave to the power loss of the upper laser state:

$$\eta = \frac{dE_w}{dt} \Big/ \frac{dE_u}{dt} = P_w/P_u \tag{H-6}$$

Using Eqs. H-3, H-4, and H-5, the efficiency may be expressed as

$$\eta = 1 - \frac{\delta}{\gamma} . \tag{H-7}$$

The gain coefficient γ is given by

$$\gamma = \frac{\lambda^2 (n_3 - n_2)\, \Delta\nu}{16\pi^2 \tau \left[(\nu - \nu_c)^2 + \left(\frac{\Delta\nu}{2}\right)^2\right]} \tag{H-8}$$

where:

n_3 = population of upper laser state

n_2 = population of lower laser state

λ = laser wavelength

τ = lifetime of upper laser state

ν_c = center frequency of fluorescent line

$\Delta\nu$ = fluorescent linewidth

ν = frequency at which gain is calculated.

If we let

$$K = \frac{\lambda^2}{16\pi^2} \tag{H-9}$$

and

$$\Delta\nu_{eff} = \frac{\left[\left(\frac{\Delta\nu}{2}\right)^2 + (\nu - \nu_c)^2\right]}{\Delta\nu} , \tag{H-10}$$

then Eq. H-8 becomes

$$\gamma = \frac{K\tau\,(n_3 - n_2)}{\Delta\nu_{eff}} \tag{H-11}$$

Substituting Eq. H-11 into Eq. H-7 allows the efficiency to be expressed as

$$\eta = 1 - \frac{\delta\,\Delta\nu_{eff}}{K\,(n_3 - n_2)\,\tau} \tag{H-12}$$

Variation of this efficiency corresponds to the modulation process. Magnetic modulation by means of the normal Zeeman effect results in varying the quantity $\Delta\nu_{eff}$, which, as indicated by Eq. H-10, varies the frequency and varies the power output.

The Hamiltonian for a state exhibiting the normal Zeeman effect can be written as:

$$\mathcal{H} = \mathcal{H}_{on} - g\beta\overline{H} \cdot \overline{L} \tag{H-13}$$

where:

$\mathcal{H}_{on}$ = energy terms independent of magnetic field
g = gyromagnetic ratio
β = Bohr magneton (9.3×10^{-21} erg/gauss)
$\overline{H}$ = magnetic field intensity
$\overline{L}$ = angular momentum of state n
n = a subscript for numbering the state.

It follows that

$$\mathcal{H}_n = \mathcal{H}_{on} - g\beta H\,M_{LM}, \tag{H-14}$$

where

$$M_{LM} = \frac{\overline{H} \cdot \overline{L}}{H}, \qquad H = \left|\overline{H}\right| .$$

For the transitions of interest in CaF_2:Dy^{2+},

$$\Delta M_{32} = M_{L3} - M_{L2} = \pm 1. \qquad \text{(H-15)}$$

Subscripts 2 and 3 refer to the states $5I_8$ and $5I_7$, respectively. When a $CaF_2:Dy^{2+}$ laser is operated while subjected to an external magnetic field parallel to the (100) direction, transitions ($^5I_7 \rightarrow {}^5I_8$) for which $\Delta M_{32} = \pm 1$ occur simultaneously with equal probability. As a result, there are shifts of the center frequency of the fluorescent line above and below the laser frequency which occurs in the absence of the field.

The energy change as a result of a laser transition expressed in terms of the difference between the proper Hamiltonian operators is:

$$\mathcal{H}_3 - \mathcal{H}_2 = (\mathcal{H}_{03} - \mathcal{H}_{02}) - g\,\beta H\,\Delta M_{32}, \qquad \text{(H-16)}$$

if the values of g are the same for the initial and final states. The second term is expressible as a frequency difference in the form

$$\Delta\nu'_H = \frac{g\,\beta H\,\Delta M_{32}}{h} \qquad \text{(H-17)}$$

where h is Planck's constant. Since ΔM_{32} assumes the values ± 1, the total effective frequency difference between the two shifted (fluorescent) line components is:

$$\Delta\nu'_H = \frac{2g\beta H}{h} \qquad \text{(H-18)}$$

which is expressible as:

$$\Delta\nu_H = \frac{g_{eff}\,\beta H}{h} \qquad \text{(H-18A)}$$

where g_{eff} represents an effective "g" for the laser transitions. This "g_{eff}" will be denoted by "g" from this point and is equal to approximately 9.6.

For the case of modulation by a homogeneous field, the field has the effect of shifting the center frequency of the fluorescent line, by the amount given by Eq. H-18. Thus, assuming that the operating frequency ν of the laser (i.e., the

cavity resonance) is set at ν_c, the center of the line, it is seen that Eq. H-10 gives $\Delta\nu_{eff}$ (in the absence of a field) as:

$$\Delta\nu_{eff} = \frac{\Delta\nu}{4} , \tag{H-19}$$

while after the field is applied,

$$\Delta\nu_{eff} = \frac{\left(\frac{\Delta\nu}{2}\right)^2 + \left(\frac{g\beta H}{h}\right)^2}{\Delta\nu} \tag{H-20}$$

Thus $\Delta\nu_{eff}$ is seen to be a quadratic function of H, as is also the efficiency.

In the case where an inhomogeneous field is applied across the crystal, the fluorescent linewidth is effectively broadened by an amount

$$\frac{g\beta\Delta H}{h} , \tag{H-21}$$

where ΔH is the variation of H across the crystal. Assuming that the operating frequency ν is at ν_c, the quantity $\Delta\nu_{eff}$ is given by

$$\Delta\nu_{eff} = \frac{\Delta\nu}{4} + \frac{g\beta\Delta H}{4h} . \tag{H-22}$$

In this case, the quantity $\Delta\nu_{eff}$, and also the efficiency, is seen to be a linear function of ΔH.

The process of modulation can now be considered. Consider a laser operating while not under the influence of a magnetic field. The threshold for this laser is given by:[1]

$$P_{th} = \frac{4\pi^2 h \nu^3 V \Delta\nu_{eff}}{c^3 t_p} \tag{H-23}$$

1. A. Yariv and J.P. Gordon, "The Laser", Proc. IEEE, Jan. 1963, Vol. 51, No. 1, p. 9.

If the same laser is operated while in an inhomogeneous magnetic field, the expression for threshold becomes:

$$P_{th}' = \frac{4\pi^2 h\nu^3 V \Delta\nu'_{eff}}{c^3 t_p} \qquad \text{(H-24)}$$

Since $\Delta\nu'_{eff} > \Delta_{eff}$, it is evident that the effect of the magnetic field is to raise the threshold for laser operation. If $\Delta\nu'_{eff}$ is sufficiently large, the threshold can be made equal to or greater than the operating input power level. When this condition is reached, laser action ceases. This corresponds to the point at which 100 per cent modulation is achieved.

The ratio of the threshold with field to the threshold without field is given by:

$$\frac{P_{th}'}{P_{th}} = \frac{\Delta\nu'_{eff}}{\Delta\nu_{eff}} \qquad \text{(H-25)}$$

Now if this ratio is equal to the pump ratio (pump power/threshold pump power), the condition of 100 per cent modulation is obtained since at this point the available pump power is less than or equal to the threshold power for the laser operated in the magnetic field. Thus, for 100 per cent modulation:

$$\frac{\Delta\nu'_{eff}}{\Delta\nu_{eff}} = \xi \qquad \text{(H-26)}$$

The statement of this condition allows the required inhomogeneous field difference ΔH for 100 per cent modulation to be related to the operating pump power level.

From Eq. H-22 it is evident that

$$\Delta\nu_{eff} = \frac{\Delta\nu}{4} \qquad \text{(H-27)}$$

$$\Delta\nu'_{eff} = \Delta\nu_{eff} + \frac{g\beta\Delta H}{4h} . \qquad \text{(H-28)}$$

Thus

$$\xi = 1 + \frac{g\beta\Delta H}{4h\,\Delta\nu_{eff}} \quad . \tag{H-29}$$

We can now determine the drive current through a modulation coil required to produce 100 per cent modulation. The field produced at the center of a short solenoid is given by[1]

$$H = \frac{LI}{kN\pi r^2} \times 10^4 \text{ gauss}, \tag{H-30}$$

where

k = 0.53 (finite solenoid compensation factor)

L = coil inductance (henrys)

N = number of turns

r = radius of coil (meters).

Combining Eqs. H-30 and H-31, we find that the current required for 100 per cent modulation is

$$I = \frac{4\pi kNr^2 h\,\Delta\nu_{eff}\,(\xi - 1)}{Lg\beta} \times 10^{-4} \text{ amperes}. \tag{H-31}$$

1. W.B. Boast, "Principles of Electric and Magnetic Fields," Harper and Brothers, N.Y., 1948.

Appendix I

DRIVE POWER FOR MAGNETIC MODULATION

A. PULSE MODULATION

According to the results given in Sec. IV. D (see Fig. 26), the peak current through a 6-μH coil for 100 per cent modulation is one ampere. Thus, the question to be answered here is "How much power is required to send 1-A pulses through a 6-μH coil at the pulse repetition rate demanded by the quantity of information to be transmitted?" The answer to this question is provided by the following analysis.

Assuming both the coefficient of coupling between L and L' and the magnitude of C in Fig. I-1 to be negligible, and letting r represent the total resistance in series with the power supply E, the left-hand loop current I(p) is given by

$$I(p) = \frac{E/p}{r+pL} \tag{I-1}$$

or, in the time domain,

$$i(t) = \frac{E}{r}\left(1-e^{-\frac{r}{L}t}\right) \tag{I-2}$$

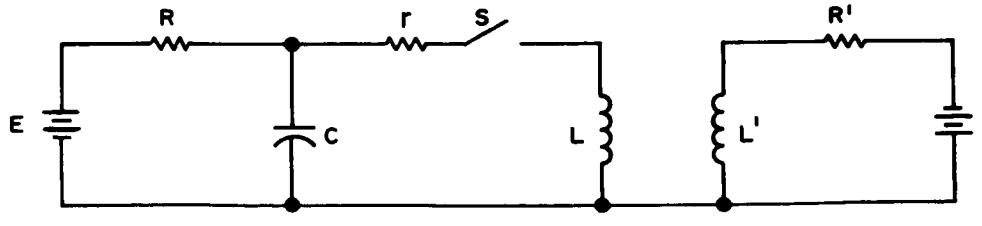


Fig. I-1. Modulation circuit.

after each closing of the switch S. In other words, i(t) is the response to a modulating voltage step.

Denoting T as the pulse period and τ as the pulse duration (i. e., duty factor = τ/T), the mean square loop current is

$$i_{rms}^2 = \frac{E^2}{r^2 T} \int_o^{\tau} \left(1-e^{-\frac{r}{L}t}\right)^2 dt \tag{I-3}$$

$$= \frac{E^2}{r^2 T} \left[\tau + \frac{2L}{r} e^{-\frac{r}{L}\tau} - \frac{L}{2r} e^{-\frac{2r}{L}\tau} - \frac{2L}{r} + \frac{L}{2r}\right] \tag{I-4}$$

If a long time constant is used (i. e., $\frac{L}{r} >> 2\tau$), Eq. I-4 reduces to

$$i_{rms}^2 = \frac{E^2 \tau}{r^2 T}$$

$$= (\hat{I})^2 \frac{\tau}{T} \tag{I-5}$$

where $\hat{I}$ is the peak drive current. On the other hand, if a short time constant is used (i. e., $\frac{L}{r} << \tau$), Eq. I-4 again reduces to Eq. I-5. In other words, the equation for input power required for magnetic modulation is independent of time-constant magnitude. Since input power is

$$P_i = i_{rms}^2 r$$

$$= (\hat{I})^2 r \frac{\tau}{T} \tag{I-6}$$

we conclude that for minimum input power, the resistance should be as small as possible.

Of course, the size of the resistance r (which includes the switch resistance) is dictated by the available switching devices. For the purpose intended, semiconductor switches are used. Reasonable state-of-the-art values for series resistance and pulse duration are one ohm and 50 nanoseconds, respectively.

Substituting $\hat{I} = 1$ A, $r = 1\ \Omega$ and $\tau = 5 \times 10^{-8}$ s into Eq. I-6, we obtain the relationship between input power and pulse repetition rate, plotted in Fig. 28 (see Sec. IV. D). Since the maximum duty factor is $\tau/T = 0.5$, the maximum input power is 0.5 watt and the maximum pulse repetition rate is 10 MHz, as seen from Fig. 28.

B. ANALOG MODULATION

In the preceding analysis, the drive power required for magnetic modulation was derived from the point of view of pulsing the magnetic field. This type binary modulation could be used for PCM (pulse code modulation), PFM (pulse frequency modulation) or any other type of binary modulation. As a comparison, consider now the power required for analog modulation with a video signal. In particular, suppose the video signal has a triangular frequency spectrum, viz.,

$$\begin{aligned} E(\omega) &= E\left(1 - \frac{\omega}{\omega_c}\right) && 0 \le \omega \le \omega_c \\ &= 0 && \omega \ge \omega_c \end{aligned} \qquad \text{(I-7)}$$

where E is a constant and ω_c is the modulation bandwidth. (This type spectral distribution is similar to that of a typical TV signal.) In this case, the circuit of Fig. I-1 becomes (with the magnitude of C and the coefficient of coupling between L and L' negligible and r representing the total resistance in series with the power supply, as before) that shown in Fig. I-2.

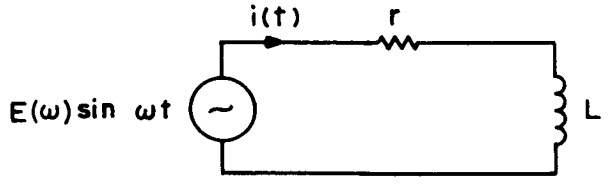


Fig. I-2. Approximate magnetic modulation circuit for analog modulation.

The modulation current i(t) is given by

$$i(t) = \frac{E(\omega)}{\sqrt{\omega^2+\omega_c^2}} \sin \omega t \tag{I-8}$$

where

$$\omega_c \triangleq \frac{r}{L} \tag{I-9}$$

Thus

$$i_{rms}^2 = \overline{i(t)^2} = \frac{E^2(\omega)}{2L^2\left(\omega^2+\omega_c^2\right)} \tag{I-10}$$

The substitution of Eq. (I-7) into this last equation yields

$$i_{rms}^2 = \frac{E^2\left(1-\frac{\omega}{\omega_c}\right)^2}{2L^2\left(\omega^2+\omega_c^2\right)} \tag{I-11}$$

and therefore, the average power input to the modulation circuit at one frequency ω is

$$\overline{P(\omega)} = i_{rms}^2 r = \frac{rE^2\left(1-\frac{\omega}{\omega_c}\right)^2}{2L^2\left(\omega^2+\omega_c^2\right)} \tag{I-12}$$

Thus, the total average power input over the bandwidth ω_c is

$$\overline{P}_{total} = \frac{1}{\omega_c}\int_0^{\omega_c} \overline{P}(\omega)\, d\omega = \frac{1}{\omega_c}\int_0^{\omega_c} \frac{rE^2\left(1-\frac{\omega}{\omega_c}\right)^2}{2L^2\left(\omega^2+\omega_c^2\right)}\, d\omega \tag{I-13}$$

The indicated integration yields

$$\overline{P}_{total} = 0.154\,\frac{rE^2}{L^2\omega_c^2} \tag{I-14}$$

In order to achieve 100 per cent modulation at low frequencies, the drive voltage must be

$$E = I(0)r \tag{I-15}$$

Where I(0) is the modulation current needed for full modulation. Thus,

$$\overline{P}_{total} = \frac{0.154\ rI^2(0)r^2}{L^2\omega_c^2}$$

$$= 0.154\ I^2(0)\omega_c L \tag{I-16}$$

A plot of this equation for I(0) = 1 A and L = 6 μH is given in Fig. 29 (see Sec. IV. D).

Appendix J

COMPARISON OF MODULATION METHODS

Assuming that the usual steps are taken to minimize the noise factor of the receiver, then overall communication system performance depends primarily on the type of modulation employed. The choice of modulation method depends on a compromise among required transmitter power, available bandwidth, minimum acceptable signal-to-noise ratio, and acceptable complexity of the modulation and detection circuits. In the following discussion, several modulation techniques suitable for transmitting TV signals are reviewed.

Modulation techniques which allow channel bandwidth to be traded for higher signal-to-noise ratio of the detected signal are referred to as wideband modulation techniques. These techniques are particularly advantageous for optical communication links because the spectral width of the unmodulated carrier is usually orders of magnitude wider than the bandwidth of the modulation signals and because extremely wide channel bandwidths are available.

The tradeoff between channel capacity, C, channel bandwidth, B_c, and signal-to-noise ratio, S/N, is given by Shannon's well known equation

$$C = B_c \log_2 \left(1 + \frac{S}{N}\right) \qquad \text{(J-1)}$$

This equation indicates that, theoretically, it is possible to code a signal prior to transmission such that bandwidth is traded for signal-to-noise ratio. In fact, because of the logarithmic relationship, just a small increase in channel bandwidth should improve the signal-to-noise ratio significantly. In existing RF communication links, channel bandwidth restrictions usually prohibit the use of wideband modulation methods, but this restriction does not exist at optical frequencies.

The tradeoff between B_c and S/N in Eq. J-1 is usually (but not always) accomplished by converting the analog information into a digital signal prior to transmission. By appropriate wideband modulation of the digital signal, the signal-to-noise ratio remains practically constant out to the threshold detection range. With narrowband analog modulation, on the other hand, the signal-to-noise ratio decreases monotonically with increase in range. Therefore, it appears advantageous to employ digital modulation techniques (or other wideband analog modulation such as FM) for information transmission on an optical carrier wave.

What price must be paid for the higher signal-to-noise ratio achieved by wideband modulation? Obviously, additional channel bandwidth is one price, but this is not the only price. All modulation methods that achieve higher signal-to-noise ratio at the expense of wider channel bandwidth, regardless of whether they employ analog modulation or digital modulation, are plagued with a threshold detection level below which performance deteriorates rapidly. Therefore, an additional price that must be paid is a shorter threshold detection range.

As an example of this prevalent threshold effect in a communication system employing wideband modulation, we considered in Appendix D the threshold detection level of systems employing binary modulation and frequency modulation. From the results of Appendix D we were able to conclude that as long as the received signal-to-noise ratio (at the input to the demodulator) is greater than 9 dB, the performance of a digital communications system will be essentially error free. This means, for example, that if a video signal, say a television signal, is transmitted as binary coded data, the signal-to-noise ratio of the resultant picture will remain at a constant level (determined by the quantizing noise of the coding process) until the signal-to-noise ratio (at the input to the receiver) falls below 9 dB.

Since noise power is directly proportional to bandwidth, it is clear that any increase in channel bandwidth will result in a proportionate increase in noise power. It follows, of course, that any modulation method that widens bandwidth also shortens the threshold detection range. Despite the shorter threshold detection range, the tradeoff still favors the use of wideband modulation methods for applications where it is necessary to maintain the postdetection signal-to-noise ratio above 10 dB (e.g., for TV applications).

To show the improvement offered by wideband modulation methods, conventional analog and wideband modulation methods are compared below. (In order to maintain realism, the comparison is based on systems having identical transmitters and receivers, the only difference being the bandwidths needed to accommodate the different modulation schemes.)

As previously mentioned, the signal-to-noise ratio with wideband modulation remains practically constant out to the range at which the threshold detection level is reached and then drops to an unusable value. On the other hand, with narrowband modulation the signal-to-noise ratio drops monotonically as range is increased. Comparing the performance of systems using these two modulation methods, one finds that beyond the threshold detection level of a wideband system, the signal-to-noise ratio of the narrowband system is much higher; however, for television picture quality of practical interest (i.e., signal-to-noise ratio higher than 15 dB), the signal-to-noise ratio available from the narrowband system is aready below the acceptable level of 15 dB. In other words, narrowband systems can provide longer range, but the quality of the received TV pictures at ranges greater than the maximum range of a wideband system is not acceptable.

As a starting point for the comparison, the required signal-to-noise ratio must be specified. Subjective tests have shown that a signal-to-quantizing noise power ratio $(S/N)_q$ of about 30 dB will result in a picture which is subjectively

noise free. That is, a TV picture produced with a 30-dB $(S/N)_q$ in a 5-MHz bandwidth, and viewed at a "normal" distance of four times picture height, appears essentially noise free.

In the narrowband-wideband comparison to follow, the relative performances of four different modulation methods will be compared, namely conventional pulse-code modulation (PCM), delta modulation, narrowband analog amplitude modulation (AM), and wideband analog frequency modulation (FM). The comparison is based on achieving the signal-to-noise ratio needed to realize a noise-free picture (namely, 30 dB).

A. PULSE-CODE MODULATION

The channel bandwidth required for PCM is

$$B_c = B_v \log_2 n \qquad \text{(J-2)}$$

where

B_v = bandwidth of television signal

n = number of gray levels in reproduced picture

As an example, consider the case of a 5-MHz television signal transmitted with an 8-level gray scale (corresponding to a signal-to-noise ratio of about 30 dB as specified above). Such transmission can be accomplished by sampling the video signal at a 10 MHz rate (minimum) and transmitting each sample as a 3-bit code. Thus, for the ideal case[1], the transmitted bit rate would be 30×10^6 bits/s,

[1] The Nyquist sampling theorem indicates that it is theoretically possible to preserve all the information in a signal by sampling it at twice its highest spectral frequency. Moreover, if the ideal sampling signal is used (e.g., a sin x/x function), the samples can be transmitted over an ideal channel (i.e., a channel having flat amplitude response and linear phase response over the passband) having a bandwidth which is one-half the sampling rate.

which could be transmitted over a 15-MHz channel. In this case the channel bandwidth would be three times as wide as the video bandwidth in accordance with Eq. J-2. In practice, however, ideal sampling signals and ideal channel bandpass characteristics are not realized, and, consequently, the required channel bandwidth is actually about twice as large as the ideal value, or 30 MHz. This realistic value will be used for the comparison of PCM with other types of modulation given later.

Once converted to a binary code, the video information can be transmitted by shifting the polarization of the optical carrier, by keying the optical carrier on and off, or by any other form of binary modulation.

Because PCM allows only discrete levels to be transmitted, the system is plagued by quantizing noise. Out to the threshold detection range, only this quantizing noise is apparent in the reproduced picture. That is, as long as the receiver SNR is above the threshold level of 9 dB, wideband modulation gain is realized and only quantizing noise is apparent in the reproduced picture. However, when the received SNR drops below the 9-dB threshold level, the signal-to-noise ratio (at the output of the receiver) drops rapidly to an unusable level.

The signal-to-quantizing noise is determined as follows. Assume that the video signal is quantized into n levels, each level being separated by E volts. Since all levels of the video signal from 0 to nE volts have equal probability of occurrence, the mean-square quantizing error is

$$\bar{e}^2 = \frac{1}{E} \int_{-E/2}^{+E/2} e^2 \, de = \frac{E^2}{12} \qquad \text{(J-3)}$$

Since the peak signal amplitude is nE, $(S/N)_q$ is

$$\left(\frac{S}{N}\right)_q = \frac{(nE)^2}{E^2/12} = 12\,n^2 \qquad \text{(J-4)}$$

It follows that the number of gray levels corresponding to an $(S/N)_q$ of 30 dB is given by

$$30 \text{ dB} = 10 \log \left(\frac{S}{N}\right)_q = 10 \log 12\,n^2 \qquad \text{(J-5)}$$

Solving for n we get $n \approx 8$.[1] In other words, an $(S/N)_q$ of 30 dB corresponds to a picture having a gray scale containing 8 levels.

It is important to note that quantizing noise is inherent with digital modulation; increasing transmitter power or increasing receiver sensitivity will not reduce it. It is also important to note that as long as the received signal-to-noise ratio at the input to the decoder is above the threshold level of 9 dB, system performance will be limited primarily by quantizing noise.

Because a PCM system can reproduce only discrete gray levels, the received TV pictures suffer an effect known as "contour" distortion. One way to overcome this effect is to add coded noise to the TV signal prior to PCM encoding. Removal of the coded noise at the receiver effectively eliminates the contour distortion. Use of noise coding and decoding circuits would, of course, make the system considerably more complex.

The delta modulation method to be described next offers a much simpler way to eliminate contour distortion. It not only surpresses contour distortion, but it

---

[1] Actually, 10 log $12n^2$ for n = 8, is equal to 28.8 dB.

also provides picture quality equivalent to that obtained with PCM with lower transmitter power and simpler coding circuits.

B. DELTA MODULATION

Delta modulation is fundamentally different from PCM in that the analog signal is compared to a reconstructed video signal at each sample or bit time. The difference between the reconstructed video and the original analog signal is used to generate the digital data. The basic delta modulation system is shown in Fig. J-1. The pulse modulator generates either a positive or a negative pulse, the polarity of the pulse being determined by the polarity of the difference between the input signal and the integrated binary output signal. In this way, the video signal is converted into a binary signal directly without using complex coding circuitry. Simplicity is one of the salient features of this type modulator, but, as will be shown, there are other attractive features.

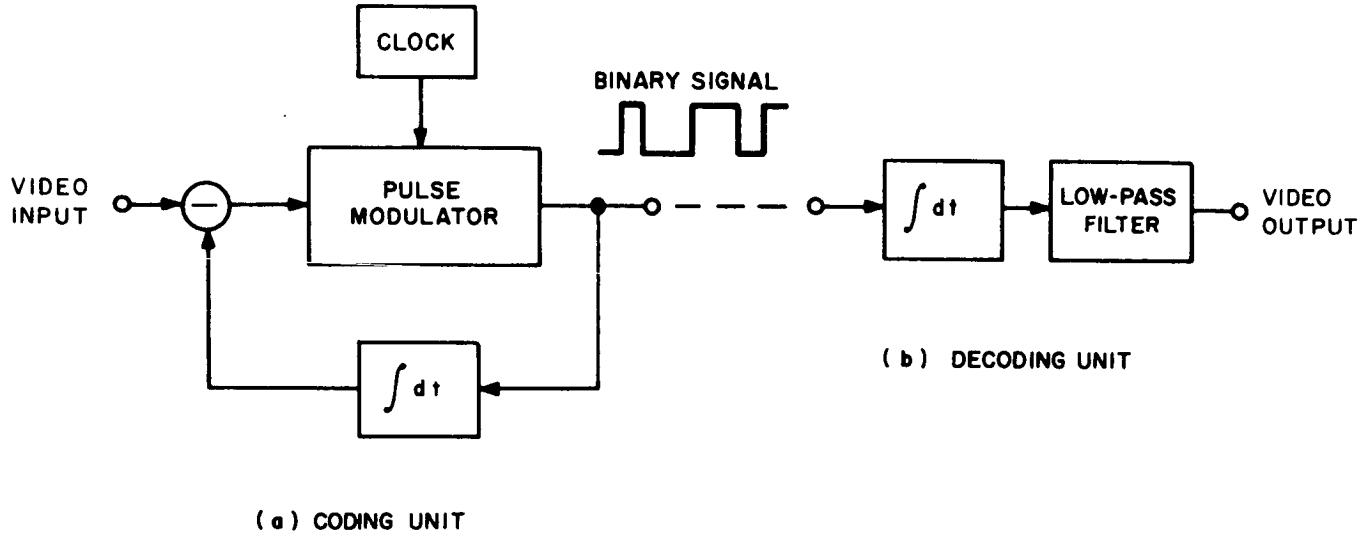


Fig. J-1. Basic delta modulation system. The analog input signal is compared to the integrated binary output signal and the polarity of the binary output signal is switched in accordance with signals from the clock.

The signal-to-quantizing noise power ratio [i.e., $(S/N)_q$] of a communication system using delta modulation has been shown[1] to be

$$\left(\frac{S}{N}\right)_q = M^2 \frac{3\beta_1^4 \gamma^2 \alpha^5}{1 + \frac{4}{3}\pi^2 (1+\gamma^2) \beta_1^2 + \frac{16}{5}\pi^4 \beta_1^4 \gamma^2} \qquad \text{(J-6)}$$

where

M = ratio of signal amplitude to one-half dynamic range

$\alpha = f_r/B_v$

f_r = sampling frequency

B_v = video bandwidth

τ_1, τ_2 = integrator time constants

$\beta_1 = B_v \tau_1$

$\beta_2 = B_v \tau_2$

$\gamma = \beta_2/\beta_1$

In a given system β_1, β_2, γ, B_v, τ_1 and τ_2 are constants. As a result, Eq. J-6 states that the signal-to-noise power ratio is proportional to α^5 or, equivalently, f_r^5. This means that for each octave increase of sampling frequency, the signal-to-noise power ratio will increase by 15 dB.

Balder[2] has shown that a signal-to-quantizing noise power ratio of 40 dB will be realized by a delta modulation system when the sampling frequency is about eight times the video bandwidth. This means that a 5-MHz TV picture would require a sampling frequency of 40 MHz to realize a 40-dB $(S/N)_q$. Therefore, since $(S/N)_q$ is proportional to f_r^5 and since the channel bandwidth required is

1. H. Inose and W. Yasuba, "A Unity Bit Coding Method by Negative Feedback," Proc. IEEE, vol. 51, p. 1527, November 1963.

2. J.C. Balder and C. Kramer, "Video Transmission by Delta Modulation Using Tunnel Diodes," Proc. IRE, vol. 50, pp. 428-431, April 1962.

approximately equal to the sampling frequency,[1] the minimum channel bandwidth of a delta modulation system is

$$B_c = 40\left[\frac{\left(\frac{S}{N}\right)_q}{1x10^4}\right]^{1/5} \text{ MHz} \qquad \text{(J-7)}$$

Eq. J-7 says that, to achieve the 28.8-dB $(S/N)_q$ of a 3-digit PCM system, the channel bandwidth must be at least 24 MHz.

The relative communication range of the PCM and the delta modulation systems can be determined as follows. As shown in Appendix K (see Eq. K-4), the threshold detection range is given by

$$R^* = \frac{D_r}{2\alpha_t}\sqrt{\frac{\rho P_t T_o}{5e}}\,\frac{1}{\sqrt{B_c}} \qquad \text{(J-8)}$$

where

R^* = threshold detection range (m)

P_t = transmitter power (W)

α_t = Transmitter beamwidth (rad)

D_r = diameter of receiver optics (m)

T_o = transmission of optics

e = charge of an electron (1.6×10^{-19} C)

B_c = modulation channel bandwidth (Hz)

ρ = responsivity of photodetector (A/W)

---

1. As previously mentioned (see footnote on page 187), the ideal channel bandwidth is only half this value, but since ideal sampling functions and ideal bandpass characteristics are not realized in an actual system, a realistic value for channel bandwidth is twice the ideal value.

To be meaningful, the comparison of the modulation methods must be based on equivalent system parameters; therefore, we will assume that D_r, ρ, P_t, T_o and α_t are identical for all systems being compared. It follows, therefore, that the ratio of the maximum range of a delta modulation system to the maximum range of a PCM system is given by

$$\frac{R^*_\Delta}{R^*_{pcm}} = \sqrt{\frac{B_{c_{pcm}}}{B_{c_\Delta}}} \qquad \text{(J-9)}$$

Now to achieve an $(S/N)_q$ of 28.8 dB we found above that the bandwidths required for the delta modulation and the PCM systems were 24 MHz and 30 MHz, respectively. Substituting these values into Eq. J-9, we find that $R^*/R^*_{pcm} = 1.12$. Thus, for an essentially noise-free received TV picture $\left[\text{i.e., } (S/N)_q = 28.8 \text{ dB}\right]$, we can conclude that delta modulation will provide a communication range that is about 10 per cent greater than can be obtained with PCM.

At this juncture it is important to point out that delta modulation does not offer improved performance for every application where data is to be coded prior to transmission. For example, if a TV picture is to be reproduced with an extremely high $(S/N)_q$, say 50 dB, PCM would provide longer communications range. However, for the proposed application where an $(S/N)_q$ of 30 dB is adequate, delta modulation does provide superior performance.

C. WIDEBAND FM

Substituting Eq. K-8 into Eq. K-7, we see that the SNR improvement of FM over double-sideband AM (on a power basis) is

$$\frac{\left(\frac{S}{N}\right)_{fm}}{\left(\frac{S}{N}\right)_{am}} \approx \frac{3}{4}\left(\frac{B_{c_{fm}}}{B_v}\right)^2 \qquad \text{(J-10)}$$

From this relationship it is clear that the SNR improvement afforded by FM is proportional to the square of the bandwidth expansion. Obviously, wideband FM has no quantization noise, since it is an analog (rather than digital) form of modulation. Consequently, the SNR at the output of the receiver decreases with range, just as in an AM system, until the 9-db threshold input SNR is reached. However, the SNR of the FM system is higher than that of the AM system by the wideband gain factor of Eq. J-10. Of course, when the input level drops below the threshold level, the wideband gain is no longer realized and the SNR at the output of the receiver drops to an unusable level, just as in the wideband digital system. The analysis given in the latter part of the next section shows that maximum usable range of the FM system is obtained when the modulation bandwidth is 30 MHz. That is, an FM modulation bandwidth of 30 MHz enables an SNR of 28.8 dB to be realized at the threshold detection level.

D. COMPARISON OF MODULATION METHODS

The relative performance of PCM, delta, wideband FM and AM systems will now be compared. The AM system has been arbitrarily chosen as the reference for this comparison. Assuming 100 per cent modulation, the range of the AM system, given in Appendix K (see Eq. K-6), is

$$R_{am} = \frac{D_r}{\alpha_t} \sqrt{\frac{\rho P_t T_o}{4e \left(\frac{S}{N}\right)_{am} B_v}} \qquad \text{(J-11)}$$

Since there is no modulation gain obtained with AM, the maximum range for a noise-free picture (i.e., a picture with an SNR $\geq$ 28.8 dB) is

$$R_o = \frac{D_r}{27.8\alpha_t} \sqrt{\frac{\rho P_t T_o}{4eB_v}} \qquad \text{(J-12)}$$

The signal-to-noise ratio vs range of the AM system, given by Eq. J-11, is plotted in Fig. J-2 (and also Fig. 31) with range normalized to R_o. It is seen that the SNR is inversely proportional to range squared, and that $R/R_o = 1$ corresponds to the range at which the SNR is 28.8 dB.

As previously mentioned, for PCM and delta modulation the $(S/N)_q$ remains essentially constant out to the threshold detection range and then drops rapidly to an unusable value. The threshold detection range for these modulation methods, given in Appendix K (see K-4), is

$$R^* = \frac{D_r}{2\alpha_t}\sqrt{\frac{\rho P_t T_o}{5eB_c}} \qquad \text{(J-13)}$$

Therefore, combining Eqs. J-12 and J-13, we find that the ratio of the maximum range of the PCM and the delta systems to the maximum range of the AM system is

$$\frac{R^*}{R_o} = 12.4\sqrt{\frac{B_v}{B_c}} \qquad \text{(J-14)}$$

Note that B_c in the above equation is the modulation bandwidth and, therefore, its value depends on the type of modulation used. It was previously shown that the channel bandwidths needed for PCM and delta modulation are 30 MHz and 24 MHz, respectively. Substituting these values into Eq. J-14 gives $R^*/R_o = 5.05$ for PCM and $R^*/R_o = 5.65$ for delta modulation.

The maximum range of the FM system is found as follows. The threshold detection level of the FM system is 9 dB, just as in the PCM and delta modulation systems. Therefore, the range improvement factor given by Eq. J-14 is also valid for the FM system, with B_c being the modulation bandwidth of the FM system. Referring to Eq. J-10, note that the SNR of the FM system is increased over that of the AM system by $(3/4)\,(B_c/B_v)^2$, the wideband gain factor. Since the SNR of both

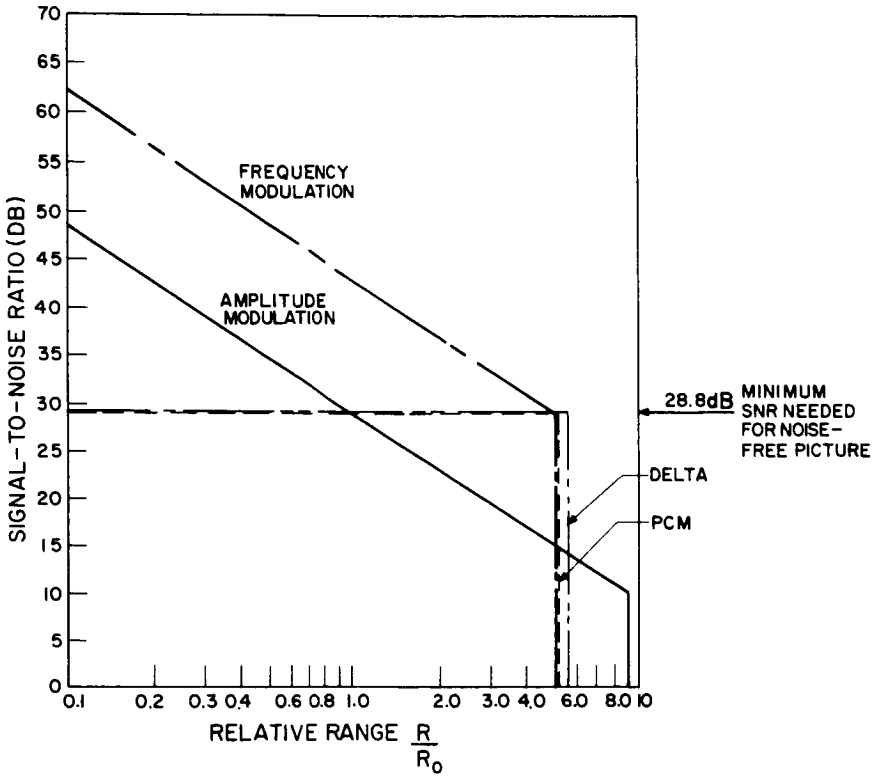


Fig. J-2. Comparison of modulation methods. R_o is the range at which the SNR of the AM system is 28.8 dB.

the FM and the AM systems is inversely proportional to range squared, and since $(S/N)_{am} = 28.8$ dB at $R = R_o$, and since it is desired to make $(S/N)_{fm} = 28.8$ dB at $R = R^*_o$, it follows that

$$\left(\frac{B_c}{B_v}\right)^2 = \frac{4}{3}\left(\frac{R_o}{R^*}\right)^2 \qquad \text{(J-15)}$$

Substituting Eq. J-14 into the above equation and $B_V = 5$ MHz, and solving for B_c we find that the optimum channel bandwidth is 30 MHz. Substituting $B_c = 30$ MHz into Eq. J-14 gives $R^*/R_o = 5.05$ for the FM system.

The results of the foregoing analysis are plotted in Fig. J-2 (and also Fig. 31), along with the SNR vs range curves of the AM system. It is seen that all the wideband modulation methods provide a noise-free picture out to a much greater range than can be realized with amplitude modulation, that delta modulation provides about 10 per cent greater range than either PCM or analog FM, and that PCM and analog FM provide equivalent performance.

One fact previously mentioned is worth emphasizing again at this point. That is, the relative merit of different wideband modulation methods depends upon the desired output SNR. If an extremely high SNR is needed, say 50 dB or higher, then a wideband modulation method such as PCM, which increases SNR exponentially with bandwidth expansion, is undoubtedly the best choice. But, if an SNR of only 30 dB is needed (as in the TV system considered here), other wideband modulation methods provide equal, or better, performance with less complex circuitry.

Appendix K

RANGE OF AN OPTICAL COMMUNICATION SYSTEM AS A FUNCTION OF MODULATION BANDWIDTH

From Eq. A-14, the ratio of signal power to noise power in the output of a multiplier phototube optical detector is

$$\frac{S}{N} = \frac{I_s^2}{I_n^2} = \frac{\rho P_s^2}{2e B_c \left(P_s + P_b + \frac{I_d}{\rho} \right)} \tag{K-1}$$

where B_c is the communication channel bandwidth.

Now as discussed in Sec. II. B, for the proposed application, received background power P_b will be negligible because the receiver will be looking into space from above (or near the top of) Earth's atmosphere. The dark current I_d will also be negligible because, for the wide bandwidths involved, the shot noise current caused by received signal power will be much greater (i. e., $\rho P_s >> I_d$). Therefore, Eq. K-1 can be expressed as

$$\frac{S}{N} \approx \frac{\rho P_s}{2e B_c} \tag{K-2}$$

Consider now the power that must be received to realize the threshold signal-to-noise ratio of 9 dB (or a power ratio of $S/N \approx 10$) as derived in Appendix D. This received threshold power is (from K-2),

$$P_s^* = \frac{20 e B_c}{\rho} \tag{K-3}$$

Substituting Eq. K-3 into Eq. B-3 (with $T_a = 1$) yields the threshold detection range

$$R^* = \left(\frac{D_r}{2\alpha_t} \sqrt{\frac{\rho P_t T_o}{5e}} \right) \frac{1}{\sqrt{B_c}} \tag{K-4}$$

Now consider the situation where narrowband amplitude modulation is used. If the received signal is amplitude modulated with modulation index m, the signal-to-noise ratio at the output of the receiver (after demodulation) is

$$\left(\frac{S}{N}\right)_{am} = \frac{m^2 \rho P_s}{4 e B_c} \tag{K-5}$$

Substituting Eq. K-5 into Eq. K-4 gives

$$R_{am} = \frac{m D_r}{\alpha_t} \sqrt{\frac{\rho P_t T_o}{4e\left(\frac{S}{N}\right)_{am} B_c}} \tag{K-6}$$

Now consider the relative performance of a wideband FM system. Black[1] has shown that the noise improvement of FM over double sideband AM (on a power basis) is

$$\frac{\left(\frac{S}{N}\right)_{fm}}{\left(\frac{S}{N}\right)_{am}} = 3\left(\frac{f_d}{B_v}\right)^2 \tag{K-7}$$

where

f_d = peak frequency deviation of FM

B_v = signal bandwidth

---

[1] H. S. Black, "Modulation Theory," D. Van Nostrand Co., New York, 1960.

Since the bandwidth occupied by wideband FM is

$$B_c \approx 2f_d \tag{K-8}$$

it is evident that the SNR improvement afforded by wideband FM is approximately proportional to the square of the bandwidth expansion.

Moreover, since the threshold detection level of an FM receiver is 9 dB, the same value as in the PCM and delta modulation systems, Eq. K-4 also gives the threshold detection range of the FM system.

Appendix L

CHARACTERISTICS OF ELECTRO-OPTIC CRYSTALS

A. THEORY

The birefringence of a crystal can be expressed in terms of its index ellipsoid, which, in a cartesian coordinate system, is given as

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_z^2} = 1 \qquad \text{(L-1)}$$

in Eq. L-1, n_x, n_y, and n_z are the three principal indices of refraction of the optical axes. When no electrical stress is present, the optical axes coincide with the crystallographic axes. In crystals of the cubic class, such as hexamine and cuprous chloride, Eq. L-1 reduces to that of a sphere; that is,

$$\left.\begin{aligned} &\frac{x^2 + y^2 + z^2}{n^2} = 1 \\ &n_x = n_y = n_z = n \end{aligned}\right\} \qquad \text{(L-2)}$$

This is because crystals of this class are isotropic without an applied field.

If an electrical field is applied to the crystal, the index ellipsoid is deformed and its equation in the original coordinate system is

$$\frac{x^2 + y^2 + z^2}{n^2} + 2r\,(E_x yz + E_y zx + E_z xy) = 1 \qquad \text{(L-3)}$$

In Eq. L-3, r is the electro-optical coefficient of the crystal and E_x, E_y, and E_z are the components of the electric field along the crystallographic axes. These axes, however, no longer coincide with the optical axes.

In order to determine the new optical axes and the new indices of refraction, it is necessary to rotate the original coordinate axes so that they coincide with the principal axes of the deformed index ellipsoid as given in Eq. L-3. In other words, in the new ($x'y'z'$) coordinate system, the index ellipsoid must be put into normal form; that is,

$$\frac{x'^2}{n_{x'}^2} + \frac{y'^2}{n_{y'}^2} + \frac{z'^2}{n_{z'}^2} = 1 \tag{L-4}$$

where $n_{x'}$, $n_{y'}$ and $n_{z'}$ are the three principal indices of refraction when the electric field is applied to the crystal. The transformation that puts the index ellipsoid into normal form is the rotation matrix given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \tag{L-5}$$

where α_1, β_1, $\gamma 1$, ---- are the direction cosines relating the two coordinate systems. The indices of refraction, $n_{x'}$, $n_{y'}$, $n_{z'}$, are obtained from the roots of Φ of the secular equation

$$\begin{vmatrix} \frac{1}{n^2} - \Phi & rE_z & rE_y \\ rE_z & \frac{1}{n^2} - \Phi & rE_x \\ rE_y & rE_x & \frac{1}{n^2} - \Phi \end{vmatrix} = 0 \tag{L-6}$$

and are related as follows: $\Phi_1 = (1/n_{x'})^2$, $\Phi_2 = (1/n_{y'})^2$ and $\Phi_3 = (1/n_{z'})^2$.

There are nine other equations available for determining the direction cosines. That is, there are three sets of equations such as Eq. L-7, one for each root of Φ.

$$\left.\begin{aligned} \left(\frac{1}{n^2} - \Phi\right)X + rE_z Y + rE_y Z &= 0 \\ rE_z X + \left(\frac{1}{n^2} - \Phi\right)Y + rE_x Z &= 0 \\ rE_y X + rE_x Y + \left(\frac{1}{n^2} - \Phi\right)Z &= 0 \end{aligned}\right\} \qquad \text{(L-7)}$$

Each root of Φ used in Eq. L-7 determines one of the three sets of direction cosines, $(\alpha_1\alpha_2\alpha_3)$, $(\beta_1\beta_2\beta_3)$, or $(\gamma_1\gamma_2\gamma_3)$. Here X corresponds to α_1, Y to α_2, and Z to α_3, when Φ_1 is used. When Φ_2 is used, X corresponds to β_1, Y to β_2, etc. When a root of Φ is applied in Eq. L-7, it is not known immediately which of the sets (α, β or γ) has been obtained. This information is obtained by trial and error by putting the sets in the transformation of Eq. L-5 and making sure that the rotated system is a right-handed coordinate system.

When applying an electric field to a crystal with cubic symmetry, three cases are of practical interest: the field perpendicular to the (001) plane, to the (110) plane, or to the (111) plane. These three cases arise because it has been found that crystals of the cubic class can be cut and polished along these planes without too much difficulty. Calculations for the field perpendicular to these three planes have been made and some of the results are given in Table L-1. The results given are the indices of refraction along the new optical axes, and the phase differences when the light is propagated along each of the new optical axes. In this situation the polarized light has components along each of the two remaining optical axes. The phase difference is defined as $\Gamma = (2\pi/\lambda)\,\ell\,(n_s - n_f)$, where λ is the wavelength of the light, ℓ is the length of the birefringent material, n_s is the index of refraction of the slower, and n_f is the index of refraction of the faster axis under consideration.

TABLE L-1. INDICES OF REFRACTION AND PHASE DIFFERENCES

Index of Refraction	$E_{\perp}$ (001) plane	$E_{\perp}$ (110) plane	$E_{\perp}$ (111) plane
$n_{x'}$	$n - \frac{1}{2} n^3 r E$	$n + \frac{1}{2} n^3 r E$	$n + \frac{1}{2\sqrt{3}} n^3 r E$
$n_{y'}$	$n + \frac{1}{2} n^3 r E$	$n - \frac{1}{2} n^3 r E$	$n + \frac{1}{2\sqrt{3}} n^3 r E$
$n_{z'}$	n	n	$n - \frac{1}{\sqrt{3}} n^3 r E$
Phase Difference with Propagation Along the...			
x' axis	$\frac{\pi}{\lambda}\frac{\ell}{d} n^3 r V$	$\frac{\pi}{\lambda}\frac{\ell}{d} n^3 r V$	$\frac{\sqrt{3}\pi}{\lambda}\frac{\ell}{d} n^3 r V$
y' axis	$\frac{\pi}{\lambda}\frac{\ell}{d} n^3 r V$	$\frac{\pi}{\lambda}\frac{\ell}{d} n^3 r V$	$\frac{\sqrt{3}\pi}{\lambda}\frac{\ell}{d} n^3 r V$
z' axis	$\frac{2\pi}{\lambda} n^3 r V$	$\frac{2\pi}{\lambda}\frac{\ell}{d} n^3 r V$	0

The phase differences were obtained using $E = V/d$, where V is the voltage across a thickness d of the crystal, and ℓ is the length of the crystal along the light path. It is seen from Table L-1 that the largest phase difference is obtained when the electric field is perpendicular to the (110) plane and the light is directed along the z' axis. This phase difference is given by $\Gamma = (2\pi/\lambda)(\ell/d)(n^3 r V)$, which can be made large if the length along the light path t is made large and the dimension d across which the voltage is applied is made small. Since the electric field and the light path are perpendicular, this is known as transverse modulation. Table L-1 also shows that when the electric field is perpendicular to the (001) plane and the light is propagated along the z' axis, the dimensions ℓ/d do not appear.

This is because the field is also along the z′ axis, so that ℓ and d are the same dimension. Since the field and the light path are in the same direction, this is known as longitudinal modulation.

B. REVIEW OF BIREFRINGENCE

A birefringent material is illustrated in Fig. L-1. This figure also illustrates the case of a cubic crystal with an electric field applied to it. Here the fast and slow axes would be two of the new optical axes in the primed coordinate system. As illustrated in Fig. L-1, the incident light has an amplitude A, and is polarized with its plane of polarization at an angle θ with respect to the fast axis. The components of the light along fast and slow axes are given by $F = A \cos \theta$ and $S = A \sin \theta$. The phase difference between the two components is given by

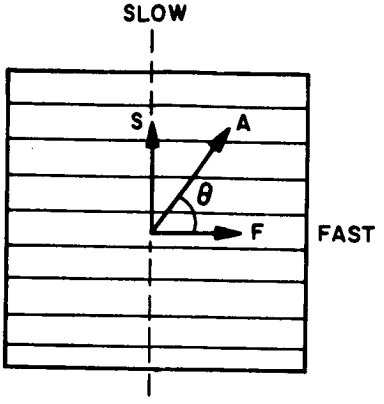


Fig. L-1. Illustration of birefringence.

$$\Gamma = \frac{2\pi}{\lambda} \ell (n_s - n_f) \qquad \text{(L-8)}$$

If the birefringent material is between crossed polarizers, the situation exists as shown in Fig. L-2. The four diagrams of the figure represent end-on views of the light (looking against it) at four different points going through the system. P_1 is the axis of the polarizer. In Fig. L-2(a) the polarized light is shown as it arrives at the surface of the birefringent material with an amplitude A and making an angle θ with the fast axis. This amplitude is broken up into two components [Fig. L-2(b)], $F = A \cos \theta$ along the fast axis, and $S = A \sin \theta$ along the slow axis. The component along the fast axis travels faster in the material, and upon emerging will be advanced in phase relative to the component traveling along the slow axis.

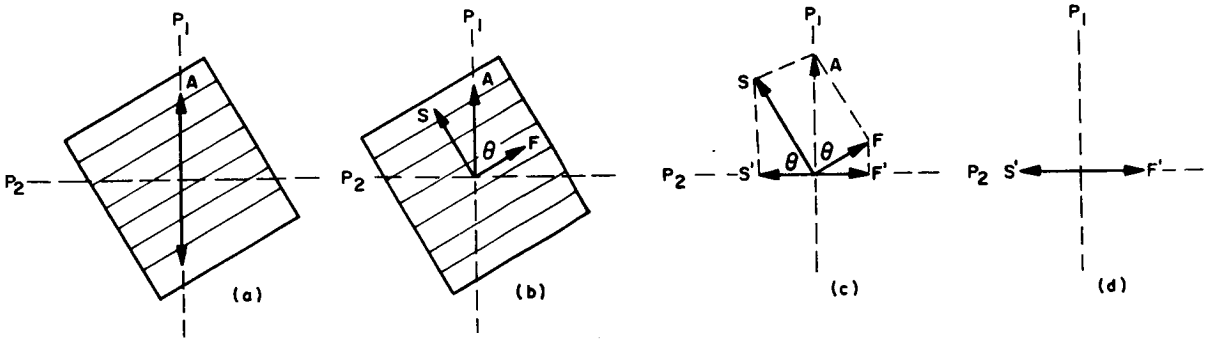


Fig. L-2. Birefringent material between crossed polarizers.

In Fig. L-2(c) these two components are shown as they arrive at the analyzer, whose axis is indicated as P_2. The analyzer transmits only vibrations parallel to P_2. In other words, only the components S' and F' get through, and they are vibrating in the same plane. These components have magnitudes

$$\left.\begin{aligned} S' &= S\cos\theta = A\sin\theta\cos\theta \\ F' &= F\sin\theta = A\cos\theta\sin\theta \end{aligned}\right\} \qquad \text{(L-9)}$$

This result shows that regardless of the angle θ, both components S' and F' transmitted by the analyzer are equal in magnitude when the polarizers are crossed. This result also shows that the maximum amount of light is transmitted by the analyzer when $\theta = 45°$, since $\sin\theta\cos\theta$ has a maximum value when $\theta = 45°$. For this reason, modulation experiments are usually performed with the plane of polarization of the light at an angle of 45° with respect to the optical axes of the birefringent material.

The intensity of the light passing through the analyzer is given by

$$R^2 = S'^2 + F'^2 + 2S'F'\cos(\Gamma + \pi), \qquad \text{(L-10)}$$

where Γ is the phase difference given by Eq. L-8 and π is added since S′ and F′ of the incident light beam are oppositely directed.[1] Inserting Eq. L-9 into Eq. L-10, and after some manipulation, the following is obtained.

$$R^2 = 4A^2 \sin^2\theta \cos^2\theta \sin^2 \frac{\Gamma}{2}$$

For the case of maximum transmission ($\theta = 45°$) this reduces to

$$R^2 = A^2 \sin^2 \frac{\Gamma}{2}$$

Since R^2 is the transmitted intensity and A^2 is the incident intensity, the previous equation can be written

$$I = I_o \sin^2 \frac{\Gamma}{2} \qquad \text{(L-11)}$$

When an electro-optic modulator is used, the phase difference, as shown in Table L-1, is controlled by the applied voltage. Eq. L-11 for this case can be written

$$I = I_o \sin^2 \frac{CV}{2} \qquad \text{(L-12)}$$

where C is the coefficient of V as given in Table L-1. If the modulating signal is a sine wave, Eq. L-12 becomes

$$I = I_o \sin^2 \left(\frac{C V \sin \omega t}{2} \right) \qquad \text{(L-13)}$$

where ω is the angular frequency of the applied voltage. It is seen from Eq. L-13 that the angular frequency of the transmitted intensity has twice the rate of that of the applied voltage. This occurs because I reaches a maximum value when sin $\omega t = 1$, and also when sin $\omega t = -1$; that is, I has a maximum value twice for every

[1] Eq. L-10 was obtained from Jenkins and White, Fundamentals of Optics, McGraw-Hill Book Company, Inc., New York, 1950, p. 207.

cycle of applied voltage. In order to eliminate this "double frequency" effect, the electro-optic modulator must be biased either electrically or optically. The latter is the easier of the two and can be accomplished by placing a quarter-wave plate, properly oriented to obtain circularly polarized light, between the polarizer and the birefringent electro-optic modulator. In Fig. L-2, this corresponds to having the F component ninety degrees ahead of the S component as the light is incident on the modulator. For this case the intensity can be expressed as

$$I = I_o \sin^2\left(\frac{C V \sin \omega t + \pi/2}{2}\right) \qquad \text{(L-14)}$$

From this equation it is seen that unless the voltage is very high the argument never goes negative and, therefore, the frequency of the transmitted intensity is the same as that of the applied voltage.

Although application of the electric field perpendicular to the (110) plane yields the highest per cent modulation, it is usually more convenient to grow the crystal such that the electric field can be applied perpendicular to the (111) plane. As shown in Table L-1, if the crystal is grown such that the field must be applied perpendicularly to the (110) plane, then the light must propagate only along the z' axis; otherwise the per cent modulation will be reduced significantly. On the other hand, if the crystal is grown such that the field can be applied perpendicularly to the (111) axis, the light can propagate along either the x' axis or the y' axis with equally good results. Thus, the reduction of the per cent modulation by the factor $\sqrt{3/2}$ is worth the convenience of not having to orient the crystal accurately.

Substituting the C factor corresponding to applying the electric field perpendicular to the (111) plane (see Table L-1) into Eq. L-14 gives

$$I = I_o \sin^2\left(\frac{\sqrt{3}\,\pi\, r\, n^3 \ell\, V \sin \omega t}{2 \lambda d} + \frac{\pi}{4}\right) \qquad \text{(L-15)}$$

Eq. L-15 shows that the peak-to-peak modulation voltage needed to achieve 100 per cent modulation is

$$V_{\lambda/2} = \frac{\lambda d}{\sqrt{3}\, n^3\, r\, \ell} \tag{L-16}$$

This voltage is known as the half-wave voltage because it results in a 180° shift of the plane of polarization. Substituting Eq. L-16 into Eq. L-15 yields

$$I = I_o \sin^2\left(\frac{\pi\, V \sin \omega t}{2\, V_{\lambda/2}} + \frac{\pi}{4}\right) \tag{L-17}$$

In accordance with Eq. L-17, the per cent modulation is given by

$$m = \frac{\Delta I}{I_o/2}\ (100)$$

$$= \frac{I - \frac{I_o}{2}}{\frac{I_o}{2}} = \frac{I_o \sin^2\left(\frac{\pi\, V \sin \omega t}{2V_{\lambda/2}} + \frac{\pi}{4}\right) - \frac{I_o}{2}}{\frac{I_o}{2}}\ (100)$$

$$= \left[2 \sin^2\left(\frac{\pi\, V \sin \omega t}{2V_{\lambda/2}} + \frac{\pi}{4}\right) - 1\right] 100$$

$$= \left[2 \cdot \frac{1}{2}\left(1 - \cos \frac{2\pi\, V \sin \omega t}{2V_{\lambda/2}} + \frac{\pi}{2}\right) - 1\right] 100$$

$$m = \left[\sin\left(\frac{\pi\, V \sin \omega t}{V_{\lambda/2}}\right)\right] 100 \tag{L-18}$$

Appendix M

OPTIMUM ELECTRO-OPTIC CRYSTAL LENGTH

Assuming a sinusoidal drive signal is applied to the electro-optic crystal, the output power of the transmitter is given by

$$P_o = P_t \left[1 + m \cos \omega_m t\right] \tag{M-1}$$

where

m = modulation factor

ω_m = modulation frequency

P_t = average output power from transmitter

The received power, being directly proportional to the transmitted power, can be expressed as

$$P_s \propto P_t \left[1 + m \cos \omega_m t\right] \tag{M-2}$$

where P_s is the average received signal power.

Since the output current of a photodetector is directly proportional to the received power (i.e., to the radiation incident on the photodetector), the DC detector current is

$$I_{dc} = \rho \left(P_s + P_b + I_d/\rho\right) \tag{M-3}$$

and the signal current is

$$i_s = \rho m P_s \cos \omega_m t \tag{M-4}$$

where

P_b = received background power

I_d = dark current

ρ = responsivity of the photodetector

It follows that the detected signal power is

$$S = i_s^2 R_\ell = \frac{(\rho m P_s)^2}{2} R_\ell \tag{M-5}$$

and the shot noise power is

$$N = 2eB\rho \left(P_s + P_b + I_d/\rho\right) R_\ell \tag{M-6}$$

where

e = charge on an electron (1.6×10^{-19} coulomb)

B = bandwidth (Hz)

R_ℓ = load resistance (ohms)

Thus, the signal-to-noise ratio is

$$\frac{S}{N} = \frac{(\rho m P_s)^2}{4\, eB\, \rho\, (P_s + P_b + I_d/\rho)} \tag{M-7}$$

For the case where the shot noise caused by signal power is negligible compared to that caused by dark current plus background power [i.e., where $(P_b + I_d/\rho) >> P_s$], the signal-to-noise ratio is given by

$$\frac{S}{N} = \frac{\rho (m P_s)^2}{4eB(P_b + I_d/\rho)} \tag{M-8}$$

Eq. M-8 shows that the optimum crystal length is that length which maximizes the factor $(mP_s)^2$.

For the usual case where the voltage applied to the crystal is not sufficient for 100 per cent modulation, the per cent modulation is directly proportional to the length of the crystal:

$$m \propto L \tag{M-9}$$

Due to the transmission loss of the crystal, the transmitter output power is given by

$$P_t = P_1 \gamma^L \tag{M-10}$$

where

P_1 = laser output power

γ = transmission/unit length of crystal

L = length of crystal

Since the received power is directly proportional to the transmitted power,

$$P_s \propto P_1 \gamma^L \tag{M-11}$$

Combining (M-9) and (M-11) we get

$$\left(mP_s\right)^2 \propto L^2 \gamma^{2L} \tag{M-12}$$

The optimum crystal length can now be determined by setting the derivative of this value equal to zero:

$$\frac{d\,(mP_s)^2}{dL} = 0$$

$$2\,K\,L^2 \gamma^{2L}\,(\ln\gamma + L) = 0 \tag{M-13}$$

$$L_{opt} = -\frac{1}{\ln\gamma}$$

Eq. M-13 shows that the optimum length is inversely proportional to the natural log of the crystal transmission. For example, if the transmission is 0.2/cm, the optimum length is 0.62 cm.

For the situation where the shot noise created by background power and dark current is negligible [i.e., where $P_s >> (P_b + I_d/\rho)$], Eq. M-7 reduces to

$$\frac{S}{N} = \frac{\rho m^2 P_s}{4eB} \tag{M-14}$$

In this case it is the factor $m^2 P_s$, rather than $(mP_s)^2$, which must be maximized.

Combining (M-9) and (M-11) we get

$$m^2 P_s \propto L^2 \gamma^L \tag{M-15}$$

Differentiating Eq. M-15 with respect to L and setting the derivative equal to zero yields

$$L_{opt} = -\frac{2}{\ln \gamma}$$

The optimum lengths predicted by Eqs. M-13 and M-15 are plotted in Fig. M-1.

From the foregoing results we conclude that there is an optimum crystal length and that the optimum length is dependent upon the transmission per unit length of the crystal and relative magnitudes of P_s and $(P_b + I_d/\rho)$.

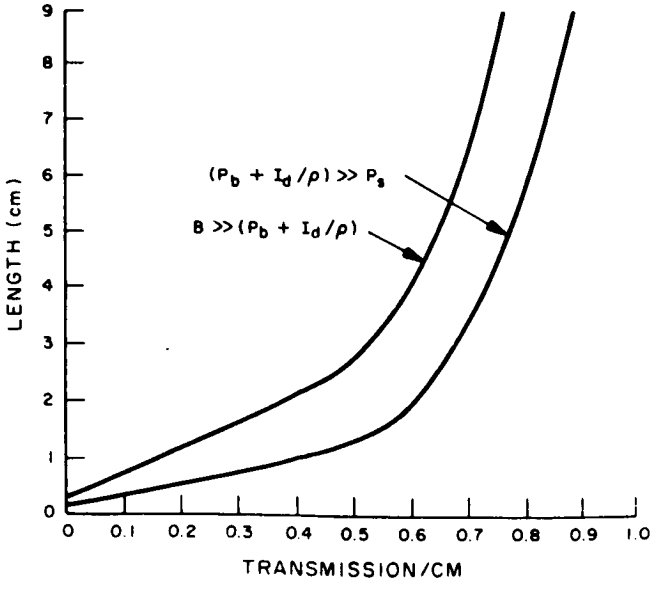


Fig. M-1. Optimum electro-optic crystal length vs crystal transmission.

Appendix N

RELATIONSHIP BETWEEN SNR AND MODULATION BANDWIDTH

The power output of the laser transmitter is

$$P_o = P_t \left[1 + m(t)\right] \tag{N-1}$$

where

P_t = average power output of laser

m = modulation factor

The modulation factor is given by

$$m(t) = \frac{\pi V_m(t)}{V_{\lambda/2}} \tag{N-2}$$

where

V_m = drive voltage across electro-optic crystal

$V_{\lambda/2}$ = half-wave voltage of electro-optic crystal

Consider the situation where it is difficult to achieve a drive voltage sufficient for 100 per cent modulation. For this case the modulation factor is proportional to the crystal drive voltage:

$$m(t) \propto V_m(t) \tag{N-3}$$

Photodetectors, being true square-law devices, generate an output current which is directly proportional to the power incident on their photosensitive surfaces. Hence, the peak AC output current is proportional to received signal power and therefore is also proportional to the transmitted power:

$$i_s \propto m P_o \tag{N-4}$$

It follows that the signal power, being proportional to signal current squared, is

$$S_c \propto m^2 \tag{N-5}$$

The noise power at the output of the detector is

$$N_c = B_m \eta_o \tag{N-6}$$

where

B_m = modulation bandwidth

η_o = noise power density

Thus, the ratio of signal power to noise power is

$$\frac{S_c}{N_c} \propto \frac{m^2 P_o^2}{B_m \eta_o} \tag{N-7}$$

It is well known that the SNR at the output of an FM receiver is related to S_c/N_c, the SNR at the input to the FM demodulator, by

$$\frac{S}{N} = 3\beta^2 \frac{S_c}{N_c} \tag{N-8}$$

where

β = deviation ratio $\triangleq \frac{\Delta f}{f_m}$

Δf = peak carrier deviation

f_m = modulation frequency

For wideband FM the modulation bandwidth is

$$B_m \approx 2 \Delta f \tag{N-9}$$

Combining Eqs. N-7, N-8 and N-9, we can express the SNR as

$$\frac{S}{N} \propto \left(\frac{B_m}{2f_m}\right)^2 \frac{m}{B_m} \tag{N-10}$$

For the case we are considering, where the drive voltage is not sufficient for 100 per cent amplitude modulation, the per cent modulation is inversely proportional to modulation bandwidth B_m (i. e., the drive voltage-bandwidth product is a constant). Therefore, Eq. N-10 can be expressed as

$$\frac{S}{N} \propto \left(\frac{B_m}{2f_m}\right)^2 \frac{1}{B_m^2} \propto \frac{1}{f_m^2} \tag{N-11}$$

which indicates that the SNR is independent of modulation bandwidth. Thus we conclude that no advantage will be gained by trading bandwidth for drive voltage unless 100 per cent amplitude modulation can be realized.