WIDEBAND OPTICAL COMMUNICATION SYSTEMS

BY

PATRICK HAROLD RIEDL

B. S., Kansas State University, 1966

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Electrical Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1968

Approved by:

Donald H. Lenhert
Major Professor

TABLE OF CONTENTS

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

The invention in 1960 of a working laser gave new life to optical communications which is probably the oldest method of communicating over distances beyond earshot. Signal fires were probably the first form of optical communication. A present day example is the blinker light system used by the Navy. A discussion of the history of optical communications is given by Beese [1961]. The above communication methods were very limited in range and in the amount of information that could be transmitted per unit time. Researchers exerted considerable effort to increase the range and information rate but were stymied by the lack of a point source which emitted a spatially coherent and single frequency wave. The laser has fulfilled this need.

The spatial coherence of a laser beam enables it to be focused into a narrow diffraction limited beam. Therefore, a low powered source can place more power per unit area on a distant receiver than a very powerful incoherent source. For instance, the focal plane intensity from a typical gas laser beam is 4 to 5 orders of magnitude greater than that obtainable from the most intense incoherent source [Rempel, 1963].

Since the laser emits single frequency radiation, optical filters may be used in optical receivers to eliminate background radiation except at the laser's frequency. Therefore significantly less optical power is needed to overcome the background noise.

It has been established above that the laser is a much better source for optical communication than an incoherent source. The laser also has advantages over lower frequency coherent radiation sources. A laser's radiation may be focused into a diffraction limited beam whose angle of divergence is $\theta = \lambda/D$, where λ is the wavelength of the laser radiation, and D is the effective source diameter. At lower frequencies λ and D are both larger, but the increase in λ is much larger than the increase in D [Ross, 1966]. For a typical X-band microwave system, $\theta \approx 10^{-2}$ radian while at optical frequencies $\theta \approx 10^{-5}$ radian. The required microwave power to place the same power on an area A as a laser is proportional to the square of the ratio of the divergence angles of the two beams. Thus the microwave source would have to be 10^6 times more powerful than the laser to place the same amount of power on a receiver of area A. Hence the extremely high directivity of laser is an important advantage.

Another important advantage of the laser compared to lower frequency coherent sources is the enormous bandwidths available to carry information. The available bandwidth at a given carrier frequency can be specified as a percentage of the carrier frequency. Assume that one can build a system whose bandwidth is 1% of its carrier frequency. Now compare an

X-band microwave system, 10GHz, and a HeNe laser system, 474 THz. The bandwidth at X-band is 10^8Hz while the laser's is 4.74×10^{12} Hz. Hence as Miller [1966] points out, the communications capacity of an optical communications system is approximately 100,000 times that of a typical microwave channel.

At present, laser communication systems may be characterized as having tremendous potential but doing little communications. This state of affairs arises primarily from economic reasons. Since communication by light is not unique, laser communication systems must compete with microwave, HF, VHF, etc. systems. If these other systems did not exist, laser communications would be feasible at the present time. However, the converse is true. Conventional systems are in a sophisticated state of development. Therefore laser communication systems will be economically feasible only in those areas where the laser's most striking properties can be successfully employed. That is, applications which require high directivity and high information capacity. Even in these areas, laser communication systems will have to be highly sophisticated to compete with microwave systems.

The three areas in which laser communication systems are best suited are secure military communications, deep space communications, and very high data rate terrestrial communications.

The military is vitally concerned with having secure communications. Due to the laser beam's extreme directivity, interception would be almost impossible. Also unlike conventional communication systems, the laser system would be far less susceptable to jamming.

The second area in which laser systems are feasible is deep space communications. The laser's wide bandwidth capability and high directivity will allow a much higher rate of information transfer from a spacecraft to Earth. There is a critical need for such a system. For instance, Mariner IV transmitted pictures of Mars back to Earth at only 8-1/3 bits/second [Brookner et al., 1967]. One might think that the advantage of using a highly directional beam also implies the near-impossibility of pointing and tracking; but Lipsett [1966] has stated that pointing and tracking systems are within the state-of-the-art. Brookner et al. [1964] and Park and Stokes [1967] give an excellent discussion of the value of lasers for use in deep space communications. Lang and Lucy [1967] describe a ground-to-space laser communication experiment.

The third area of application of laser communications, high information rate terrestrial communication, is the topic of this report. These systems will probably take the form of a gigahertz bandwidth, time or frequency multiplexed systems. This application logically belongs to the telephone companies. They will need much more information capacity than at present when video telephones are introduced. The transfer of vast amounts of data between computers is another application for a laser communication system.

The general characteristics of the problem have been covered by Ross [1966]. Miller and Tillotson [1966] have reviewed the results of research on terrestrial optical communications. An excellent comparison between optical and conventional systems is given by Miller [1966]. Research work in this area is generally published in the Proceedings of the IEEE, Applied Optics, and Microwaves.

1.1 Summary of Chapter Development

A communications system is only as good as its receiver. Therefore the system is discussed in reverse order. That is, laser receivers are discussed first in Chapter II. Hence the other parts of the system can then be discussed in terms of the receiver's capability. Of many possible receiver configurations, only those that are applicable to wideband systems are discussed.

Transmission mediums are discussed in Chapter III. Propagation through the atmosphere and through optical transmission lines is described and evaluated to determine what limitations are placed on the system by the transmission media.

In Chapter IV a simple lumped constant phase retardation modulator is described so that the reader may get a basic understanding of optical modulators. The characteristics of advanced modulators necessary for wideband operation are then described.

The general characteristics of the laser are described in Chapter V. After an introduction to some general properties that apply to all types of lasers, specific characteristics of crystalline, glass, semiconductor, and gas lasers are discussed. Finally laser stability, laser noise, the FM laser, and the phase-locked laser are discussed for the gas laser.

Chapter VI describes some feasible laser communication systems and evaluates their relative value.

The conclusions reached in the previous six chapters are summarized in Chapter VII.

CHAPTER II

LASER RECEIVERS

2.1 General Considerations

Many configurations are possible in laser receivers. The primary concern of this chapter is the transformation of the optical signal into a RF signal which may be handled by conventional techniques. This transformation may be accomplished by two general methods. The first is a direct method which uses photodetection. The second is a heterodyne technique which uses photomixing. Both methods are illustrated by Figure 1 in their simpliest forms. The heterodyne receiver requires a local oscillator (LO) at a frequency near the optical carrier.

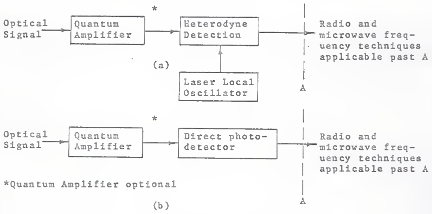

Figure 1. Two laser receivers using a) heterodyne, b) direct detection [Ross, 1966].

The direct detection system responds to variations in the intensity of the laser beam. It doesn't rely on the coherence of the signal while the coherent character of the signal is essential to the heterodyne detection system.

Each of these methods may in principle use an optical preamplifier (quantum amplifier). However no low noise units are presently available [Bloom, 1965].

2.2 Direct Detection System

The direct detection system consists of detecting the incident energy within the frequency response of the detector with the output signal variation following the variations in intensity of the input signal. In order to develop its properties, consider Figure 2 which illustrates a basic direct detection system.

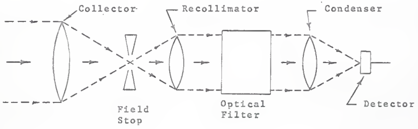

Figure 2. Optical schematic of a direct detector [Fried and Seidman, 1967].

The signal incident on the collector is brought to a focus where it passes through the field stop. The field stop limits the field of view of the receiver. In an electrical engineering sense the directivity of the receiver's antenna is increased. Thus the background noise level may be reduced.

The signal is then recollimated and passed through an optical filter. This filter further reduces the background noise by passing only noise whose wavelength is near the signal's. The field stop (i.e. spatial filter) and the optical filter (i.e. spectral filter) are very important to optical receivers since the photodetector cannot distinguish photons due to the noise and signal provided each are within the spectral response of the photodetector. Also the arrival angle of a photon on the surface of the photodetector cannot be distinguished.

After passing through the optical filter, the signal passes through the condenser which focuses the beam onto the photodetector's surface. As has been noted above photodetectors are square law devices, thus the direct detector responds only to intensity variations in the signal. All phase and frequency information are lost. However Fried and Seidman [1967] point out that the loss of the phase information is not a serious disadvantage. In fact they note that there is no fundamental reason to avoid intensity modulation since the large communication theory advantage that results through the use of FM, PCM, PPM, etc., schemes of coding information may be realized with intensity modulation. For instance the optical carrier intensity may be varied at a high frequency and this frequency changed corresponding to the information to be transmitted (i.e. FM coding but intensity modulation).

The value of a receiver is determined primarily by the signal-to-noise-power ratio (S/N) that may be achieved with the receiver. It is desirable for a receiver to have as high a sensitivity (ability to detect low level signals) as possible. In order to do this one must reduce the noise level as much as possible.

Noise in the detection system is composed of background noise, detector noise, and signal shot noise. Background noise is formed by radiation from the sun, stars, moon, lightning flashs, etc. Ross [1966] gives an excellent discussion of background noise. Some methods of overcoming background noise have already been discussed. Detector noise consists of shot noise due to dark current, thermal noise of the output resistance, and excess noise due to imperfections in the current gain mechanism (if it has one) of the detector [Anderson and McMurtry, 1966]. Signal shot noise is due to the quantized nature of the signal. That is, the signal is composed of individual photons of energy. Thus a natural limit to the sensitivity of a detector is to be able to count individual photons.

An ideal detector may then be defined as one for which background noise and detector noise are zero. Ideal detectors are not possible, but quasi-ideal detectors, for which the background noise and detector noise are less than the inherent signal shot noise are possible. Fried and Seidman [1966] show that quasi-ideal detection can be attained if the quantum efficiency of the photodetector is at least 10% and if the required S/N and bandwidth are large. Thus for wide band operation background noise is easier to overcome than in the narrow band case. Kerr [1967] obtains the same results as Fried and Seidman by a different method.

The S/N for an ideal detector (i.e. signal shot noise limited case) with amplitude modulation is [Kerr, 1967]

$$\frac{S}{N} = \frac{m^2 \eta P_s}{4h\nu B} \quad , \qquad (1)$$

where
m = modulation index
η = quantum efficiency
B = signal bandwidth
P_s = optical power required
$h\nu$ = photon energy.

The instantaneous S/N in a pulsed direct-detection system is

$$\frac{S}{N} = \frac{\eta \ P_s}{2h\nu B} \quad , \qquad (2)$$

where P_s is the average instantaneous power. It is important to note that these S/N's apply only to the ideal case. Many other factors affect S/N when the detector is not signal shot noise limited.

2.3 Photodetectors

Photodetectors to detect optical radiation have been available for many years. However they were not capable of high frequency, wideband, low noise response which is necessary to take advantage of the properties of the laser in a communications system. A figure-of-merit that may be used to evaluate the relative value of photodetectors is $\eta^2 M^2 R_{eq}$ [Anderson and McMurtry, 1966], where η is the quantum efficiency, M is the current gain, and R_{eq} is the interaction resistance between the photodetector and the following amplifier. It is desirable to make the product $\eta^2 M^2 R_{eq}$ as large as possible.

Photodetectors may be divided into two general types. The first is the vacuum detector which uses an external photoelectric effect. Solid-state photodetectors form the second type which employs an internal photoelectric effect.

2.3.1 Vacuum Photodetectors

In the vacuum detector the photon to electron converter is the photocathode (i.e. absorption of photons results in emission of electrons). After a discussion of relevant characteristics of photocathodes and secondary emitters, the relative merits of several high-speed photodetectors will be discussed.

The sensitivity of the photocathode depends on its quantum efficiency (number of free electrons generated per incident photon). The quantum efficiency is a function of the photocathode material and the wavelength of the incident radiation. Quantum efficiencies of 10% to 20% are available in the visible spectrum from 0.3 to 0.6 μm. Efficiencies on the order of 1% are available from 0.6 μm to 0.75 μm. In the near infrared (0.75 μm to 1.5 μm), efficiencies are poor (<0.1%). Above about 1.3μm photoemission does not take place due to photon energies which are insufficient to liberate electrons. Therefore at 10.6μm (CO_2 Laser) vacuum detectors cannot be used. Sensitivity curves for various photocathode materials may be found in Ross [1966] or Anderson and McMurtry [1966].

A major problem of the photocathode, especially when followed by a high gain multiplier section, is the differences in velocity of the emitted electrons. This effect results in smearing-out of the high frequency electron beam and the consequent loss of the modulation signal. Photocathodes by themselves have no internal gain and low R_{eq} which severely limits their sensitivity.

By using secondary emission, electron multipliers can be built which have an internal current gain of 10^6. Thus one of the big disadvantages of the photocathode may be overcome. Probably the most important characteristic of electron multipliers is that they are inherently low noise devices. Their frequency response is limited only by the transit-time dispersion caused by the spread of emission velocity from the photocathode. The transit-time dispersion may be partially eliminated by proper design.

Some specific photodetectors may now be discussed. Vacuum photodiodes have extremely high frequency response. Frequenciea as high as 50 GHz can be detected [Anderson and McMurtry, 1966] if the photodiode is placed in a suitable microwave circuit. However the device lacks an internal gain mechanism and has a low equivalent resistance. Therefore it is suitable to detect modulation on high intensity laser beams only. It would be of little use in a laser receiver which needs high sensitivity.

One detector that has a large internal gain ia the dynamic crossed-field electron multiplier (DCFEM). This detector substantially reduces the transit time dispersion by using spatially uniform crossed fields. One field is a uniform magnetic field and the other field at right angles to it is a microwave frequency electric field [Anderson and McMurtry, 1966]. The output of the detector consists of train of short pulses with repetition frequency equal to the applied microwave electric field. In other words the output consists of an amplified sampling of the photocathode current. Therefore the device may be operated as a baseband detector with an upper cutoff frequency equal to 1/2 the frequency of the microwave electric field, or it may be operated as a bandpass device at a frequency which is higher than the microwave pump frequency. Bandpass operation has been achieved at 3GHz [Anderson and McMurtry, 1966]. The DCFEM has an M^2R_{eq} of about 10^{15} which is high enough to assure signal shot noise limited operation.

Another multiplier detector is the static crossed-field photomultiplier. Like the DCFPM this photomultiplier uses crossed electric and magnetic fields. But in this detector both fields are static. This detector is very effective in removing much of the transit time dispersion which limits frequency response. Power outputs from this device have been reported which were 105 db greater than the output of a simple photodiode using the same output coupler [Anderson and McMurtry 1966]. Uniform baseband response to 5GHz with signal shot noise limited response has been obtained.

A broadband photodetector may also be obtained by using a helix (slow wave circuit) with a photocathode to form a traveling-wave phototube (TWP). In a conventional TWT an ac signal is impressed on a large direct current, and ac amplification is brought about by increasing the depth of modulation (ratio of the difference between the maximum and minimum signal to the maximum signal) of the combined currents. For direct detection there is no direct current; hence the TWP has no amplification. However, the helix behaves as a series of resonant cavity couplers whose output electrical signals are added in phase [Anderson and McMurtry, 1966]; thus one obtains a useful impedance transformation. Very high values of R_{eq} ($>10^5 \Omega$) can be obtained. Furthermore these high equivalent resistances are accompanied by very large bandwidths. Anderson and McMurtry [1966] report R_{eq} $>10^5 \Omega$ over a 10 to 2[illegible] GHz band in a single tube.

One important difference between the TWP and other detectors discussed in this report is that R_{eq} and bandwidth are mostly independent of one another. The R_{eq} bandwidth product is not even approximately constant.

In spite of the TWP's high R_{eq}, signal shot noise operation does not occur since the gain is small. The TWP is more useful in a heterodyne receiver since the local oscillator generates a direct current in the TWP. Thus amplification will occur by increasing the depth of modulation in the direct currents [Lasser, 1966].

By combining the broadband traveling-wave phototube with a high gain electron multiplier, one obtains a high gain wideband detector called the multiplier traveling-wave phototube (MTWP). The MTWP has a passband response extending from 1 to 4 GHz with an M^2R_{eq} of 10^{10} [Kerr, 1967].

Anderson and McMurtry [1966] and Ross [1966] give excellent discussions of the above devices and list extensive references.

In order to achieve signal shot-noise limited detection with a bandwidth of 1GHz or greater, a M^2R_{eq} of 10^9 is required. The DCFEM, static crossed-field photomultiplier, and the MTWP meet the above requirement. The static crossed-field photomultiplier is probably best suited for baseband detection and the MTWP best for bandpass detection [Anderson and McMurtry, 1966].

2.3.2 Solid-State Photodetectors

Solid-state photodetectors useful at high frequencies may be divided into two types: depletion layer devices and photoconductive devices. Basically a depletion layer detector consists of a reverse biased semiconductor junction whose reverse current is modulated by the charge carriers produced near the depletion layer by the absorption of photons [Anderson and McMurtry, 1966]. The depletion layer devices fall into three classifications: junction photodiodes, point-contact photodiodes, and surface-barrier photodiodes. Anderson and McMurtry [1966] discuss these photodetectors in detail. Only the general characteristics of those applicable to wideband high frequency receivers are discussed here.

The junction photodiode is the simplest solid-state detector. It usually consists of a reverse biased surface illuminated p-i-n structure. The diodes frequency response, equivalent resistance, and sensitivity are discussed below.

The quantum efficiencies for silicon and germanium vary from 10% to 50% over a band from 0.5 to 1.7 μm. Thus in the near infrared, the quantum efficiency for the solid-state photodetector is much better than the quantum efficiency for the photocathode devices.

The frequency response of the junction photodiode is strongly dependent on device geometry. Gigahertz frequency response can be easily achieved. The equivalent resistance, R_{eq}

is only about 250Ω. Hence photodiodes are always thermal noise limited. The sensitivity therefore is about 4 orders of magnitude poorer than the ideal photodetector [Anderson and McMurtry, 1966].

The point contact photodiode is capable of extremely high frequency operation (>50GHz); however its active area is only 5μm in diameter. The surface-barrier photodiode has quantum efficiencies greater than 70%. The limitations of these devices are the same as those on the junction-photodiode.

2.3.3 Solid State Detectors With Current Gain

A partial solution to the problem of poor sensitivity is to build solid state detectors with some sort of current gain. Three types will be discussed. They are avalanche photodiodes, photoconductors, and the photoparamp.

The most promising detector with internal gain is the avalanche photodiode. The avalanche process consists of impact ionization in the high field region of a reverse-biased p-n junction. Unlike the gain mechanism in the electron multipliers, avalanche gain is quite noisy. However, the avalanche photodiode has about 50 times better sensitivity than the simple photodiode with 1 GHz bandwidth. Nevertheless, an ideal photodetector is still about 200 times more sensitive than the avalanche photodiode.

Current gain also may be achieved by parametrically pumping a photodiode. This device is called a photoparametric detector or photoparamp. The process does not have true internal gain, but amplifies both the signal and the thermal noise associated with the diode's series resistance. The amplification introduces very little additional noise and gains of 25 db have been reported [Ross, 1966]. A complete discussion of the photoparamplifiers characteristics is given by Penfield and Sawyer [1965]. However the device is narrow band and therefore not suited for wideband communications.

A photodetector useful in the middle infrared is the photoconductor. The transformation of photons to electrons occurs by the generation of charge carriers in a semiconductor or insulator by the absorption of light. Ross [1966] gives an extensive treatment of photoconductors. A common photoconductor is copper-doped germanium (Ge:Cu). In order to increase the photoconductor's sensitivity, it is cooled to liquid nitrogen temperatures and below. The big disadvantage of photoconductors is that their gain-bandwidth product is small. They are not very good wideband detectors but at 10.6μm they are better than most other detectors. Theoretically optimum operation has been achieved with a heterodyne receiver at kHz IF frequencies using Ge:Cu at 4°K [Teich, 1968].

Photoconductors may also be used in the near infrared. Sommers and Gatchell [1966] give an extensive treatment of the photoconductor for use in the near infrared. They see no reason why signal-shot noise limited operation cannot be achieved as the state-of-the-art improves. Gigahertz response might be possible.

An effect similar to the photoconductive effect, the photovoltaic effect in which the photons produce a voltage that can be detected without need of bias or a load resistor, may also be used in the middle infrared. Teich [1968] reports the use of a $Pb_{1-x}Sn_xSe$ photovoltaic detector cooled to 77°K to achieve optimum sensitivity in a heterodyne receiver with MHz IF frequencies and a carrier wavelength of 10.6μm. He also predicts that operation in the GHz range may soon be possible.

2.3.4 Photodetector Comparison

The best photodetectors presently available are the cross-field photomultiplier and traveling wave phototube for the photoemission devices, and the avalanche photodiode for the solid-state photodetectors. When sensitivity is the criterion for selection, the choice of photodetector depends primarily on the specified wavelength. This is amply illustrated in Figure 3. The curves in the figure are drawn for room temperature operation and a bandwidth of 1 GHz has been used for the solid state detectors.

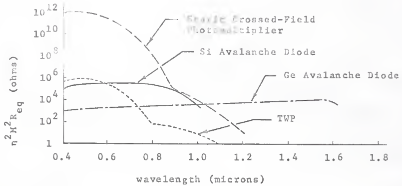

Figure 3. Figure of merit ($n^2M^2R_{eq}$) versus wavelength for some photodetectors. Note that performance of the MTWP and DCFEM is comparable to the static crossed field multiplier [Anderson and McMurtry, 1966].

In the visible spectrum, the vacuum photodetector is by far the best detector since the ideal quantum limits are nearly reached. The solid state photodetector is best in the near infrared region, but its sensitivity is still far from an ideal detector. Prospects for improved photocathodes in this region are very poor, while better solid state detectors are expected.

2.4 Heterodyne Detection Receiver

The basic advantage of heterodyne detection is that by mixing the signal with a strong local oscillator (LO) the output IF signal may be raised above the internal noise, while the LO shot noise swamps the background shot noise. It has been shown by Oliver [1961] that the SNR for an "ideal" optical heterodyne with single sideband (SSB) modulation is

$$\frac{S}{N} = \frac{\eta P_s}{h\nu B} \tag{3}$$

where the terms are defined as in equation (1). This is twice the S/N for the "ideal", pulsed direct detection system. However in practice the ideal heterodyne cannot be approached.

Consider the S/N for a heterodyne receiver which is not perfectly aligned. Miller and Tillotson [1966] state that this S/N is

$$\frac{S}{N} = \cos\phi \, \frac{A_c}{A} \, \eta \, \frac{P_s}{h\nu B} \, \frac{\sin\frac{\pi d\alpha}{\lambda}}{\left(\frac{\pi d\alpha}{\lambda}\right)} \quad , \tag{4}$$

where A_c = common area of signal and local oscillator spots at the detector surface (see Figure 4a)

A = the larger (LO or signal) spot area

η = quantum efficiency

d = diameter of the signal spot

λ = light wavelength (signal = LO)

h = Planck's constant

B = bandwidth of signal

P_s = signal power

ν = light frequency

α = spatial angle defined in Figure 4b

ϕ = spatial angle defined in Figure 4c.

The LO and signal spots should coincide. If the LO spot is larger than the signal spot, the shot noise increases. Wavefront missalignment at the detector's surface introduces a variable phase relationship between the LO and signal waves which yields a (sin x)/x term in equation (4). For an α of one minute of an arc the S/N is reduced 3 db for a 1 mm spot size at 0.63μm [Miller and Tillotson, 1966]. Note that the (sin x)/x term is directly proportional to λ. Thus the S/N at 10.6μm would be much less sensitive to wavefront missalignment than in the visible spectrum. In a similar fashion, polarization missalignment yields the cos ϕ term. Small errors in polarization however do not severely degrade the S/N. Finally if the signal is composed of several modes (quite common in high powered lasers), only the mode which has on the average a component aligned with the LO field will contribute to IF power output.

The above alignment tolerances to achieve ideal heterodyne S/N are very severe.

Another problem arises when heterodyne receivers are used for communications in the atmosphere. It will be shown in Chapter III that turbulence in the atmosphere distorts the signal's wavefront. Therefore the receiver's aperture A_r is limited to the phase coherence area of the signal. The receiver signal power P_s is proportional to A_r, and S/N is proportional to P_s. Therefore the S/N ratio is limited by the receivers aperture. The effect of atmospheric distortion of an optical wavefront on the performance of an optical heterodyne detection system is examined theoretically by Fried [1967]. Mandel [1966] also discusses the limitations placed on a heterodyne receiver by wavefront distortions.

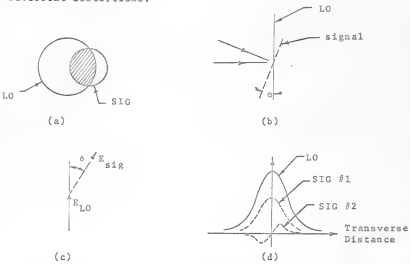

Figure 4. Heterodyne receiving detector alignment effects. (a) Spot coincidence. (b) Wavefront alignment. (c) Field rotational alignment. (d) Mode purity. [Miller and Tillotson, 1966].

The antenna properties of optical heterodyne receivers have been theoretically examined by Siegman [1966]. He proves that the effective aperture of the receiver, A_r, and the solid angular field of view, Ω_r, are related by the antenna law:

$$A_r \Omega_r = \lambda^2 \quad .$$

Consequently there is an inverse trade-off between the directional tolerance and the effective aperture of the receiver. Note that λ^2 is a very small quantity. Thus pointing problems often become difficult. It will be shown in Chapter V that the frequency stability of lasers is not too good. The percentage change in frequency is extremely small but since the frequency is so high a small percentage shift is significant compared to the heterodyne IF frequency. Hence some form of automatic frequency control (AFC) is essential. The AFC system is usually divided into a slow loop and a fast loop (see Figure 5). The slow loop is used to compensate for long term shifts ($t > 1$ sec) in frequency. Slow frequency control of the LO is accomplished by mounting one of the resonator mirrors on a piezoelectric crystal. Varying the voltage across the crystal changes the cavity length and hence changes the output frequency.

Fast loop frequency control is accomplished by feedback to the optical frequency translator. Targ and Bush [1965] show that frequency translation may be obtained by single sideband suppressed carrier (SSBSC) modulation of the LO. However the modulation efficiency is only 0.1%. Kerr [1967] describes some other optical frequency translators that promise to be more efficient.

Targ and Bush [1965] have built and tested an AFC system using a difference frequency of 2.5 GHz. The optical heterodyne receiver is capable of receiving SSB, FM, AM, or pulse modulation.

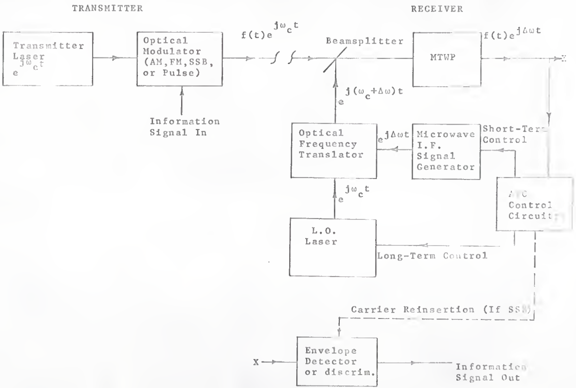

Figure 5. Basic microwave-bandwidth optical heterodyne communication system. The information modulation is denoted by f(t), and the microwave-frequency offset between the transmitter and LO lasers by Δω. [Kerr, 1967]

Conclusions as to the relative value of each type of modulation will be delayed until Chapter 6.

2.5 Other Types of Receivers

In RF systems, the effect of noise in the LO signal is often reduced by using a balanced detector. A receiver using a balanced optical mixer is possible and theoretically the LO noise could be reduced more than 20 times [Ross, 1966]; however, it would be more than twice as complicated as a heterodyne receiver and thus it is not practical at this time.

Nonlinear crystals may be used for parametric amplification or generation of harmonics of the signal frequency. The second harmonic may be more easily detected than the fundamental signal. The major disadvantage is that a powerful source is needed. Lasser [1966] describes several nonlinear detection systems.

Homodyne detection, a special case of heterodyne detection in which the IF is zero, has some advantages. In this case the reference signal is transmitted along with the modulated signal. Its major advantage is that the atmosphere perturbs both the reference signal and modulated signal in the same fashion. Thus the signal remains coherent over a larger part of the photodetector's surface. However it has the disadvantage of less sensitivity since the reference signal does not raise the modulated signal above the internal noise of the detector as in the heterodyne receiver. Further information is given by Aircraft Armaments [1967].

2.6 Summary

The "ideal" direct receiver has a theoretical SNR of 1/2 that of "ideal" heterodyne with SSB modulation. Photodetectors are available in the visible spectrum that are quasi-ideal (i.e. the S/N is close to the ideal S/N). However the heterodyne receiver is beset by a number of practical difficulties. Alignment problems, coherence problems, AFC problems, SSB problems all degrade the heterodyne S/N. In addition the heterodyne is far more complex and expensive than the direct receiver. Hence the direct receiver would be the best choice in the visible and near infrared spectrum.

As the frequency of operation is lowered, quasi-ideal operation is more difficult to achieve because of decreasing quantum efficiency. However tolerances are eased for heterodyne detection receivers. The middle infrared seems to be the most promising spectral range for the application of heterodyne detection.

CHAPTER III

TRANSMISSION MEDIA

The effects of the transmission media on a laser beam are extremely important in the design and successful operation of a laser communications system. The selection of modulation methods, detectors, frequency of operation, information bandwidth, and range between transmitter and receiver are strongly influenced by the transmission medium. In fact the transmission medium is probably the limiting factor in laser communications at present.

There are two possible modes of transmission: atmospheric propagation and guided propagation. Each of these modes are separately discussed below.

3.1 Atmospheric Propagation

Atmospheric propagation of a laser beam has been a subject of extensive research in the past several years. Extensive data for incoherent light propagation is available but very little research has been done for coherent light. Unlike incoherent light transmission which is affected primarily by absorption and scattering, coherent light transmission is also affected by refractive index fluctuations of the atmosphere. These effects are influenced by a large number of variables (e.g. temperature, pressure, precipitation, wavelength, altitude, wind velocity). Since most of these variables are random in nature, deterministic calculations are difficult, if not impossible. The general aspects of propagation in a random medium have been treated by Chernov [1960], while propagation in a turbulent medium has been covered by Tatarski [1961].

The atmospheric absorbs, scatters, changes the direction of, causes loss of wavefront coherence, frequency modulates, and disperses laser beams. The causes and the significance of each effect are discussed below.

3.1.1 Absorption

Light may be severely attenuated by molecular absorption. Carbon dioxide, water, and other gaseous molecules cause most of the attenuation. The amount of absorption is a function of frequency, as is apparent from Figure 6. Therefore when choosing the wavelength of operation of a laser communication system, one must avoid regions of low transmission.

Sufficient power for communications is presently available at 0.488, 0.633, 1.06, 1.15, 3.39, 3.5, and 10.6 μm. Although Figure 6 does not show it, there is a wide atmospheric window at 10.6μm. From Figure 6, it is apparent that 3.39μm lies in a low transmission range where the slope of the transmission curve changes rapidly. When this 3.39μm beam is modulated, sidebands appear which may have different attenuation than the carrier. This variable attenuation causes distortion. The distortion may be reduced by limiting the information bandwidth.

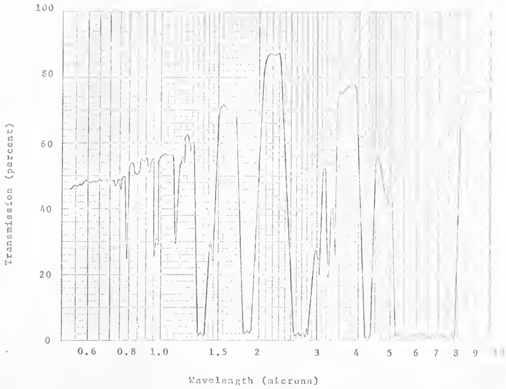

Figure 6. Transmission of Atmospheric Air [Moller, 1964].

3.1.2 Scattering

Atmospheric scattering consists of two types: scattering by air molecules, and scattering by particles. Air molecules are smaller than the wavelength of light; therefore scattering by air molecules is a form of Rayleigh scattering. The scattering coefficient is inversely proportional to the fourth power of the wavelength. Thus the attenuation of shorter wavelength radiation will be much greater than for the longer wavelengths. For example, blue light is severely attenuated.

Scattering by particles, which are approximately one wavelength and larger in diameter, is called Mie scattering. Smoke, fine dust, very small water droplets, and other aerosols are examples of such particles. The attenuation due to particle scattering varies with time. Also note that unlike absorption, scattering is a broad band effect and cannot be minimized by wavelength selection.

Scattering by molecules and small particles is important but the greatest attenuation is due to forms of precipitation (rain, snow, fog). Precipitation attenuation is due to a combination of absorption and scattering. Miller and Tillotson [1966] report the results of experiments to determine transmission loss due to precipitation. The experiment was conducted over a 2.6 km path. The results show that moderate rainfall can introduce more than [illegible] db attenuation, while fog attenuation may be over 60db. Loss due to fog is dependent on the wavelength, with the shorter wavelengths sustaining the higher loss. Snow attenuation ranged from 3.5db/km for light snow to 17.9db/km for heavy snow. Snow, rain, and fog also broaden laser beams which results in loss of power at the receiver. Thus reliable atmospheric communications will require closely spaced repeaters or some other form of diversity to overcome attenuation due to precipitation.

3.1.3 Index of Refraction Fluctuations

Index refraction fluctuations are very important to coherent optical communication systems. The signal degradation introduced by refraction fluctuations cannot be removed directly by an increase of transmitter power as in the case of attenuation by absorption and scattering. Signal degradation takes the following forms: random beam scanning, loss of coherence, random amplitude modulation, random frequency modulation, and polarization fluctuations.

The fluctuations of the index of refraction are caused by thermal turbulence in the atmosphere. An incremental change in temperature of 1°K causes a change in the index of refraction of 10^{-6} [Hondara, 1966]. A turbulent atmosphere contains air blobs or turbulons of different index of refraction which produce angular deviations of the propagated beam. These deviations may affect the entire beam or only a part of it.

When index of refraction discontinuities are large compared to the laser beam cross-section, the result is random beam scanning. Beam scanning is the time-varying change in direction of a beam. The change in direction may be great enough that it completely misses the receiver [Hondara, 1966]. This may be remedied by increasing the receiver's collecting aperture and increasing the beam divergence at the transmitter. However this is done at the expense of received signal intensity. The random beam scanning also degrades heterodyne receiver's performance in a way which cannot be compensated for by increasing the collecting aperture. As a result of bending, the beam's wavefront arrives at the receiver at a small angle. Thus the LO wavefront and signal wavefront are not parallel. Consequently the signal intensity at the IF frequency is reduced. An angular deviation of 10 mrad is sufficient to destroy the signal in a typical system [Hondara, 1966]. This effect may be compensated for by tracking the incident wave. Chase [1966] has shown that the S/N may be increased by a maximum of 11.5db.

Hondara [1966] shows that there is no random frequency modulation in random beam scanning.

Another effect caused by large scale refraction discontinuities is random frequency modulation. Unlike the transverse variation which caused beam scanning, frequency modulation is caused by a refractive index variation parallel to the direction of propagation. The time-varying refractive index yields a time-varying velocity of propagation which implies a loss of time coherence of the wave. Hence the end result is unwanted frequency modulation. The changes in refractive index may be caused by variations in the temperature of the medium or crosswind currents. The frequency modulation introduced in typically between 10kHz and 100kHz [Hondara, 1966].

When the discontinuities in refractive index are small compared to the beam diameter, only parts of the signal wavefront are acted upon. The result is the loss of spatial coherence of the laser beam. That is, the wavefront of the laser beam is no longer planar. Spatial phase fluctuations (loss of spatial coherence) can seriously the performance of both the heterodyne and the direct receiver.

A major requirement of a heterodyne receiver is that the signal be spatially coherent over the collecting aperture. The atmospherically induced loss of spatial coherence of the signal requires that the receiver's collecting aperture be reduced in an attempt to keep the signal coherent over as much of the aperture as possible. Hondara [1966] states several inequalities that must be satisfied for good detection. He also reports that the loss of spatial coherence can be severe enough to cause complete signal break-up.

Interference, both constructive and destructive, occurs in the laser beam due to its loss of spatial coherence. The result of this interference is variation in the intensity of the laser beam which is called scintillation. This effect is also called "breathing" since the beam cross-section changes corresponding to changes in intensity of the beam. Scintillation corresponds to unwanted random intensity modulation. The scintillation noise requires that the modulation depth (ratio of the difference between the maximum and minimum signal to the maximum signal - usually expressed as a percentage) of the signal be greater than some certain value. Hondara [1966] shows that the modulation depth must be greater than 15% for long path lengths and greater than 30% for relatively short path lengths. This is somewhat of a problem since large modulation depths are difficult to achieve at present. A survey of optical scintillation is given by Meyer-Arendt and Emmanuel [1965].

Hondara [1966] has theoretically analyzed polarization fluctuations but does not make a conclusion on the magnitudes of these fluctuations in an actual system.

3.2 Guided Propagation

The adverse effects of fog, rain, snow, and air turbulence may be avoided by using guided propagation. The principal advantage of an optical guide is its extremely low transmission loss. An experimental 970 meter optical guide developed by the Army Electronics Command transmitted a signal with a loss of only 0.5 db/km. [Goubau, 1966]. However guided propagation, like most alternatives, solves the problems of atmospheric propagation but introduces new system limitations. These limitations are the high mechanical tolerances required and the cost of the optical guide.

There are three types of waveguides: tubular guides, surface guides, and beam guides.

3.2.1 Beam Optical Guides

Beam waveguides repeat the field distribution of the light beam at periodic intervals. There are two types of beam waveguide: the lens-type, shown in Figure 7a, and the reflector type, shown in Figure 7b.

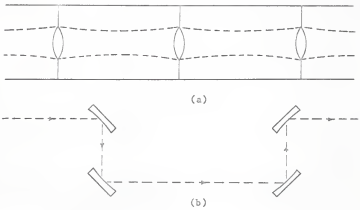

Figure 7. Beam waveguides, (a) lens-type, (b) reflector type [Goubau, 1966].

Miller and Tillotson [1966] classify the lens guide as having continuous guidance if the lenses are closely spaced (about 1 meter) and having intermittant guidance for a lens spacing of about 100 meters. Continuous guidance allows a much sharper curving of the waveguide than intermittent guidance.

However for a radius curvature of less than several thousand meters, reflectors must be used.

Low loss solid lenses are available that may be used in the intermittant guidance system with a loss of only 0.2 db/km [Gloge, 1967]. However for continuous guidance many more lenses are needed and the total loss becomes prohibitive. A solution to this problem is to use a gas lens. Using gas lenses, a loss of 2 db/km has been achieved [Miller and Tillotson, 1966]. A gas lens is formed by passing a gas through a heated tube. The gas along the axis of the tube is cooler than the gas along the heated edge. The resultant temperature gradient acts like a convex lens. Because the beam does not pass through any surfaces the losses are extremely low.

A major problem of beam waveguides is the variation of the index of refraction of the medium within the waveguide. These variations are caused by thermal gradients within the guide. One method of reducing the refraction variations is to partially evacuate the waveguide [Goubau, 1966]. Another is to insulate the guide by burying it underground where the large thermal inertia of the soil holds the temperature approximately constant [Gloge, 1967].

3.2.2 Other Optical Guides

Two other waveguides will be discussed: tubular and surface waveguides. Tubular waveguides are commonly used for microwave transmission. Tubular optical guides take the form of a glass or aluminum tube whose diameter is a large number of wavelengths. Losses are on the order of 1.8 db/km for perfectly straight tubes which have an optical finish [Goubau, 1966]. Any bending greatly increases the attenuation. Therefore tubes are not suited for long distance optical transmission.

Surface waveguides at optical frequencies are glass fibers. The attenuation for large diameter fibers (10μm to 100μm) is about 210 db/km [Goubau, 1966]. For a 0.3μm fiber the loss would still be 20 db/km. Hence glass fibers are not suited for optical transmission.

3.2.3 Summary

Considerable research is being done to determine the mechanical tolerances necessary to build a practical optical waveguide. Present indications are that tolerances will be very stringent and the cost of building the guide therefore very high.

A mathematical analysis of guided propagation is given by Gordon [1966].

CHAPTER IV

MODULATION

A laser may be modulated internally or externally. Internal modulation is accomplished by changing the pump or internal driving power of the laser (i.e. turning the laser "more on" and "more off"). This form of modulation is discussed in Chapter V with lasers. External modulation is accomplished by changing the output coupling (i.e. the laser is always "on").

Almost all optical modulators developed thus far use the variation of the index of refraction of some medium to control the signal. This index may be varied by the electrooptic effect, the magnetooptic effect, acoustic effects, variation in charge carrier concentration in a semiconductor, and combinations of these effects [Miller and Tillotson, 1966]. The most promising and best developed technique for wideband communications is the electrooptic effect. Therefore the electrooptic effect is the modulation technique of major concern in this chapter. Ross [1966], Miller and Tillotson [1966], and Spencer et al. [1967] briefly discuss the other effects and cite many references.

Many different modulation structures have been used in communication experiments. They can be classified as lumped constant, traveling-wave, p-n junction, and cavity-tuned structures. Modulators using the above structures are all basically phase retardation modulators. They differ in the method used to apply the modulating field. Hence the basic characteristics of a simple phase retardation modulator are first described.

Then specific materials and modulators useful in wideband laser communications systems are discussed.

4.1 Phase Retardation Modulators

The operation of phase retardation modulators depends on two factors: interference between two parts of the incident wave, and a method for varying the phase retardation between these two parts. In order to illustrate the modulator's operation, consider a basic retardation modulator shown in Figure 8.

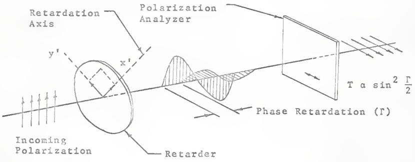

Figure 8. Phase retardation modulator [Hobart, 1966].

The retarder is composed of birefringent material; that is, the index of refraction for a wave polarized in x' direction is different from a wave with y' polarization. In another sense, one would say that the velocity of propagation of a light wave along the x' plane is different than the velocity along the y' plane. Note that in Figure 8 the incoming light is vertically polarized and is oriented at 45° to the two axes of birefringence. The input signal may be resolved into two in-phase equal amplitude components along these axes. After passing through the retarder, the two components will be out of phase; hence the output is an elliptically polarized wave. The eccentricity of the output wave is proportional to the phase retardation Γ which is in turn proportional to the birefringence of the retarder. In an electrooptic material, the birefringence is zero when no field is applied, and increases with the field. Thus the eccentricity of the output wave is proportional to the applied field and a polarization modulator results.

By following the retarder with a polarization analyser which passes the horizontal component, one obtains an amplitude modulator. The intensity of the transmitted wave may be shown to be proportional to $\sin^2 \frac{\Gamma}{2}$ [Hobart, 1966]. When amplitude modulation is desired, it is common to bias the retarder for 50% transmission [Hobart, 1966]. This may be done electrically or optically by introducing $\pi/2$ radians of retardation either before or after the retarder. The retardation is easily accomplished optically by introducing a $\lambda/4$ plate in the beams path. The modulator then works along the linear portion of the transfer curve.

In order to become more familiar with terminology associated with phase-retardation devices, consider a typical transfer characteristic, as shown in Figure 9.

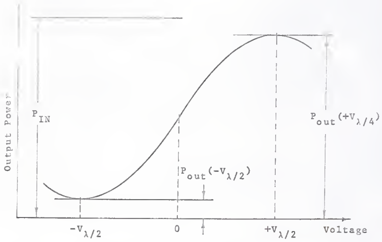

Figure 9. Voltage power output transfer characteristic for a retardation modulator [Hobart, 1966]

The terms appearing in the figure may be defined as follows:

P_{in} is the optical power entering the modulator,

$P_{out}(+V_{\lambda/2})$ is the (maximum) optical power exiting the modulator when the positive quarter-wave voltage is applied,

$P_{out}(-V_{\lambda/2})$ is the (minimum) optical power exiting the modulator when the negative quarter-wave voltage is applied,

$P_{out}(+V_{\lambda/4})/P_{in}$ is the maximum transmission of the cell,

$P_{out}(+V_{\lambda/4})/P_{out}(-V_{\lambda/4})$ defines the extinction ratio which is a measure of the dynamic range of operation.

Manufacturers of voltage retardation cells usually specify the half-wave voltage $V_{\lambda/2}$, which is the voltage necessary to change the retardation by one-half wavelength [Hobart, 1966].

Retardation that utilizes the electrooptic effect is characterized by two effects which depend on the direction of the applied field. The Kerr effect is characterized by retardation which is proportional to the square of the applied field which is parallel to the retardation axis (i.e. the field is usually perpendicular to path of the light beam). The Kerr effect is also sometimes called the quadratic electrooptic effect since the phase retardation is proportional to the square of the applied field. The Pockels effect, or linear electro-optic effect, is characterized by a retardation proportional to the applied field which is at right angles to the retardation axis. The applied field is sometimes parallel to the optical path (longitudinal modulator) and sometimes perpendicular (transverse modulator). The transmission curves corresponding to the two devices are shown in Figure 10.

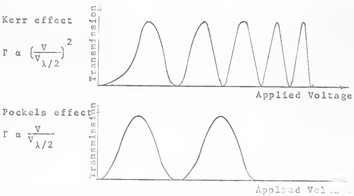

Figure 10. Comparison of Kerr and Pockels Effects [Hobart, 1966].

Before considering electrooptic materials and modulator structures, consider some of the desirable characteristics of a good modulator. These characteristics are low driving power, high optical transmission, and large dynamic range.

Most modulators present essentially a capacitive load to the source. The equivalent circuit for a retardation modulator is shown in Figure 11.

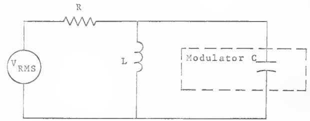

Figure 11. Driving power equivalent circuit for capacitive modulators [Hobart, 1966].

Hobart [1966] shows that the required driving power is

$$P = \frac{2\pi f C V_{rms}^2}{Q} ,$$

where f is the center frequency, Q the quality factor, and V_{rms} the voltage required to attain a specified modulation level. The bandwidth between half power points is $\Delta f = f/Q$, therefore $P = 2\pi(\Delta F) C V_{rms}^2$. Hence the required power is proportional to the bandwidth. The quantity πCV_{rms}^2 is a modulator figure-of-merit and corresponds to the power per unit bandwidth to drive the device. [illegible] minimize this quantity. The factor V_{rms} is inversely proportional to the product of the electrooptic coefficient and the third power of the index of refraction, rn^3, of the material used as the modulator [Ross, 1966]. Therefore by choosing a material with a large electrooptic coefficient and a large index of refraction, the driving power required can be reduced.

High optical transmission through the modulator is another important property for obvious reasons. Low losses require that all optical components be of high quality and that the electrooptic material be transparent to the laser beams.

At 10.6μm (CO_2 laser) losses are a serious problem since most electrooptic materials are opaque at this wavelength. In the visible spectrum the maximum transmission through a typical modulator is 25% [Ross, 1966].

One more characteristic that is important to modulators is the relative response as a function of frequency. A typical response curve is given by Figure 12. At low frequencies the electrooptic effect aided by the piezoelectric response of the crystal. At high frequencies the piezoelectric response is lost because the crystals are mechanically unable to follow the rapid variation in applied voltage. At intermediate frequencies, crystal resonances occur which lead to the dashed portion of the curve. Therefore the crystals must be [illegible] so that they cannot

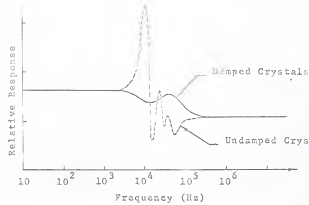

Figure 12. Response of an electrooptic crystal versus modulation frequency [Hobart, 1966].

vibrate. Note that even if the crystals are damped the response is not uniform over the entire frequency range.

4.2 Specific Materials and Modulators

The lumped modulator may use the Kerr effect or the Pockels effect. The Kerr effect is characterized by high dc bias voltages (kilovolts) and large power bandwidth quotient. The common electrooptic material used has been nitrobenzene and carbon disulfide. Stone et al. [1966] have reported continuous modulation at 5 MHz with a modulation depth of 50%. Highly purified nitrobenzene was used as the electrooptic material. The nitrobenzene needed to be highly purified in order to withstand the high dc bias voltage applied to it. A more promising electrooptic material for use in Kerr cells is potassium tantalate niobate

($KTa_{.65}Nb_{.35}O_3$) which is a perovskite mixed crystal [Kaminow and Turner, 1966]. However devices using this material still require considerable development.

Pockels effect modulators use primarily potassium dihydrogen phosphate (KDP) or ammonium dihydrogen phosphate (ADP). Other materials with better properties are known but are extremely difficult to obtain in large enough crystals [Kaminow and Turner, 166]. Commercial Pockel's effect modulators, such as the Spectra-Physics Model 320, are available. The Model 320 has a bandwidth of 75 MHz and a power/bandwidth quotient of about 10 watts/MHz. The required driving power has been reduced from the large values previously needed by using several crystals which are optically in series and electrically in parallel.

The lumped constant modulator's major disadvantage is that its high frequency cutoff is about 100 MHz. This limitation is due to the transit time needed by the incident wave to pass through the modulator. When the period of the modulation field becomes approximately equal to the transit time, modulation is severely degraded because the interaction time between the light and the modulation field becomes very small.

The transit time restriction may be removed if the modulating field is a traveling wave with phase velocity approximately equal to the optical group velocity [Kaminow and Turner, 1966]. The resulting modulator is called a traveling-wave modulator. Broadband operation at GHz frequencies is possible. Peters [1963] first demonstrated a traveling-wave phase modulator capable of

1 GHz bandwidth. He used a strip of electrooptic material between the plates of a parallel plate transmission line. Recently Bicknell et al. [1967] constructed a 0 to 3 GHz traveling-wave electrooptic modulator that requires only 5 watts drive power to obtain 30% amplitude modulation. The modulator uses potassium-dihydrogen-arsenate crystals and has been tested at .63μm. The optical transmission loss was 4db.

Presently the major limitation of traveling-wave modulators is the lack of adequate broadband RF matching sections necessary to couple the modulating signal into the modulator.

If one is willing to settle for a narrower bandwidth, the impedance matching problem may be overcome by placing the electrooptic crystal in a resonant cavity. Band et al. [1966] have reported X-band modulation using a microwave cavity.

They achieved a 3db bandwidth of 150 MHz with a modulation power of 3 watts. KDP was employed as the electrooptic material. Kaminow and Sharpless [1967] have reported intensity modulation at 4.2 GHz using a reentrant cavity. Both $LiTaO_3$ and $LiNbO_3$ were used in the experiment. Trevelyan and Pursey [1967] have demonstrated X-band modulation using KDP in a microwave cavity.

The last modulation structure that will be considered is the p-n junction modulator. Common materials are GaP and GaAs. Very high field strengths across the junction are available to modulate the signal by the electrooptic effect. Kaminow and Turner [1966] estimate that phase modulation with [illegible] GHz bandwidth and 200 mw driving power may be possible in the future.

Further development is necessary to determine the value of the p-n junction modulator.

Thus far in this chapter only polarization and intensity modulation have been discussed. Single sideband and FM are not very well covered in the literature. When the other forms of modulation have been better developed, more interest will probably be shown in SSB modulation. Frequency modulation does not seem to be practical at present. Further information on frequency modulation is given by Ohm [1967], and Aircraft Armaments [1967].

For the reader interested in electrooptic materials an excellent review of the present state-of-the-art is given by Spencer et al. [1967]. They also discuss some of the recent developments in microwave bandwidth optical modulators.

4.3 Summary

Presently the traveling-wave type of modulator is the only one that is capable of microwave response. The main limitation of the traveling-wave device is that efficient broadband RF couplers have not been developed.

The modulation depth capability of optical modulators hinges on the development of better electrooptic materials. There are materials known that have better electrooptic properties than those presently in use, but they can't be fabricated into crystals large enough to be used in modulators. The lack of adequate electrooptic materials is especially critical at 10.6μm. The most promising material at this wavelength is GaAs. A modulation technique for use in the middle infrared that may be better than the electrooptic effect has been described by Peters [1967]. He shows that AM modulation may be accomplished using induced free carrier absorption in n-type germanium. He obtained a modulation depth of 1.2% with a 200 MHz bandwidth; however by refining the system, the modulation depth is expected to reach 40%.

CHAPTER V

LASERS

5.1 General Considerations

A laser oscillator, which is often called a laser, is a regenerative oscillator which is formed by placing an amplifying atomic system in a resonant cavity formed by end reflectors. The amplifying atomic system may consist of excited impurity atoms of a crystal, excited gases, or other systems which involve atoms in an excited state. The amplifying mechanism is called stimulated emission which is characterized by emission of radiation from an excited atom which is in phase with and the same frequency as the radiation which caused the emission.

The resonant cavity is formed by two mirrors, one of which is only partially reflecting in order to couple radiation from the cavity. To determine the cavity resonances, the radiation may be regarded as a uniform plane wave which bounces back and forth between parallel mirrors at a distance L. The resonance condition is that L is equal to an integral number of half wavelengths in the material. (i.e. $L = n\lambda/2$, where n is an integer). Since frequency equals the average velocity of propagation c divided by the wavelength λ, the resonant frequencies are $nc/2L$.

Unlike most microwave cavities whose dimensions are on the order of a wavelength, the dimensions of the laser cavity are on the order of hundreds of thousands of wavelengths. That is the integer n is on the order of 10^5 to 10^6. The resonant frequencies defined by nc/2L are called the longitudinal, axial, or primary modes. In addition to these modes, other resonant frequencies are possible as is illustrated in Figure 13. These are called off-axis or transverse modes and they depend on the geometry of the laser cavity. The optical path length for these modes is slightly different from the axial modes; therefore their frequencies lie between the axial frequencies. The

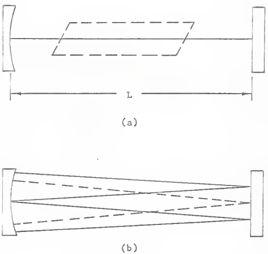

Figure 13. Laser cavity configurations. (a) Basic cavity [illegible] longitudinal mode. (b) [illegible] modes (two shown). [Kussmen[illegible]].

transverse modes are usually undesirable and can be eliminated by proper design of the laser cavity.

Although the laser cavity is resonant for many frequencies, only those frequencies, for which the amplifying medium has a gain greater than losses, appear in the laser's output spectrum. The amplifying medium exhibits gain at frequencies corresponding to the energy difference between energy levels for which stimulated emission occurs according to the equation $E = h\nu$. The medium has a gain over a certain range of frequencies. Thus as shown in Figure 14 only those frequencies within the laser gain profile oscillate. The laser gain profile for the gas laser is due to the Doppler effect. The atoms in a gas laser have random velocities as high as 10^6 cm/sec due to thermal motions caused by the gaseous discharge used to excite the laser. These random velocities shift the atom's emission frequency over a 1000

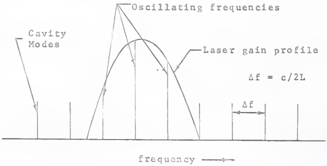

Figure 14 Cavity modes and laser gain profile [Nussmeier, 1967].

MHz band [Warren, 1965]. For a laser 1 meter long, the longitudinal mode spacing is approximately 150 MHz; thus many axial modes oscillate.

In order to generate high power with a laser, it is necessary to use longer lasers. Since the axial mode interval, c/2L, is inversely proportional to the laser length, the frequency spectrum of high powered lasers has more frequencies than low power lasers.

It is desirable to have a high powered transmitter, but the multimode output limits the modulation bandwidth. Consider the special case in which the laser output consists of three modes. The modulated output is shown in Figure 15. Clearly, interference would result if the modulation sidebands overlap. Thus

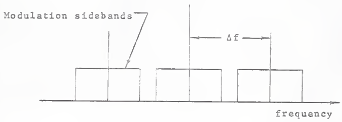

Figure 15. Frequency spectrum of a modulated multimode laser beam.

the useful bandwidth is limited to less than one half of the axial mode spacing. Also note that if the laser has off-axis or transverse modes the modulation bandwidth is further restricted since the frequencies of these modes lie somewhere between the primary modes.

The terminology concerning the laser output spectrum is somewhat confusing due to overuse of the word "mode". Therefore a review is in order. If the laser has no off-axis or transverse modes then it is a called single transverse mode laser. However a single transverse mode laser may operate at several frequencies since the laser may exhibit a gain for several different longitudinal or axial modes. It is possible to build a single transverse mode, single frequency laser at the expense of power and efficiency (e.g. Spectra Physics Model 119 laser which has 100μ watts output at 0.6328 μm).

The field strength of a single transverse mode laser has a Gaussian distribution with respect to radial distance from the beam axis. However if the laser has several transverse modes the field strength is no longer Gaussian. The reason for this is easily seen by referring to Figure 13. Note that the wavefronts associated with the transverse modes are not parallel to the axial mode. Therefore the wavefronts will interfere constructively and destructively across the beam cross-section. This behavior is unsuitable for communication purposes. An excellent discussion of the electric field of a laser is given by Kogelnik and Li [1966]. They also show photographs of various higher order laser beams.

There are four types of lasers available for use in communications. Each has its own particular advantages and disadvantages.

At present most communications research is being done with gas lasers; however, semiconductor lasers, glass and crystaline solid lasers all have some desirable characteristics. Each of these will be discussed below, concentrating primarily on the gas laser. The technique of phase locking a laser to generate a train of nanosecond pulses is discussed. The multimode spectrum of a high powered laser may be transformed to a single frequency output by using an FM Laser and an external modulator. This technique is considered in some detail. Laser noise and stability are also considered.

Tomiyasu [1967a, 1967b] has compiled an extensive bibliography on lasers and laser devices.

5.2 Crystalline Solid State Lasers

The crystalline solid state laser was the first type of laser developed. Laser action in ruby ($Al_2O_3:Cr^{3+}$) was first achieved in 1960 by Maiman [Kiss and Pressley, 1966]. Since then crystalline lasers have been developed that lase from 5985 angstroms to 2.61 microns. Kiss and Pressley [1966] list 29 crystalline lasers specifying their host crystal, dopant, frequency, temperatures of operation, and references.

Crystalline solid lasers are optically pumped with flash lamps (i.e. the flash lamp supplies the energy to excite the impurity atoms in the laser crystal). One of the major reasons for the low efficiency of the laser is that the impurity atoms absorb energy only over a small range of wavelengths. However the radiant energy from the flash lamps is spread over a much larger range. The energy that is not absorbed must be dissipated as heat. Since typical efficiencies are on the order of 0.1%, heat dissipation is a major problem. Hence crystalline lasers are usually pulsed. The pulse repetition rate is limited by the heat dissipation problem as well as the availability of flash lamps that will operate at high frequencies. Present flash lamps are limited to repetition frequencies on the order of 10 kHz [Brookner et al., 1967].

Therefore the principal characteristic of the pulsed crystalline laser, very high pulse power, does not appear to be applicable to high information rate communication systems due to the low pulse rate.

However if one is willing to sacrifice high pulse power, the laser may be operated continuously. Kiss and Pressley [1966] list 8 CW solid lasers. One watt output at .69 μm has been obtained using ruby with a mercury arc lamp. At 1.06 μ, neodymium in yttrium aluminum garnet ($Y_3Al_5O_{12}:Nd^{3+}$) developes up to 200 watts with an efficiency of 0.2%. Power output depends on the type of optical pump used. Also holmium-doped yttrium aluminum garnet ($Y_3Al_5O_{12}:Ho^{3+}$) has been reported to lase at 2.12 μm with an efficiency of 5% and output of 15 watts [Kiss and Pressley, 1966]. Its major disadvantage is that it must be cooled to liquid nitrogen temperatures (i.e. 77°K). This laser has a theoretical efficiency of 30%. Power limitations have been set by cooling capability rather than pump power.

Continuously operated crystalline solid lasers have a multimodal output spectrum. Single frequency operation has been achieved at reduced power and efficiency.

Perhaps the most important property of the CW crystalline solid laser in regard to wideband communications is that it may be phase locked. The time domain output of a phase-locked laser, which is discussed in a later section, consists of a train of nanosecond pulses.

In summary the CW crystalline laser has many desirable advantages for communications applications. However not enough research has been done to fully evaluate its capability as a laser transmitter.

5.3 Glass Solid Laser

The glass laser is formed by doping glass with ions of one of the following rare earth elements: neodymium (Nd^{3+}), holmium (Ho^{3+}), ytterbium (Yb^{3+}), and erbium (Er^{3+}). Of these neodymium is most important since it lases at room temperature and with high efficiency. Glass, as a laser host material, has the advantage of being easily fabricated to a high degree of precision. Its major disadvantage is its low thermal conductivity [Snitzer, 1966]. This requires that the laser be pulsed and also limits the pulse rate which is usually slower than the pulse rate of the crystalline solid laser. [illegible] CW operation will be more difficult with a glass laser due to the low thermal conductivity.

The frequency spectrum consists of many modes and is extremely unstable. Thus at present the glass laser is not suitable for use in a communications system. An excellent discussion of glass lasers is given by Snitzer [1966].

5.4 Semiconductor Lasers

Laser action in GaAs, a semiconductor, was first demonstrated in 1962 [Nathan, 1966]. Since that time, semiconductors have been found that lase at frequencies ranging from ultraviolet to the far infrared.

The semiconductor laser differs from other lasers in two important respects. In solid lasers and gas lasers a photon of energy is emitted when an electron falls from a one energy level to a lower one. The frequency of the radiation depends on the energy difference between the two levels according to the relation $E = h\nu$. However in the semiconductor discrete energy levels do not exist, the allowable energy levels are continuous bands. Therefore the laser will exhibit gain over a wider range of frequencies than other types of lasers.

The other important difference is that a variety of pumping methods may be employed to excite the electrons in the semiconductor. The four methods commonly used are injection using a p-n junction, electron beams, optical pumping, and avalanche breakdown injection. The p-n junction method is the simplest. It requires a properly doped p-n junction and a low voltage dc power supply. It is [illegible]. The primary difficulty is that the [illegible] difficult to build.

The electron beam pumping method requires a high energy stream of electrons, 10 kV or higher, directed onto the semiconductor. Optical pumping uses another laser as the pump. Avalanche breakdown injection generates so much heat the laser can only be operated in a pulse mode. In the last three methods all the compactness and simplicity of the p-n junction laser system are lost; therefore they will not be considered further. Nathan [1966] describes each of the above pumping techniques; furthermore he discusses the state-of-the-art of semiconductor lasers.

The resonant cavity for the semiconductor injection laser (i.e. pumping by means of a p-n junction) is formed by polishing two opposite ends of the semiconductor material such that the ends are perpendicular to the junction. The wavelength of the axial modes is given by $m\lambda = 2n_oL$ where m is the mode number, L is the distance between the reflecting ends, and n_o is the index of refraction [Nathan, 1966]. The spacing between modes is then approximately 1 to 2 angstroms for a typical L of 10 to 20 mils. This spacing corresponds to a frequency difference of 30 GHz for a semiconductor laser whose wavelength is 1 μm. Most lasers have a gain over a 100 angstrom interval which corresponds to an interval of 3000 GHz in the frequency domain. These semiconductor lasers have a multimode output which precludes their use with a heterodyne receiver.

The directivity of the semiconductor laser is poor compared to other lasers. The laser radiates energy over about 10° in the horizontal plane by 16° in the vertical plane. Thus a collimating lens will be necessary to focus the radiation into a narrow beam for communications purposes.

Heat dissipation is one of the major limitations of injection semiconductor lasers. In order to operate an injection laser CW, it is necessary to provide cyrogenic cooling. By operating at 4°K (liquid helium), CW outputs on the order of 10 watts have been obtained [Nathan, 1966]. However increasing the temperature to 77°K (liquid nitrogen) reduces the output to only 1 watt.

The laser may be operated at room temperature in a pulsed mode with a very low duty cycle. For example the RCA TA2930 GaAs injection laser diode has a pulse power output of 50 watts with a duty cycle of only 0.02% and a repetition frequency less than 1 kHz.

Perhaps the greatest advantage of the injection laser is the ease with which it may be amplitude modulated. Amplitude modulation is accomplished by simply modulating the pumping current. The frequency response limitations have not yet been determined due to the lack of adequate high speed detectors. However amplitude modulation of a GaAs laser cooled to 4°K has been reported at 11 GHz [Nathan, 1966]. The laser may also be pulse modulated [illegible] frequency modulation is possible using [illegible] temperature control since the laser's frequency shifts 20 GHz/°K

The semiconductor laser's high frequency response makes it very attractive for communications systems. Potentially its efficiency would be as great as 30 to 60% at room temperatures. Its major disadvantage at present are its poor frequency stability, multimode spectrum, and the requirement for cyrogenic cooling. It has been shown by Schiel et al. [1965] that the transmission of 24 voice channels using pulse code modulation (PCM) over a distance of 13 km is possible using a GaAs injection laser. The semiconductor laser might find its greatest application in deep space communications where its simplicity and compactness can be employed to great advantage.

5.5 Gas Lasers

The gas laser has received the most attention in communications research. Continuous operation at a single frequency has been achieved and many wavelengths are available. In fact, for most purposes, a gas laser can be found that will be superior to optically pumped solid state lasers or semiconductor lasers. The exception being a requirement for high pulse power. Gas lasers have been developed that lase from the ultraviolet near 2000 angstroms to the far infrared at 400 microns. The gas laser is also easier to design than other lasers because the active gas provides an optically homogeneous medium with a low index of refraction. This allows application of simple mathematical formulas to determine the probable output mode structure. The qualities of the output laser beam (i.e. spatial coherence, noise level, and frequency stability of a gas laser are better than the other lasers.

Gas lasers may be pumped or excited by a RF field or a dc discharge through the gas. An excellent review of gas laser state-of-the-art is given by Bloom [1966].

Gas lasers may be divided into three general classes: neutral atom gas lasers, ion lasers, and molecular gas lasers. Each class will be discussed separately below.

5.5.1 Neutral Atom Lasers

The neutral atom lasers are primarily infrared lasers. The important exception is the HeNe laser which lases at 6328 angstroms 1.15 microns, and 3.39 microns. A table listing the different types of gas lasers and their parameters is given by Bloom [1966]. The neutral atom lasers are generally low power devices with 100 milliwatts regarded as a high figure with powers on the order of 1 mw more typical. Power efficiencies are generally low, about 0.1%. Single frequency operation can be achieved with decreased efficiency and power output. Single frequency power is typically about 100 microwatts. At higher powers the frequency spectrum becomes multimodal. However techniques are available to convert this multimode output to a single frequency or to a pulsed output. These techniques will be discussed later.

5.5.2 Ion Lasers

Ion lasers radiate primarily in the visible spectrum, but the spectrum also extends into the ultraviolet. Probably the most common ion laser is the argron ion laser which radiates at 4880 angstroms (blue) and 5145 angstroms (green). Power outputs of 1 to 10 watts being practical from a efficiency viewpoint. Optimum efficiency ranges between 0.3 and 0.03%. Single transverse mode operation can be achieved by proper design. The output from high power ion lasers is multimodal.

It should be noted that the ion lasers are rather heavy, bulky, and expensive. They require water cooling to dissipate the heat due to the poor conversion efficiency of electrical power to light.

5.5.3 Molecular Gas Lasers

The molecular gas lasers operate in the infrared region of the spectrum, usually beyond 5 microns. The recently developed CO_2 laser is the most important molecular gas laser. More than 100 watts output power is easily attained at 10.6 microns from a CO_2 laser which has small amounts of nitrogen and helium added. This output was attained for CW operation with an efficiency of about 10%. [Whitehouse, 1967]. In fact, power outputs on the order of a kilowatt with efficiencies on the order of 20% are feasible. Unlike other lasers, the CO_2 laser may be used as an amplifier to increase output power. The laser may be designed for multimode or single mode operation. An excellent paper by Whitehouse [1967] describes the design, theory, and state-of-the-art of CO_2 lasers.

The use of the CO_2 laser for communications hinges on the development of efficient detectors and modulators. Present materials absorb to much of the laser beam leading to the problem of cooling the modulator crystal. However, the possibility of being able to beam kilowatts of power in a narrow beam is incentive for continued research.

5.5.4 Frequency Stability

The frequency stability of gas lasers percentagewise is good but the frequency variation compared to microwave frequencies is substantial. The frequency of a typical gas laser may change 10^7 to 10^8 Hz in a few seconds. Much larger variations occur over longer time intervals. This instability is caused by variations in the resonant frequency of the cavity. The frequency of operation is critically dependent on optical path length changes. For instance, a minute change in optical path length of 4 angstroms will shift the laser's frequency by 1 MHz. Path length changes of this order are easily generated by acoustic vibrations. The optical path length also depends on the ambient temperature, since the laser's mirrors are separated with spacers whose length depends on the temperature. Hence the spacers are usually made of Invar which has a small temperature coefficient of expansion. Even with Invar spacers, a change in ambient temperature of 1°C will shift the frequency of the laser 500 MHz [White, 1967]. The laser is also sensitive to changes in ambient pressure with a figure of 20 MHz/Torr being typical.

The stability problem consists of both short term (time < 1 sec) and long term (> 1 sec) stability. Thus stabilization techniques must have a fast response for the short term instability and a wide dynamic range to stabilize the long term shifts. Hence stabilization is usually accomplished by using two closed sero loops (a fast loop and a slow loop). The laser's frequency is corrected by changing the optical path length which is usually accomplished by mounting one mirror on a piezoelectric transducer.

In order to detect changes in the laser's output frequency, an optical discriminator is needed. White [1967] describes in some detail each of the following four types: laser-atomic-line, passive-cavity, Zeeman-cell, and passive-cavity Zeeman-cell combination discriminators. With these discriminators it is now possible to control a laser's output frequency to better than 1 part in 10^9 over long time intervals.

In communication systems an optical discriminator is seldom used since in heterodyne receivers where frequency stability is needed the difference frequency is used as the error signal. Direct detection systems do not require a high degree of frequency stability.

5.5.5 Gas Laser Noise

In laser communications systems one must detect highly attenuated laser radiation. Noise is the determining factor whether the signal can be detected. Therefore it is important to know the noise properties of the laser.

Unlike detector noise which is proportional to the final system bandwidth, laser noise is proportional to the emission linewidth (i.e. wavelength range over which laser gain is greater than losses) [Bloom, 1965]. Laser noise may be divided into three general classes: spontaneous-emission noise, plasma noise, and mode-interference noise.

The spontaneous-emission noise is due to the excess noise of a laser acting as an oscillator. This noise has an effect on both the amplitude of the signal and on its instantaneous frequency. The effect on frequency is small compared to path length changes, so only amplitude effects will be discussed. Bloom [1965] indicates that the amplitude component of spontaneous-emission noise is detectable only when the gas laser is operating very close to threshold, otherwise it is very small. Therefore spontaneous-emission noise should be no problem with high powered communication transmitters since they operate well above threshold.

Plasma noise is noise introduced into the laser beam by fluctuations in the plasma current density and represents a macroscopic fluctuation in the gain of the laser. This type of noise is a serious problem in dc excited gaseous lasers since the noise intensity may be as much as 20% of the lasers output [Bloom, 1965]. However by using RF to drive the plasma, the plasma noise is reduced to an insignificant level.

Mode-interference noise occurs in lasers operating simultaneously in several modes. This is not "noise" in the usual sense of white noise, but corresponds to discrete frequency outputs that are uncontrolled and cause interference with the desired signal. Single frequency lasers are not bothered by mode-interference noise.

As laser power is increased these unwanted and uncontrolled frequencies become more of a problem. Various methods of internal perturbation of the laser cavity can eliminate this type of noise. These methods result in a FM laser or phase-locked laser.

The power in these noise modes is a function of the laser's design. No simple quantitative equation may be given for this noise. The magnitude of the problem of mode-interference depends upon the specific laser under consideration.

5.6 Phase-Locked Lasers

A high powered laser usually operates in many axial modes with random phase relations and unstable mode amplitudes. This laser is not well suited to obtain the goal of very wide bandwidths and high information capacity since the bandwidth is limited to the axial mode interval spacing. A technique which increases the information capacity of the multimode laser is called phase-locking the laser. Figure 16 illustrates the basic configuration of a phase locked laser.

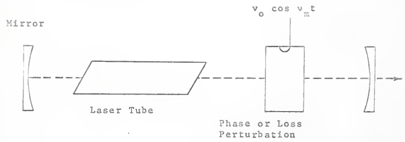

Figure 16. Configuration of a phase locked laser.

Yariv [1965] has shown that, by introducing either a time-varying phase or loss perturbation into the cavity of a multimode laser, the axial modes can be locked to a common reference phase provided the frequency of the perturbation equals the axial mode interval frequency c/2L. Under these conditions the time domain output of the laser is a series of pulses whose repetition frequency is c/2L. The pulse width compared to the period of the pulse train is quite small. Figure 17 illustrates the phase-locked laser's time domain output. Note that the peak intensities are approximately 6 times the intensity of the free-running laser.

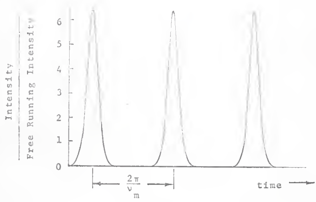

Figure 17. Output intensity vs. time for a phase locked laser [Harris, 1966].

The phase locked laser may be used with direct or heterodyne detection systems. Either pulse amplitude or pulse code modulation may be used. Denton et al. [1966] have demonstrated pulse code modulation of a phase-locked HeNe laser with a bit rate of 224 Mbit/sec. Harris [1966] gives a detailed discussion of the phase-locked laser.

5.7 FM Lasers

The FM laser like the phase-locked laser is obtained from a multimode laser by introducing a time-varying internal perturbation into the laser's cavity. In this case however the perturbation is driven at a frequency which is slightly different from the axial mode interval frequency.

The basic configuration of the FM laser is the same as that of a phase-locked laser.

The time-varying perturbation locks the phases of the modes to a common reference. Harris [1966] shows that, if the perturbation is sufficiently strong, the laser modes oscillate with FM phases and Bessel function amplitudes thereby comprising the sidebands of a frequency modulated signal. The modulation depth δ is determined by the strength of the phase perturbation and by the difference between the driving frequency of the phase perturbation and the axial mode interval.

This type of laser operation was first reported by Harris and Targ [1964] who used a HeNe laser. The first theoretical explanation was given by Harris and McDuff [1964]. A more complete discussion of the FM laser is given by Harris [1966].

Kerr [1966] has pointed out that the output of a FM laser can be used as the optical carrier in a wideband, direct-detection optical communications system. He notes that almost as much power is available from an FM laser as from a free-running multimode laser without the bandwidth limitation of the free-running laser. Furthermore although the laser has an inherent periodic frequency variation, this variation does not limit the laser in a direct detection system. Also the noise associated with mode pulling (mode-interference noise) is eliminated since the modes are all coherent.

A heterodyne detection system however requires a single frequency source. This has been a major limitation on the range of a heterodyne detection communications system. The commonly used source in communications experiments with heterodyne detection is the Spectra-Physics Model 119 which has a power output of only 100 microwatts.

5.8 Methods for Generation of Single Frequency Light

Two techniques are available to produce single frequency output from a high powered multimode laser without the inherent loss of power associated with mode suppression techniques. These are the "supermode" technique and the frequency selective coupling technique.

Massey et al. [1965] first reported the generation of single frequency light using the supermode technique. The basic diagram of a "supermode" laser is shown in Figure 18.

The light output of an FM laser is passed through an external modulator which is driven 180° out of phase and with the same optical phase deviation as the output of the FM laser. The output of the external modulator is then in principle a monochromatic signal.

The output of an FM laser may be approximated by the following equation:

$$E = E_o \cos(\omega_c t + \delta \sin \omega_m t) \quad ,$$

where E_o is the peak amplitude of the optical field, ω_c is the

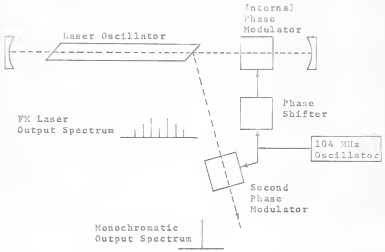

Figure 18. Block diagram of "super-mode" laser, showing FM laser with external phase modulator [Massey et al. 1965].

center frequency of the optical output spectrum, ω_m is the modulation frequency, and δ is the phase deviation of the FM oscillation [Massey et al., 1965]. If the FM signal is passed through an external modulator driven at the same frequency as the internal modulator, the output signal is of the form:

$$E' = E_o \cos[\omega_c t + \delta \sin \omega_m t + \delta' \sin(\omega_m t+\theta)],$$

where δ' is the peak phase deviation and θ is the phase of the external modulator relative to the internal modulator [Massey et al., 1965]. Thus if δ' is equal to δ and θ equals $(2n+1)\pi$ radians, where n is an integer, the output E' is monochromatic.

The arrangement in Figure 18 is for a HeNe laser whose mode separation is 104 MHz. The reflection from the Brewster angle window was used to obtain the FM signal but the signal may be obtained equally as well through a partially transmitting mirror.

This process is theoretically 100% efficient, but there is some distortion in the FM signal and loss in the modulators. Harris [1966] has calculated that efficiencies of about 98% are possible provided the modulators are lossless. Thus high efficiencies are dependent on the development of low loss modulators.

Osterink and Targ [1967] have reported 350 mw of single frequency light at 5145 angstroms using an argon FM laser and the supermode technique. They obtained a conversion efficiency from multimode power to single mode power of 45%.

The second technique for obtaining single frequency light from a multimode laser is called frequency selective coupling. A time-varying internal perturbation is used along with a Fabry-Perot output coupling etalon which passes only the desired mode. Harris and McMurtry [1965] have successfully demonstrated this technique; however there is considerable practical difficulty associated with the technique.

CHAPTER VI

SYSTEMS

Theoretically the modulation of an optical carrier can take the same form as that of an RF carrier. However because of quantum effects, atmospheric turbulence, receiver circuit complexities and component problems, the conclusions drawn for the RF case are not directly applicable to optical communications systems.

The systems may be divided into two classes. The first are direct detection systems which use intensity modulation. The second class is the heterodyne system which may use all forms of amplitude and frequency modulation.

6.1 Direct Detection Systems

The direct detection system is the simplest system for optical communications. Although only intensity modulation may be used, it was shown in this report that this is not a significant disadvantage since subcarriers with other forms of modulation may be used.

All the characteristics of direct detection systems have been covered in the previous chapters except the use of pulse modulation.

6.1.1 Pulse Systems

The information capacity of continuous wave modulation systems is limited by the bandwidths of the present optical modulators. Presently the bandwidth is limited to several gigahertz. However if laser communications systems are to be technically useful, higher information capacity will be necessary. The only method developed thus far that has a very large information capacity is a digital optical communications system using a pulse code modulation (PCM) format. The PCM format consists of a series of time slots occuring at a fixed frequency, with the information conveyed by the presence or absense of a pulse in each time slot. With this approach the information capacity is not set by the modulator's bandwidth but by the width of the pulses from a phase-locked laser. Denton and Kinsel [1968] show that with existing devices the information rate may reach 10^{10} bits/sec by time-multiplexing many PCM channels onto a single monochromatic laser beam.

A single PCM channel will be described and the results extrapolated to the multiple channel case.

In Chapter V it was shown that a phase-locked laser emits a train of narrow pulses whose repetition frequency is c/2L. Denton et al. [1966] show that if a phase-locked laser is used with an electrically controlled optical gate, a PCM modulated output may be obtained. The basic system is shown in Figure 19.

The optical source used by Denton et al. [1966] was a HeNe laser operating at 6328 angstroms which was phase-locked by a KDP phase modulator operating at 224 MHz. The output consists of a train of pulses which are 0.6 ns wide and separated by 4.46 ns. The beam splitter passes vertically polarized pulses from the laser that reflects the horizontally polarized pulses from the modulator by 90°. The modulation of the laser pulses is accomplished by the electrooptic effect.

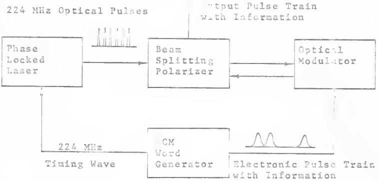

Figure 19. Schematic diagram of an optical PCM system [Denton et al., 1966].

Note that since the optical pulses are narrow and separated by a wide interval, the electrical pulses do not need to be nearly as narrow as the optical pulses.

Denton and Kinsel [1968] have built and tested a PCM optical communications system using the configuration of Figure 19 along with a germanium avalanche photodiode as a detector. They evaluated the system using live program material consisting of a broadcast-quality color video signal, a reduced bandwidth video signal and 36 telephone channels. This information was provided by an electrical PCM transmitter terminal at a 224 Mbit/sec rate to drive the PCM optical modulator. The results indicated that the error rate of the entire system was not affected by the presence of the optical link.

A 224 Mbit/sec information rate would not justify a laser communications system since the same amount of information may be transmitted electrically by a coaxial cable [Mayo, 1968]. However the width of the optical pulses is so small compared to the pulse spacing that there exists the possibility of time multiplexing many PCM optical channels onto a single laser beam.

Kinsel and Denton [1968] show that, by using a phase-locked solid-state laser such as Nd:YAlG which has a pulse width from 30 to 100 ps. wide, 24 PCM optical channels may be time-multiplexed together with a resultant information rate of 5.376 Gbit/sec. Furthermore if horizontal polarization is used in addition to vertical polarization a information rate of 10^{10} bit/sec may be obtained.

Kinsel and Denton [1968] have built a two channel system and found that it performs satisfactorily.

Ross [1966] shows that pulse modulation systems may offer a number of advantages. These can be efficiency in operation, noncritical modulation circuits, better discrimination against the bandground, noncritical receiver circuitry in that distortion-free detection and amplification may not be required. Pulse modulation systems are more efficient if the transmitter laser can be internally modulated (i.e. laser turned on and off). External pulse modulation wastes energy since the laser is always on. Also recall that the pulse intensity of a phase-locked laser is approximately 6 times the free-running intensity. Thus a phase-locked laser is effectively 6 times as powerful as the same laser which is not phase-locked.

The modulation circuits are less critical since the linearity requirements are not as strict as for amplitude modulated CW systems. Distortion-free receivers are not required since the receiver must only determine the presence or absence of a pulse. Furthermore pulse modulation is less susceptible to atmospheric disturbances for the same reason.

Other forms of pulse modulation besides PCM are pulse position modulation, pulse width modulation, and pulse frequency modulation. These are all possible in an optical communications system but are inferior to PCM systems [Ross, 1966].

6.2 Heterodyne Detection Systems

A basic microwave bandwidth optical heterodyne communications system is shown in Figure 5. In principle AM, FM, SSB, or pulse modulation may be used. However SSB and FM modulation are not practical at present. Wideband SSB modulation is difficult to achieve and results in serious attenuation the signal [Kerr, 1967]. Also the demodulation of SSB signal requires the exact reinsertion of the carrier which imposes severe tolerances on the stability of the transmitter and LO lasers and on the performance of the AFC system.

[Kerr, 1967] states that the required modulation depth and discriminator bandwidth for FM modulation is difficult to achieve for the information rates of interest.

The use of intensity or pulse modulated light with an optical heterodyne communications system will greatly reduce the requirements on the AFC system and the frequency stability of the laser. This reduction is very important since, at present, frequency stability requirements for FM and SSB modulation are very difficult to achieve in the laboratory and next to impossible in the field.

Lucy et al. [1967] have built and tested a heterodyne receiver that may be used in the field. They list data on the signal to noise ratios obtained under various conditions.

CHAPTER VII

CONCLUSIONS

The heterodyne detection receiver's S/N is ideally 3 db better than the direct detection receiver; however ideal operation of the heterodyne receiver has not been approached. Conversely quasi-ideal operation of a direct detection receiver may be obtained in the visible and near infrared. The fact that direct detection receivers respond only to intensity modulation is not a major disadvantage. The operation of a direct detection receiver in a field environment is easily achieved; however operation of a heterodyne receiver in the field is difficult at present. Heterodyne detection receivers will probably find their greatest application in the middle infrared where mechanical tolerances are eased somewhat.

The best detectors currently available in the visible spectrum are photoemission devices with a current gain mechanism. Signal-shot-noise-limited response for a bandwidth of 1 GHz has been obtained. The static crossed-field photomultiplier is best suited for baseband detection while the multiplier traveling-wave phototube is best for bandpass detection.

In the near infrared solid state detectors are available; however their sensitivity is at least two orders of magnitude away from the ideal case. The avalanche photodiode with [illegible] hertz bandwidth is the best detector available in the near infrared.

Detectors for the middle infrared are not well developed, are narrow band, and have poor sensitivity. Photoconductive and photovoltaic devices are currently being used.

The signal degradation caused by the atmosphere is many times severe enough to completely destroy the signal. Present indications are that reliable communications through the atmosphere without the use of very closely spaced receivers are not probable. Low loss guided wave propagation is possible but the mechanical tolerances seem to be severe. Further research is necessary to determine the tolerances necessary and the cost of constructing a guide. The most flexible guides consist of spaced lenses mounted in a pipe.

Gigahertz bandwidth modulators for use in the visible and near infrared are available. They usually have a traveling-wave structure to achieve wideband operation. The major limitation of traveling-wave modulators is the lack of adequate broadband RF matching sections necessary to couple the modulating signal into the modulator. Losses are also a major problem which may be overcome by the development of better materials. There are no adequate modulators presently available for use in the middle infrared.

Crystalline solid state lasers are most suitable for wideband communications when they can be phase-locked and thus be used in a PCM system. Continuous operation of the laser at relatively high powers has been achieved but the output is too unstable to be useful at this time.

Glass solid state lasers may only be pulsed at a slow rate and hence are not suited for use in high information rate communication systems.

Semiconductor lasers are presently of limited usefulness in wideband communication systems. They require cyrogenic cooling and have a multimode output, poor stability, and poor directivity. However they are small, compact, and easily modulated at high frequencies. Further development is necessary before it can be considered adequate for wideband communication.

Gas lasers are currently the most popular lasers for communications research. The helium-neon laser has been used extensively in communications research; however it has relatively low power output. Therefore the more powerful ion laser's will probably be used in more advanced communication systems. The CO_2 laser, which is a molecular gas laser, has a very powerful output in the middle-infrared; however at this wavelength efficient modulators and detectors are not available.

The major problem with lasers is that, to attain the high power outputs necessary for wideband laser communication systems, the lasers must be operated in a multi-frequency mode. This type of operation limits bandwidth and precludes the use of heterodyne detection receivers. The recently developed FM laser removes the bandwidth limitation and is suitable for use with a direct detection receiver. The FM output may be converted to single frequency output by using the supermode technique or frequency selective coupling. Single frequency oscillation is necessary for heterodyne detection.

Kerr (1967) has calculated the general influence of information capacity requirements, range, background level, optical wavelength, and detection system on the required transmitter power for seven systems. Six of the systems are presently realizable while the seventh, a heterodyne system operating at 10.6 μm, uses components which should become available in the future. The results of Kerr's calculation show that a heterodyne system at 10.6 μm requires almost an order of magnitude less power than the next best system. However, at present, adequate modulators and demodulators are not available at 10.6 μm. The best system that is realizable at present is a medium-aperature direct-detection system operating with an S-20 MTWP at the argon laser wavelength (4880 angstroms).

REFERENCES

Aircraft Armaments, Inc. FM [illegible] a CW Laser Beam. National Aeronautics and Space Administration Report CR-668. Washington, D.C., Feb. 1967.

Anderson, L. K. and B. J. McMurtry. "High Speed Photodetectors". Proc. IEEE, Oct. 1966, 54(10):1335-1349.

Beese, N. C. "Light Sources for Optical Communication". Infrared Physics, Mar. 1961, 1(1):5-16.

Bicknell, W. E., B. K. Yap, and C. J. Peters. "0 to 3Gc Traveling-Wave Electrooptic Modulator". Proc. IEEE (Letters), Feb. 1967, 55(2):225-226.

Bloom, A. L. Noise in Lasers and Laser Detectors. Spectra-Physics Laser Technical Bulletin Number 4. Mountain View, Calif.: Spectra-Physics, Inc., 1965.

Bloom, A. L. "Gas Lasers". Proc. IEEE, Oct. 1966, 54(10):1262-1275.

Brand, L., B. Hill, E. Holtz, and G. Wencker. "External Light Modulation with Low Microwave Power". Electron. Letters, Aug. 1966, 2(8):317-318.

Brookner, E., M. Kolker, and R. M. Wilmotte, "Deep-Space Optical Communications." IEEE Spectrum, Jan. 1967, 4(1):75-82.

Chase, D. M. "Power Loss in Propagation through a Turbulent Medium for an Optical Heterodyne System with Angle Tracking". J. Opt. Soc. Am., Jan. 1966, 56(1):33-40.

Chernov, L. A. Wave Propagation in a Random Medium. New York: McGraw Hill, 1960.

Denton, R. T. and T. S. Kinsel. "Terminals for a High-Speed Optical Pulse Code Modulation Communications System: I. 224-M bit/s Single Channel". Proc. IEEE, Feb. 1968, 56(2):140-145.

Denton, R. T., T. S. Kinsel, and F. F. Chen. "224-Mc/s Optical Pulse Code Modulator". Proc. IEEE (Letters), Oct. 1966, 54(10): 1472-1473.

Fried, David L. "Optical Heterodyne Detection of an Atmospherically Distorted Signal Wave Front" Proc. IEEE, Jan. 1967, 55(1):57-67.

Fried, D. L. and J. B. [illegible]idman. "Heterodyne and Photon-Counting Receivers for Optical Communications" Appl. Opt., Feb. 1967, [illegible]):245-[illegible].

Gloge, D. "Experiment with an Underground Lens Waveguide". Bell System Tech. J., Apr. 1967, 46(4):721-735.

Gordon, J. R. "Optics of General Guiding Media". Bell System Tech. J., Feb. 1966, 45(2):321-331.

Goubau, Georg. "Lenses Guide Optical Frequencies to Low-Loss Transmission". Electronics, May 16, 1966, 39(10):83-89.

Harris, S. E. "Stabilization and Modulation of Laser Oscillators by Time-Varying Perturbation". Proc. IEEE, Oct. 1966, 54(10): 1401-1413.

Harris, S. E. and O. P. McDuff. "FM Laser Oscillation Theory". Appl. Phys. Letters, Nov. 15, 1964, 5(10):205-206.

Harris, S. E. and B. J. McMurty. "Frequency Selective Coupling to the FM Laser". Appl. Phys. Letters, Nov. 1, 1965, 7(10): 265-266.

Harris, S. E. and Russell Targ. "FM Oscillation of the He-Ne Laser". Appl. Phys. Letters, Nov. 15, 1964, 5(10):202-204.

Hobart, James L. Retardation-Type, Laser Modulators. Spectra-Physics Technical Memorandum. Mountain View, Calif.: Spectra-Physics, Inc., 1966.

Hodara, H. "Laser Wave Propagation Through the Atmosphere". Proc. IEEE, Mar. 1966, 54(3):368-375.

Kaminow, I. P. and E. H. Turner. "Electrooptic Light Modulators". Proc. IEEE, Oct. 1966, 54(10):1374-1390.

Kerr, J. Richard. "The FM Laser and Optical Communication System". Appl. Opt., Apr. 1966, 5(4):671-672.

Kinsel, T. S. and R. T. Denton. "Terminals for a High Speed Optical Pulse Code Modulation Communication System: II. Optical Multiplexing and Demultiplexing". Proc. IEEE, Feb. 1968, 56(2): 146-154.

Kiss, Z. J. and R. J. Pressley. "Crystalline Solid Lasers". Proc. IEEE, Oct. 1966, 54(10):1236-1248.

Kogelnik, H. and T. Li. "Laser Beams and Resonators". Proc. IEEE, Oct. 1966, 54(10):1312-1328.

Lang, K. and R. Lucy. "Continuous Wave Laser Ground-to-Space-to-Ground Laser Experiment". Appl. Opt., Sept. 1967, 6(9):1579.

Lasser, Marvin E. "Detection of Coherent Optical Radiation". IEEE Spectrum, Apr. 1966, 3(4):73-78.

Lipsett, M. S. Laser/Optics Technique Summary Report. National Aeronautics and Space Administration Report CR-77482. Washington, D.C., Apr. 29, 1966.

Lucy, R. F., K. Lang, C. J. Peters, and K. Duval. "Optical Superheterodyne Receiver". Appl. Opt., Aug. 1967, 6(7):1333-1342.

Mandel, L. "Heterodyne Detection of a Weak Light Beam". J. Opt. Soc. Am., Sept. 1966, 56(9):1200-1206.

Massey, G. A., M. K. Oshman, and Russell Targ. "Generation of Single-Frequency Light Using the FM Laser". Appl. Phys. Letters, Jan. 1, 1965, 6(1):10-11.

Mayo, J. S. "Pulse-Code Modulation". Sci. Am., Mar. 1968, 218(3): 103-108.

Meyer-Arendt, J. R. and C. B. Emmanuel. Optical Scintillation: A Survey of the Literature. U.S. Department of Commerce, National Bureau of Standards Technical Note 225. Boulder, Colo., April 5, 1965.

Miller, Stewart E. "Communication by Laser". Sci. Am., Jan. 1966, 214(1):19-27.

Miller, S. E. and L. C. Tillotson. "Optical Transmission Research". Proc. IEEE, Oct. 1966, 54(10):1300-1311.

Moller, F. "Optics of the Lower Atmosphere". Appl. Opt., Feb. 1964, 3(2):157-166.

Nathan, Marshall I. "Semiconductor Lasers". Proc. IEEE, Oct. 1966, 54(10):1276-1289.

Nussmeier, T. A. "Laser Communication". IEEE Student Journal, Sept. 1967, 5(5):18-22.

Ohm, E. A. "A Linear Optical Modulator with High FM Sensitivity". Appl. Opt., July 1967, 6(7):1233-1235.

Oliver, B. M. "Signal-to-noise Ratios in Photoelectric Mixing". Proc. IRE (Correspondence), Dec. 1961, 49(12):1960-1961.

Osterink, L. M. and Russell Targ. "Single-Frequency Light from an Argon FM Laser." Appl. Phys. Letters, Feb. 15, 1967, 10(4): 115-117.

Pask, E. C. and L. S. Stokes. "Lasers vs. Microwaves for Deep-Space Communications". Microwaves, May 1967, 6(5):78-9[illegible].

Penfield, P. and D. E. Sawyer. "Photoparametric Amplifier". Proc. IEEE, Apr. 1965, 53(4):340-347.

Peters, C. J. "Gigacycle Bandwidth Coherent Light Traveling-Wave Phase Modulator". Proc. IEEE, Jan. 1963, 51(1):147-152.

Peters, D. W. "Infrared Modulator Utilizing Field-Induced Free Carrier Absorption." Appl. Opt., June 1967, 6(6):1033-1040.

Rempel, Robert C. Optical Properties of Lasers as Compared to Conventional Radiators. Spectra-Physics Laser Technical Bulletin Number 1. Mountain View, Calif.: Spectra-Physics, Inc., 1963.

Ross, Monte. Laser Receivers. New York: John Wiley & Sons, 1966.

Schiel, E. J., E. C. Bullwinkel, and R. B. Weimer. "Pulse-code Modulation Multiplex Transmission over an Injection Laser Transmission System". Proc. IEEE (Correspondence), Dec. 1965, 53(12): 2140-2141.

Siegmen, A. E. "Antenna Properties of Optical Heterodyne Receivers". Appl. Opt., Oct. 1966, 5(10):1588-1594.

Snitzer, E. "Glass Lasers". Proc. IEEE, Oct. 1966, 54(10):1249-1261.

Sommers, H. S. Jr. and E. K. Gatchell. "Demodulation of Low-Level Broad-Band Optical Signals with Semiconductors". Proc. IEEE, Nov. 1966, 54(11):1553-1568.

Spencer, E. G., P. V. Lenzo, and A. A. Ballman. "Dielectric Materials for Electrooptic, Elastooptic, and Ultrasonic Device Applications". Proc. IEEE, Dec. 1967, 55(12):2074-2108.

Stone, Julian, George Lynch, and Richard Pontinen. "A dc-Biased Kerr Cell Light Modulator". Appl. Opt., Apr. 1966, 5(4):653-657.

Targ, Russell and W. D. Bush. "Automatic Frequency Control of a Laser Local Oscillator for Heterodyne Detection of Microwave Modulated Light". Appl. Opt., Dec. 1965, 4(12):1523-1527.

Tatarski, V. I. Wave Propagation in a Turbulent Medium. New York: McGraw-Hill, 1961.

Teich, M. C. "Infrared Heterodyne Detection". Proc. IEEE, Jan. 1968, 56(1):37-46.

Tomiyasu, K. "Laser Bibliography V". IEEE J. Quantum Electron., July 1967, QE-3(7):296-32[illegible].

Tomiyasu, K. "Bibliography of Laser Devices". IEEE J. Quantum Electron., Dec. 1967, QE-3(12):67[illegible]-[illegible]8[illegible].

Trevelyn, B. and H. Purdey. "An Optical Communication Link for Laboratory Demonstration using Microwave Modulation of the Light Beam". Radio Electron. Engr., June 1967, 33(6):384-386.

Warren, Grover. "Laser Practice and Applications". Electronics World, Sept. 1965, 74(3):45-[illegible]0.

White, Alan D. "Gas-Laser Frequency Stabilization". Microwaves, Jan. 1967, 6([illegible]):51-[illegible]

Whitehouse, David R. "Understanding CO_2 Lasers". Microwaves, July 1967, 6(7):A6-A14.

Yariv, A. "Internal Modulation in Multimode Laser Oscillators". J. of Appl. Phys., Feb. 1965, 36(2):388-391.

VITA

Patrick Harold Riedl

Candidate for the Degree of

Master of Science

Report: WIDEBAND OPTICAL COMMUNICATION SYSTEMS

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Great Bend, Kansas, August 17, 1944, the son of Harold F. and Thersa A. Riedl.

Education: Attended grade school in Otis, Kansas; graduated from Otis Rural High School in 1962; received the Bachelor of Science degree from Kansas State University, with a major in Electrical Engineering, in June, 1966; completed requirements for Master of Science degree in August, 1968.

Honorary Societies: Initiated into Phi Kapp Phi in 1966.

Licenses: Author holds a First Class Radio Telephone Operator License.

WIDEBAND OPTICAL COMMUNICATION SYSTEMS

by

PATRICK HAROLD RIEDL

B. S., Kansas State University, 1966

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Electrical Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1968

The purpose of this report is to determine the present status of high information rate terrestrial laser communication systems. Due to its extreme directivity and extremely high frequency the laser has the capability to transmit enormous amounts of information.

Laser receivers, transmission mediums, laser modulators, and laser transmitters are each treated in detail. It is shown that direct detection receivers are generally superior to heterodyne detection receivers in the visible and near infrared spectrum. However, in the middle infrared the heterodyne receiver is more promising. Gigahertz bandwidths are presently available using photoemissive detectors with an internal gain mechanism.

Atmospheric disturbances can severely degrade laser signals. Guided propagation eliminates the problems due to the atmosphere but introduces the problems of high cost and sometimes high attenuation.

Laser modulators for wideband communications are traveling-wave structures which utilize the electooptic effect. Relatively efficient microwave bandwidth amplitude modulators are available.

The gas laser is found to be the most suitable laser for communication systems. A major disadvantage of the high powered gas laser is that the output consists of many frequencies. However, methods are described by which the multiple frequency output may be converted to a single frequency output or to a train of very narrow pulses.

The train of pulses may be pulse code modulated with a resultant information rate on the order of 10^2 Mbit/sec. Information rates of 10^{10} Mbit/sec are possible by time-multiplexing the signal.