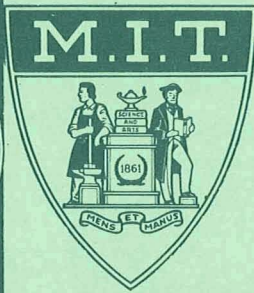


NASA  
SP  
217  
c.1



# Optical Space Communication



LOAN COPY: RETURN TO  
AFWL (WLOL)  
KIRTLAND AFB, N MEX

Proceedings of an MIT-NASA  
Workshop held at Williams College,  
Williamstown, Mass., August 4-17, 1968



NASA SP-217  
TECH LIBRARY KAFB, NM



0063331

# OPTICAL SPACE COMMUNICATION

Proceedings of a workshop sponsored jointly  
by the NASA Electronics Research Center  
and the Research Laboratory of Electronics,  
Massachusetts Institute of Technology, and  
held August 4-17, 1968, at Williams Col-  
lege, Williamstown, Massachusetts

Edited by

ROBERT S. KENNEDY AND  
SHERMAN KARP

*Massachusetts Institute of Technology*



*Scientific and Technical Information Division*  
OFFICE OF TECHNOLOGY UTILIZATION  
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION  
Washington, D.C.

1969

---

For Sale by the Superintendent of Documents,  
U.S. Government Printing Office, Washington, D.C. 20402  
Price \$1.50 (paper cover)

## Preface

The advent of coherent sources at optical frequencies opened new and important possibilities for increased capacity and efficiency in communication systems. These increases have yet to be realized. The delay is due in part to the time required to develop an adequate technological base and in part to the emergence of fundamental problems that do not arise at lower frequencies. To assess some of these problems and to delineate a path toward their solutions, NASA/ERC and MIT/RLE invited some of the leaders in the field of optical communication to a two-week Workshop that was held at Williams College.

Broadly stated, the Workshop was concerned with the knowledge that must be acquired to determine the structure and performance of near-optimum optical systems. The feasibility of implementing these systems with existing devices and the characteristics of necessary, but unavailable, devices were also considered. We hope that the Workshop, and this report of its deliberations, will contribute to the further development of optical communication systems.

The success of the Workshop was ensured by the expertise and energy that all of the participants brought to it, but it was achieved more easily because of the guidance provided by the Steering Committee. Paul L. Smith's efficiency as business manager and general administrator and the efforts of Charles Janke, Mary Owens, Marilyn Pierce, and Bunny Frank also contributed greatly to the smooth operation and enjoyment of the Workshop.

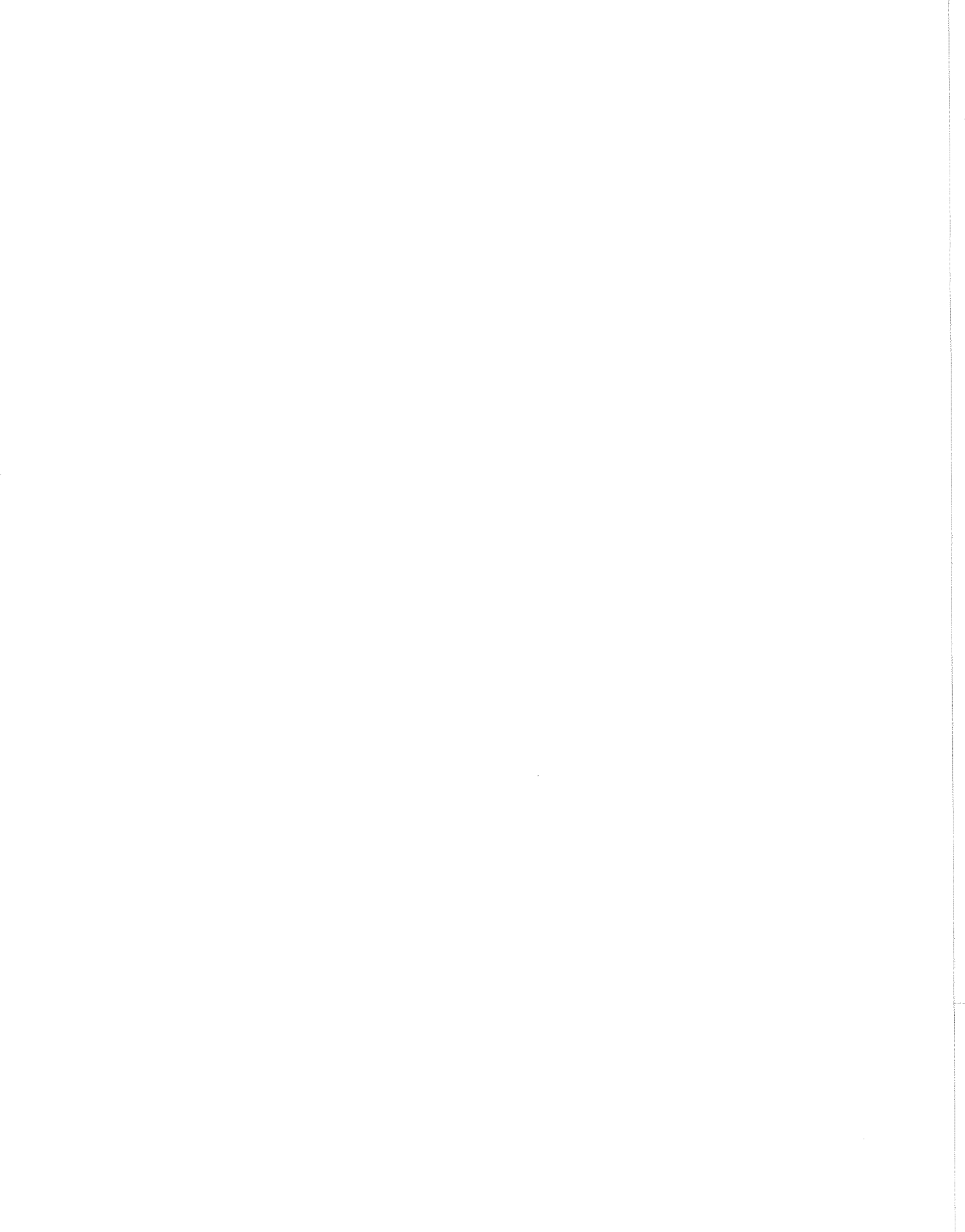
We particularly thank Joseph Hull, whose early commitment to our suggestion for the Workshop encouraged us to pursue it, and Estil V. Hoversten, whose contribution to its preparation, operation, and summation was a principal factor in its success.

Sherman Karp  
NASA/Electronics Research Center

Robert S. Kennedy  
M. I. T./Research Laboratory of Electronics

### Steering Committee

Henry L. Anderton  
Estil V. Hoversten  
Joseph A. Hull  
Sherman Karp  
Robert S. Kennedy  
Henry Plotkin  
Joseph L. Randall  
Lawrence G. Roberts  
Michael S. Shumate  
Louis D. Smullin  
Lester C. Van Atta  
Henry J. Zimmermann



# Contents

|                                             |     |
|---------------------------------------------|-----|
| PARTICIPANTS IN THE WORKSHOP                | vii |
| I. INTRODUCTION                             | 1   |
| II. SUMMARY AND CONCLUSIONS                 | 3   |
| 2.1 General Conclusions                     | 3   |
| 2.1.1 Large Collecting Optics               | 3   |
| 2.1.2 Device Needs                          | 3   |
| 2.1.3 System Analysis                       | 5   |
| 2.2 Atmospheric Channel                     | 6   |
| 2.3 Scattering Channels                     | 9   |
| III. FREE-SPACE SYSTEMS                     | 11  |
| 3.1 Existing Receiver Components            | 12  |
| 3.1.1 Energy Detectors                      | 12  |
| 3.1.2 Preamplifiers                         | 14  |
| 3.2 Existing Receiver Structures            | 14  |
| 3.2.1 Heterodyning                          | 15  |
| 3.2.2 Homodyning                            | 15  |
| 3.2.3 Direct Detection                      | 15  |
| 3.2.4 System Performance                    | 16  |
| 3.3 Large-Aperture Receivers                | 17  |
| 3.4 Quantum Communication Theory            | 18  |
| IV. ATMOSPHERIC CHANNEL                     | 21  |
| 4.1 Turbulent Effects                       | 22  |
| 4.1.1 Theoretical Approaches                | 22  |
| 4.1.2 Intensity Scintillation               | 24  |
| 4.1.3 Phase Fluctuations                    | 34  |
| 4.1.4 Beam Time Behavior                    | 41  |
| 4.1.5 Areas Requiring Further Investigation | 42  |
| 4.2 Channel Models and Representations      | 46  |
| 4.3 Communication Systems                   | 49  |
| 4.3.1 Direct-Detection Techniques           | 50  |
| 4.3.2 Heterodyne-Detection Techniques       | 58  |
| 4.3.3 New Avenues of System Development     | 65  |

|            |                                                                                                       |     |
|------------|-------------------------------------------------------------------------------------------------------|-----|
| V.         | SCATTERING CHANNELS                                                                                   | 67  |
| 5.1        | Communication Considerations                                                                          | 67  |
| 5.2        | Channel Characterization                                                                              | 67  |
| 5.2.1      | Scattering Theory                                                                                     | 71  |
| 5.2.2      | Physical Characteristics of Clouds, Fog, and Haze                                                     | 72  |
| 5.2.3      | Complex Refractive Index for Water and Ice<br>Particles                                               | 73  |
| 5.3        | Proposed Experiments                                                                                  | 73  |
| 5.3.1      | Laboratory Experiments on Artificial Clouds<br>(Ice and Water)                                        | 74  |
| 5.3.2      | Natural Cloud Experiments                                                                             | 74  |
|            | REFERENCES                                                                                            | 76  |
| APPENDIX A | Abstracts of Briefing Sessions                                                                        | 83  |
| APPENDIX B | Designing "Real" Optical Communication Systems                                                        | 95  |
| APPENDIX C | On the Reliability of Quantum-Mechanical Com-<br>munication Systems                                   | 98  |
| APPENDIX D | Some Quantum-Mechanical Aspects of Optical<br>Communications                                          | 107 |
| APPENDIX E | Estimation of Modulation Frequency of a Light Beam                                                    | 117 |
| APPENDIX F | Comments on Quantal Theory of Communication                                                           | 120 |
| APPENDIX G | Equivalence of Transmitter and Receiver Antenna<br>Gain in a Random Medium                            | 129 |
| APPENDIX H | Atmospheric Turbulence Considerations for a Ground-<br>Based Laser Transmitter in a Ground-Space Link | 133 |
| APPENDIX I | Spatially Distributed Coherent Heterodyne Reception<br>through a Fluctuating Atmosphere               | 137 |

## Workshop Participants

Sherman Karp, Co-Chairman  
Electronics Research Center  
National Aeronautics and Space Administration  
Cambridge, Massachusetts

Robert S. Kennedy, Co-Chairman  
Department of Electrical Engineering and  
Research Laboratory of Electronics  
Massachusetts Institute of Technology  
Cambridge, Massachusetts

Darrell E. Burch, Group Coordinator  
Philco Ford Corporation  
Newport Beach, California

Estil V. Hoversten, Group Coordinator  
Department of Electrical Engineering and  
Research Laboratory of Electronics  
Massachusetts Institute of Technology  
Cambridge, Massachusetts

Joseph A. Hull, Group Coordinator  
Electronics Research Center  
National Aeronautics and Space Administration  
Cambridge, Massachusetts

Henry Plotkin, Group Coordinator  
Goddard Space Flight Center  
National Aeronautics and Space Administration  
Greenbelt, Maryland

Henry L. Anderton  
Office of Advanced Research and Technology  
National Aeronautics and Space Administration  
Washington, D.C.

William B. Bridges  
Hughes Research Laboratories  
Malibu, California

Anthony Devaney  
Electronics Research Center  
National Aeronautics and Space Administration  
Cambridge, Massachusetts

Edward C. Fraser  
Information and Control Laboratory  
Stanford Research Institute  
Menlo Park, California

David L. Fried  
Autonetics  
Anaheim, California

Robert M. Gagliardi  
Department of Electrical Engineering  
University of Southern California  
Los Angeles, California

Joseph E. Geusic  
Bell Telephone Laboratories Incorporated  
Murray Hill, New Jersey

Frank E. Goodwin  
Hughes Research Laboratories  
Malibu, California

James P. Gordon  
Bell Telephone Laboratories  
Holmdel, New Jersey

Paul E. Green  
Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington, Massachusetts

Robert O. Harger  
Radar and Optics Laboratory  
University of Michigan  
Ann Arbor, Michigan

Carl W. Helstrom  
Department of Applied Electrophysics  
University of California, San Diego  
La Jolla, California

Alfred C. Holland  
Electronics Research Center  
National Aeronautics and Space Administration  
Cambridge, Massachusetts

J. Richard Kerr  
Oregon Graduate Center for Study and Research  
Portland, Oregon

Robert H. Kingston  
Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington, Massachusetts

Robert S. Lawrence  
Optical Propagation Program  
U.S. Department of Commerce  
Environmental Scientific Services  
Administration Research Laboratories  
Boulder, Colorado



Jane W.S. Liu  
Research Laboratory of Electronics  
Massachusetts Institute of Technology  
Cambridge, Massachusetts

Nelson McAvoy  
Goddard Space Flight Center  
National Aeronautics and Space Administration  
Greenbelt, Maryland

Edward O'Neill  
Physics Department  
Worcester Polytechnic Institute  
Worcester, Massachusetts

Athanasios Papoulis  
Department of Electrical Engineering  
Polytechnic Institute of Brooklyn  
Brooklyn, New York

Joseph L. Randall  
George C. Marshall Space Flight Center  
National Aeronautics and Space Administration  
Huntsville, Alabama

Monte Ross  
McDonnell-Douglas Astronautics  
St. Louis, Missouri

Edward Sarachik  
Electronics Research Center  
National Aeronautics and Space Administration  
Cambridge, Massachusetts

Michael S. Shumate  
Jet Propulsion Laboratory  
Pasadena, California

Lester C. Van Atta  
Electronics Research Center  
National Aeronautics and Space Administration  
Cambridge, Massachusetts

Ralph V. Wood, Jr.  
Lincoln Laboratory  
Massachusetts Institute of Technology  
Lexington, Massachusetts

# I. Introduction

Optical communication systems for space applications are potentially of great importance to the National Aeronautics and Space Administration. The high data rate requirements for space-to-space, space-to-ground, and computer-to-computer communication, as well as the requirements for deep-space applications, constitute a strong justification for the pursuit of a research program in optical communication.

Many system studies for deep-space applications have been performed by contractors for the National Aeronautics and Space Administration. The coordination and effective utilization of this effort has been complicated by the multidisciplinary nature of the work. For example, the technical language barrier between the optical communication theorist and the quantum physicist, or between the solid-state device researcher and the physicist working on the statistical properties of the atmosphere, is a significant problem in the advancement of the program.

In an attempt to improve this coordination, the Massachusetts Institute of Technology and NASA jointly organized a Workshop, which was held at Williams College, Williamstown, Massachusetts, from August 4 to 17, 1968. A selected group of participants from the areas of the physical sciences having a direct bearing on optical communication were invited to participate under the guidance of communication engineers. Group interactions were emphasized rather than individual work. The participants were selected from industry, universities, and NASA centers. The object of the Workshop was to identify some of the problems that represent the most significant obstacles in the development of optical communications systems to their full potential.

To provide a common basis for work, the first two and one-half days were devoted to the presentation of status papers on various aspects of optical communication by members of the group. Abstracts of these papers are presented in Appendix A. After these presentations, the participants formed working groups pertaining to three classes of channels: atmospheric, cloud, and free-space. In each of these groups, the objective was to assess the present level of understanding and to identify problem areas in which research should be concentrated to enhance this understanding. This report is intended to present the conclusions of the working groups.

The emphasis of the Workshop was directed toward digital, rather than analog, signaling schemes, partly because digital systems provide greater flexibility in signal selection and partly because they are more easily understood. Many of the conclusions are applicable to both digital and analog systems. The approach to the problems reflects the perspective of contemporary communication engineering. The following questions are typical of those that were posed.

1. What is and must be known about the channel in order to determine the optimum receiver and transmitter? How can the missing information be acquired?
2. Given the channel description, is the (mathematically) optimum receiver already known? If not, how can it be determined?

3. What new devices are needed if near-optimum systems are to be realized? What performance can be realized with existing devices? What signal sets should be used?

4. Given the (mathematically) optimum receiver, can it be implemented easily? If not, what are some possible avenues along which "good" practical systems might be developed?

The principal conclusions and recommendations of the Workshop are summarized in Section II. In Section III system problems associated with free-space channels are discussed. Many of the problems encountered in the free-space system are common also to the atmospheric and cloud channels. Section IV represents the largest group effort of the Workshop; the atmospheric channel considered there poses many challenging problems, as indicated by the suggested work areas. In Section V the problem of communicating through clouds is considered.

## II. Summary and Conclusions

Several general conclusions, which are applicable to all of the channels discussed, emerged from the workshop. We present them first and then consider the atmospheric and cloud channels. The free-space channel is not discussed separately, since the conclusions reached by that group applied to all three classes of channels.

For emphasis, sentences suggesting important areas for further investigation have been underscored. These suggestions reflect the bias of the participants, as well as the objectives of the workshop. Omission of an area in no way implies that it is unimportant.

### 2.1 GENERAL CONCLUSIONS

#### 2.1.1 Large Collecting Optics

With proper design, the performance of a system operating through free space, the atmosphere, or a cloud will continue to improve as the collecting area of the receiver is increased (for a given field of view). The investigation and development of large-aperture systems should be intensified.

The improvement in performance that can be gained by proper design (diversity combining) is substantial and does not require that the collecting optics be diffraction-limited. As an example, for a given aperture area,  $A_r$ , and a given performance, a heterodyne system employing proper combining will require on the order of  $A_c/A_r$  times the transmitter power required for a (single) heterodyning system, where  $A_c$  is the area over which the field is coherent in the receiver aperture. Comparable performance can be realized with direct-detection systems. Difficulties in implementation occur because either a large number of parallel heterodyne receivers or a very narrow-band optical filter and large detector area is required. These difficulties will be discussed further in connection with device needs.

#### 2.1.2 Device Needs

A continued development of space qualified optical components such as sources, modulators, and detectors is vital to the progress of optical communication systems. It is important, however, to recognize the limitations of the components which will emerge from this development and the importance of some devices not now under development.

Signal Generation: A wideband modulator is certainly a necessity for wideband FM and is very useful in many other communication systems, but it is not a necessity for high-rate digital systems. The required set of signals can often be generated by exploiting the pulse properties of lasers (mode locking, Q-switching, etc.). In fact, such direct generation has some distinct advantages, when it can be employed, because the electronic bandwidth of the drivers need only be the information bandwidth rather than the channel bandwidth. The point here is that the extreme versatility of a wideband

optical modulator is not always required. When it is not required substantial system simplifications often can be realized by using a signal-generation technique that is tailored to the application; for example, using a fast optical switch to select one of  $M$  possible waveforms. The development of specialized signal generation techniques should be pursued.

Signal Processing: Some promising systems for communication through the atmosphere would employ devices that have yet to be developed (see section 4.3). These same systems would be useful through clouds, or through free space with large receiving antennas that are not diffraction-limited (see sections 3.3 and 5.2). For example, one near-optimum system uses frequency and spatial filtering to reject as much background noise as possible, then independently detects the signal energy arriving from different angles, and finally combines the detector outputs. Such systems can be implemented as a collection of parallel direct detection or heterodyning systems, but the number of subsystems required is often enormous ( $10^3$ - $10^8$ ). An alternative implementation would consist of large focusing optics, an aperture stop, a single very narrow-band optical filter, and an ideal photon counter.

At present, neither of these implementations is very appealing. If the paralleling approach is to be practical, techniques for the fabrication of arrays containing a large number of parallel systems must be developed. In principle, this might be done by a microelectronic array in which each element consists of an optical detector followed by an IF filter and a square-law detector. The processed outputs would then be added together. No such cascades have been developed for parallel operation.

The second implementation avoids the array problem by employing a large area photodetector preceded by an optical filter. For the atmospheric and cloud channels, the product of the filter's area and field of view should be large relative to that of a diffraction-limited system, while its bandwidth should be very "small." Such devices are possible, in principle, and have been realized for some isolated applications by exploiting atomic effects. The optical filters now available fall far short of the ideal requirements of communication systems. The possibility of exploiting novel physical phenomena to gain improved filter performance should be explored.

Another obstacle to the development of systems along this second avenue is that relatively large-area, semi-ideal, photon counters are required. At present, these are only available at wavelengths shorter than  $\sim 0.9 \mu\text{m}$ . Thus an infrared photon counter would be very desirable.

Receivers and/or transmitters that measure and compensate for the phase distortion introduced by atmospheric turbulence, clouds, and antenna imperfections are also promising (see sections 3.3 and 4.3.3). The implementation of such systems requires, in essence, a lens, or phase grating, whose phase delay can be varied in time and space. As in the system just discussed, this variation can be achieved by paralleling a large number of simple systems, for example, the active mirror concept. It may be possible, however, to obtain this variation more simply and efficiently at the focal plane of a lens

with a distributed electro-optic system. The feasibility of a temporally and spatially adjustable phase grating should be explored.

Another promising system employs a transmitted reference, or "pilot tone," which is filtered out of the signal at the receiver and "spatially amplified" with the spatial phase kept intact. In some situations, the amplified reference can be used as a local oscillator which is in phase with the signal at every point on the detector surface (see sections 3.3 and 4.3.3). Consequently, the need for a large number of parallel systems or for a narrow-band optical filter for, say, the atmospheric channel can be avoided. The characteristics of wide-angle optical spatial amplifiers should be explored.

Multiplexing: The electronic bandwidth of existing and contemplated modulators and detectors significantly limits the information rate that can be realized with these devices. When higher information rates are required, multiplexing of the optical waveforms will be necessary. The development of optical multiplexing techniques and devices is probably a prerequisite to operation at data rates exceeding  $10^{10}$  bits/sec.

### 2.1.3 Systems Analysis

Many valuable studies of optical communication systems, particularly free-space systems, have been undertaken. The quantitative understanding of such systems is still meager. Thus, just as a continued improvement of optical devices is important to optical communication, so also is a more complete understanding of system performance characteristics.

The inadequacy of our present understanding is due, in part, to the novel characteristics of some optical channels, for example, the atmosphere; but the dominance of quantum effects at optical frequencies has been a more fundamental obstacle to quantitative results. Thus, although their output signal-to-noise ratios are comparable, direct detection and heterodyning systems yield different performance levels. We shall very briefly summarize the present understanding of these two systems and then consider the problem of optimum quantum systems.

Heterodyning Systems: Broadly stated, the analysis of heterodyning systems poses no fundamentally new difficulties. This is not to say that all systems of interest have been, or can easily be, analyzed. The required analysis techniques are well understood, and results are available (see sections 3.2.1 and 4.3.2) for some important situations. These analyses have made it possible to compare different signaling and post heterodyning processing schemes.

Direct Detection Systems: Unlike heterodyne detection, the analysis of direct detection systems involves a class of problems that have not been investigated previously. Consequently, the performance of only a few direct-detection systems is now known (see sections 3.2.1 and 4.3.1). The lack of quantitative results has made it difficult to determine optimum waveforms and post detection processing. It has also impeded performance comparisons between heterodyning and direct-detection systems.

Optimum Quantum Systems: It appears, at least in the binary case, that the optimum receiver for combating quantum noise is sometimes neither a direct-detection nor a heterodyning system. On the other hand, in all of the available results, either the direct-detection or heterodyning system is essentially as good as the optimum system. Thus, the question arises: Is the optimum quantum system ever significantly superior to both the heterodyning and direct-detection system?

If it can be shown that either a direct-detection or heterodyning system is "essentially" optimum for all channels of interest, the question becomes relatively unimportant to practical space communication. If, however, any example can be found for which the optimum quantum receiver is far superior to heterodyning and direct detection, the question becomes extremely significant. Even if essentially optimum systems can be realized by heterodyning or direct detection, the understanding of such systems will be enhanced and simplified when a unified quantum communication theory is available.

In addition to the general conclusions just discussed, there were conclusions pertaining specifically to the atmospheric and cloud channels.

## 2.2 ATMOSPHERIC CHANNEL

While a great deal of effort has been expended, for several years past, on the problem of obtaining an adequate theoretical and experimental understanding of the effects of atmospheric turbulence on optical propagation, the present state of knowledge is still not entirely satisfactory from a communication theory viewpoint. In particular, at the present time there is no general consensus of opinion on an adequate theoretical solution to the turbulent optical propagation problem. The workshop participants felt that the Rytov approximation (see section 4.1.1) was often adequate for communication system studies, particularly when an Earth-Space path was involved; this feeling was based largely on the reasonable agreement between this theory and existing measurements. On the other hand, there was no great confidence that this approximation technique adequately extrapolates to answer new questions where confirming measurements are not available. A better understanding of the limitations, as far as answering questions of interest in communication system design, of the Rytov technique is required; also, solution techniques that are both complete, in the sense of answering the communication questions of interest, and also more general, in the sense of being applicable to most physical situations, should be explored. Specific theoretical questions of interest in the channel characterization problem are discussed in section 4.1.5.

At the present time, data on the microscale changes of the atmosphere's refractive index are scarce and their interrelationships with other atmospheric conditions are not well understood. There are almost no direct data on the altitude variation of the refractive index process. The lack of confidence in the theoretical propagation results has also made indirect refractive index variation measurements suspect. Effort should be expended to understand the relationship of the refractive index variation process to other

atmospheric conditions and to better characterize this process, for example, examine the time behavior and ergodicity of the process. Measurements of the height distribution of  $C_N^2$ , the refractive index structure constant, should be made; the measurements should include direct temperature measurements, and if feasible, the development of some remote sensing technique.

Measurements of various parameters of the received optical field have been made. Many of these measurement programs have been concentrated on a single or small number of parameters, however, and often have not been guided by a communication objective. In particular, there are relatively few phase measurements available. There is a need for a comprehensive measurement program with the objective of supplying the information required to design and analyze communication systems. The emphasis should be on characterizing as completely as possible the received field, including the effects of variations in various transmitter and receiver parameters. Thus such quantities as the statistics of the received intensity and phase processes and the intensity and phase coherence bandwidths should be measured. Furthermore, the variations in the statistical properties of the intensity and phase processes with such parameters as the transmitting aperture size and illumination, the receiver aperture size and field of view, and the source wavelength are needed. Particular measurements of interest are listed in section 4.1.5.

The measurements discussed above should be made over propagation paths that are typical of the communication links that are of interest. In particular, data are needed for ground-space paths if a ground-space link is the one of interest. With a better theoretical understanding of turbulent propagation and measurements of the turbulence altitude profile, it may be possible to extrapolate careful ground measurements to ground-space paths. It will always be preferable to have measurements over the actual path, however. Thus careful consideration should be given to using stellar measurements when they are applicable. Such consideration should include performing additional stellar measurements if it is found that appropriate data are not now available, for example, statistical measurements and daylight measurements. Obviously as sources and detectors become available in space they should be used to make the desired measurements. Indeed the sources and detectors put into space in the near future should be so designed that they maximize the number of meaningful channel characterization measurements that can be made.

The analysis and synthesis of optical communication systems for the atmospheric channel has been impeded by the lack of a tractable representation for the received field process. Often the appropriate process model for the received field is, conditionally, the sum of log-normal and Gaussian processes, which presents analysis difficulties. Work is needed on alternative channel representations and appropriate approximations. In particular, the relationship between aperture and focal plane representations should be examined with emphasis on the form in which the degrees of freedom appear. An angular spectrum representation for the received field should be investigated both



theoretically and experimentally. (See section 4.2 for more details.) It should be emphasized that this channel representation work is not independent of the measurement program suggested above, but rather should both be guided by measurement results and used to suggest appropriate measurement and analysis questions.

Communication systems for the atmospheric channel must process the received field in such a way as to properly treat the spatial diversity and combat the fading. For both direct-detection and most heterodyning systems, the best processing will require an array of detectors. At the present time, however, neither the general performance of optimized single-detector systems nor the performance of array systems is known for the case of direct detection; further, the optimum processing is also, in general, unknown. Thus the tradeoff between performance and complexity for direct-detection systems is not known.

For heterodyning systems, more results are available. Some performance results exist for both single-detector and array systems. These results indicate that proper spatial diversity combining is critical when heterodyne detection is used and that useful aperture sizes may be relatively small or the number of array elements large, particularly at visible frequencies. Phase-front tracking receivers can be used to increase the effective aperture area for a single heterodyne detector (by approximately a factor of 10) or reduce the number of array elements required for a given aperture. Heterodyning systems also have many unanswered questions. For example, the data processing is not known for signaling times that are long relative to the channel coherence time, or for analog modulation. As discussed in section 2.1.2, the important question of practical implementation of heterodyning array systems is unsolved. Analysis work should be performed to provide a better understanding of both direct-detection and heterodyning systems for the atmospheric channel. As the design tradeoffs become better understood, experimental work should be performed to validate the approximations and assumptions that went into the system designs.

The use of adaptive systems for the atmospheric channel was discussed at the Workshop. Such systems include both channel-measurement receivers and adaptive transmitters (see section 4.3). Channel-measurement techniques look promising, because of the channel's long coherence time. With respect to measurement receivers, the important issues are how well the channel variations can be measured, the performance improvement that is possible by using these noisy measurements, the system implementations required, and the signal design problem.

With respect to transmitter adaptation, the question is: To what degree can a transmitter compensate for distortions that are introduced by the channel? The answer involves the issues of channel measurement and reciprocity, as well as transmitter strategy. Specific examples of adaptive transmitter techniques are variable-rate systems, adaptive signal design systems, and adaptive antenna (beam-shaping) techniques. The performance improvement that might be realized by both channel-measurement receivers and adaptive transmitter techniques and practical means of implementing them

should be investigated. The questions of channel "state" determination should be studied and measurements should be made to determine the validity of the concept of atmospheric reciprocity (see section 4.3).

The use of transmitted reference signals and 'imaging' preamplifiers, as discussed in sections 3.3 and 4.3, are techniques with promise for the atmospheric channel. Thus transmitted reference signaling schemes and imaging preamplifier receiving techniques should be studied further to determine their performance and feasibility.

### 2.3 SCATTERING CHANNELS

The feasibility of optical communication through strongly scattering media such as clouds, fog, and haze appears to have received very little attention. The available knowledge and experience with related channels suggests, however, that such communication may be feasible – at least at modest information rates.

The effect of a given nonabsorbing scattering medium depends upon its location relative to the system terminals. If it is located "near" the transmitter, its major effect is to spread the transmitted beam and hence to reduce the energy that can be collected by the receiver. This reduction imposes a basic limitation upon the potential for communication through scattering media near the transmitter. If the scattering medium is located "near" the receiver, its major effect is to destroy the spatial and temporal coherence properties of the received field; however, its presence does not lead to an enormous reduction in received energy if the receiver is properly designed. Consequently, a non-absorbing cloud near the receiver, for example, space-to-ground communication, does not limit the potential for communication as severely as does a cloud near the transmitter, for example, ground-to-space communication. Since all-weather operation of optical systems is sometimes important, the feasibility of communication through clouds, fog, and haze, should be investigated.

The initial effort should be to obtain a description of the radiation emerging from these channels that is adequate for the analysis of communication systems. Some means of obtaining this description are discussed in Section V. The difficulties of an exact and complete analytical description dictate that approximate analysis, simulation, and experimental data be employed. Even in the absence of a complete description, valuable estimates of the potential for communication through clouds, fog, and haze, can and should be obtained from some simple experiments (section 5.3).

A question that emerged at the Workshop concerned the absorption coefficient of water and ice at  $10.6 \mu\text{m}$ . The tabulated values of absorption are larger than the values suggested by recent measurements on natural fog and rain, as well as on artificial clouds. This uncertainty is important because the tabulated data suggest that communication through clouds at  $10.6 \mu\text{m}$  is not competitive with visible wavelengths, while the recent measurements suggest that it may be. Experiments to resolve the uncertainty in the absorption of ice and water at  $10.6 \mu\text{m}$  should be performed.

A thorough investigation of the design and performance of systems for scattering channels must await a more complete description of the channel characteristics. Preliminary estimates of the performance characteristics and device requirements for optical communication through clouds, fog, and haze can, and should, be performed. The general structure of these systems will probably be similar to atmospheric systems, but the device requirements and performance limitations will be more severe (section 5.1).

### III. Free-Space Systems

Free-space systems, as typified by satellite-to-satellite communication, were not major topics of discussion at the Workshop, but many of the unresolved problems encountered in atmospheric and cloud systems are also encountered, in a simpler form, in free-space systems. Because of this, and since these systems are important in their own right, we shall first consider free-space systems.

By a free-space system we mean one in which the propagating medium only attenuates the transmitted signal, for example, interplanetary space or a nonturbulent atmosphere. Such systems are in many ways the optical equivalent of line-of-sight microwave systems. The quantum and pointing issues, however, that are unimportant at microwave frequencies become fundamental at optical frequencies.

At microwave frequencies the limitations upon antenna size and beamwidth are such that adequate pointing accuracy can be attained without limiting the communication system's performance; at optical frequencies pointing accuracy and antenna stability impose important constraints upon the usable antenna sizes. Although this suggests that antenna pointing should be considered as an integral part of the communication system, a comprehensive discussion of the pointing problem was beyond the scope of this Workshop and the problem was not discussed in detail.

Traditionally, it has been possible to view a "radio receiver" as a device that adds some noise to its input, or inputs, and then performs noise-free operations (mixing, filtering, detection) upon the resultant sum. The important point is that the nature and intensity of the added noise is independent of the operations performed. In particular, it has often been possible to characterize a receiver by a temperature that is a measure of the quality of the receiver front end. Consequently, questions of optimum system design could, to a great extent, be separated from questions of receiver noise characteristics. At optical frequencies the situation is fundamentally different because of quantum effects. In particular, the character of the noise associated with a receiver is strongly dependent upon the operations that it performs. Moreover, this situation is a consequence of fundamental (quantum) physical considerations; it cannot be relieved by improved component design.

Three approaches to optical communication systems have been fruitful thus far. The first approach employs heterodyne detection and introduces a quantum noise that is a Gaussian random process. Broadly stated, these systems are well understood. The second approach uses energy-detection devices and introduces a "quantum noise" that is a shot-noise process. There is renewed interest in problems of signal design, detection, and parameter estimation for such processes. The final approach, and probably the one least understood, is the reformulation of the communication problem in quantum-mechanical terms. As we shall discuss, there are fundamental problems in this approach, but it appears to be the only way of determining the ultimate performance limits of optical communication systems. Before considering this approach we shall review the present technology in components and receiver structures.

### 3.1 EXISTING RECEIVER COMPONENTS

All optical receivers that have been constructed, or contemplated, have employed direct detection (energy detection), heterodyning, homodyning, or preamplification. Since heterodyning and homodyning have only been implemented efficiently with energy detection and amplification, we only consider energy detectors and amplifiers. Components, such as dielectric filters, which are well understood, are not discussed nor are some of the sophisticated components, such as parametric converters and second-harmonic generators.

#### 3.1.1 Energy Detectors

Many parameters of an energy detector are useful in the design of optical communication systems: sensitivity, bandwidth, internal noise power, area, and so forth. A more complete statistical description of the detector is required, however, if its ultimate utility in a communication system is to be judged. One of the questions discussed at this Workshop was whether or not such a description exists, in principle, for the various classes of available detectors. The conclusion was that it does – at least in broad outline. Since this general description has not been presented in a form that is accessible to many communication engineers, we present it here.

Ideal Energy Detectors: It is generally agreed that the output of an ideal energy detector of quantum efficiency  $\eta$  is a conditional Poisson process.<sup>1-5</sup> That is, given a linearly polarized classical field\*,  $E(\vec{r}, t) = \text{Re} [V(\vec{r}, t) \exp j2\pi\nu t]$ , incident upon the detector, the sequence of detections forms a time-varying Poisson process of events whose rate parameter  $\lambda(t)$  is

$$\lambda(t) = \frac{\eta \int |V(\vec{r}, t)|^2 d\vec{r}}{h\nu}, \quad (1)$$

where the integral is over the detector area. If the process  $V(\vec{r}, t)$  is not narrow-band, one should also include the wavelength dependence in  $\eta$  and  $\nu$ . An equivalent description of the detected process could be stated thus: Given the incident field, the numbers of detected photons in any two disjoint time intervals are independent of each other, and the probability of detecting  $k$  photons in the time  $t$  to  $t+T$  is

$$P(k) = \frac{\bar{K}^k}{k!} \exp -\bar{K}, \quad (2)$$

where

---

\* If a photon description of the field is used, an identical result can be obtained by considering the detector as a binary branching process.

$$\bar{K} = \int_t^{t+T} \lambda(\tau) dt. \tag{3}$$

It is important to note that the classical field incident upon the detector is presumed to be known. If it is random, the conditional Poisson process must be averaged over the statistics of the incident field. Thus, for example, if the incident field is an extremely narrow-band, zero-mean, Gaussian random process, the probability of  $k$  counts in a time that is small compared with the inverse of the optical bandwidth will obey a Bose-Einstein distribution.

Real Photon Detectors: Real detectors differ from ideal detectors in several ways. Although the details of these differences depend upon the specific detector that is considered, it appears that the following characterization is generally accepted by physicists.

Given the classical field incident upon the detector, the output of the detector and its associated amplifier can be expressed as

$$f(t) = \sum_i h(t-t_i, t_i) + n_T(t) + n_S(t) \tag{4}$$

and can be interpreted as illustrated in Fig. 1. In this expression  $n_T(t)$  is a thermal noise (zero-mean, stationary, and Gaussian) that is typically white over the band of interest. The  $t_i$  are the detection times for an ideal photon detector, that is, given  $\lambda(t)$  in Fig. 1, they comprise a Poisson random process with rate parameter  $\lambda(t)$ . The function  $h(t-t_i, t_i)$  is the shape of the detector output pulse which is caused by a photon

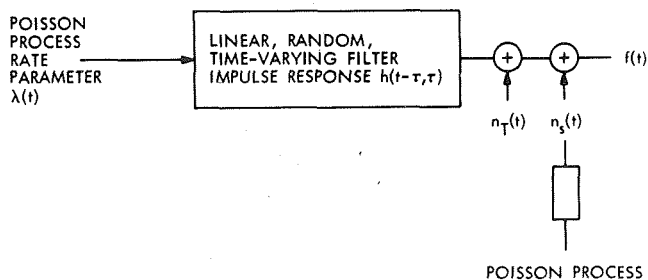


Fig. 1. Photon detector model.

detected at time  $t_i$ . In general, this pulse can itself be a random function. As is shown in Fig. 1,  $\sum_i h(t-t_i, t_i)$  can be interpreted as the output of a random time-varying filter driven by a conditionally Poisson process. Finally,  $n_S(t)$  is a signal-independent "shot noise" that can also be viewed as the output of a random time-varying filter driven by an (unconditional) Poisson process. This can represent detector dark current.

To illustrate the application of Eq. (4), we consider several examples in the absence of thermal noise and dark current.

- i. Ideal photon detector:  $h(t-t_i, t_i) = \delta(t-t_i)$ , where  $\delta(\cdot)$  is a unit impulse.
- ii. Photodiodes: to first order  $h(t-t_i, t_i) = g(t-t_i)$ , where  $g(\cdot)$  is a fixed known function whose Fourier transform is the detector frequency response.
- iii. Photomultiplier: to first order  $h(t-t_i, t_i) = \sigma_i g(t-t_i)$ , where the set of  $\sigma_i$  are independent, identically distributed, random variables.
- iv. Photoconductor: to first order  $h(t-t_i, t_i) = g[(t-t_i)/\sigma_i]$ , where the  $\sigma_i$  are independent, identically distributed, random variables. The probability density of  $\sigma_i$  is

$$p(\sigma_i) = \frac{1}{\tau} \exp - \frac{\sigma_i}{\tau}.$$

- v. Avalanche Photodiode: Similar to iii, except that the probability density of  $\sigma_i$  differs. Typically, the variance of  $\sigma_i$  is of the form  $(\bar{\sigma})^a$  with  $a < 2$ , where  $\bar{\sigma}$  is the mean value of  $\sigma_i$ .

For a first approximation,  $n_T(t)$  is the dominant additive noise in detectors with small internal gain (photoconductors), while  $n_S(t)$  is the dominant internal noise in devices with large internal gain (photomultipliers and avalanche diodes). At present, photodiodes and photomultipliers are limited to wavelengths less than approximately  $0.9 \mu\text{m}$ , avalanche detectors are limited to wavelengths shorter than approximately  $5 \mu\text{m}$ , and photoconductors are limited to wavelengths shorter than approximately  $200 \mu\text{m}$ . It is reasonable to say that photomultipliers are close approximations to ideal photon detectors with a quantum efficiency of less than approximately 30%. The detailed noise characteristics of available devices have been discussed extensively.<sup>6-12</sup>

### 3.1.2 Preamplifiers

Although not used extensively, optical preamplifiers of modest gain are available at all laser wavelengths (every laser can be used as an amplifier). For an ideal, high-gain, four-level preamplifier, the equivalent input noise can be considered to be zero-mean and Gaussian with a power density of  $h\nu$  W/Hz per spatial mode<sup>\*13</sup> [If the field of view is too large, a high gain cannot be realized.] The total noise power per Hz is then  $h\nu A\Omega/2$  where  $A$  is the area,  $\lambda$  is the wavelength, and  $\Omega$  is the solid-angle field of view of the receiver. The noise levels per mode for lower gain, or three-level, devices are larger than  $h\nu$ .

## 3.2 EXISTING RECEIVER STRUCTURES

Existing receivers employ optical amplifiers, local oscillators, and energy detectors in one of the following system configurations: heterodyning, homodyning, or direct

---

\* All power densities are unilateral or "measured."

detecting. Since the general characteristics of these systems have been discussed extensively, we review them only briefly.<sup>10, 14-24</sup> This review will provide a basis for the subsequent discussion of some important unresolved questions pertaining to all of the channels that we consider.

### 3.2.1 Heterodyning

At present, heterodyning is accomplished by adding a strong local-oscillator field to the received field, detecting the sum with an energy detector, and passing the detected waveform through a bandpass filter. If the received and local-oscillator fields are linearly polarized plane waves propagating in the same direction, and if the local oscillator is strong enough, the output signal will, in effect, be a frequency-translated replica of the incident "waveform" plus a zero-mean, stationary, Gaussian noise of power density  $h\nu$ , where  $\nu$  is the carrier frequency. The assumption here is that the bandwidth of the received field is small relative to the carrier frequency.

If the local-oscillator field is a plane wave, but the received field is not, the system, in essence, extracts those plane-wave components of the received field whose angle of propagation differs from that of the local oscillator by no more than an angle  $\lambda/D$ , where  $\lambda$  is the wavelength, and  $D$  is the diameter of the heterodyning area (which is assumed to be circular). More generally, one can say that heterodyning extracts the temporal variation of one spatial mode of the received field, translates it to an intermediate frequency, and adds to it a zero-mean Gaussian noise of power density  $h\nu$ .<sup>25,26</sup> The particular mode that is extracted depends upon the spatial variation of the local-oscillator field.

One of the important characteristics of a heterodyning system is that internal detector noise ( $n_s(t)$  and  $n_T(t)$  in Eq. (1)) is unimportant. It is only the quantum, or zero-point fluctuation, noise that contaminates the received field. Thus, for example, if the input "signal" power in the relevant mode is  $P$ , the output signal-to-noise-power-density ratio will be  $P/h\nu$ , provided that the local oscillator is strong enough.

### 3.2.2 Homodyning

A homodyning system can be viewed as a heterodyning system in which the intermediate frequency is zero. That is, the original optical field variation is translated to base-band. If the system is coherent, that is, the local-oscillator field is in phase with the signal component of the received field, the equivalent input noise power density is  $h\nu/2$  W per Hz. Roughly stated, the noise power reduction can be attributed to the phase information that is available to the system.

### 3.2.3 Direct Detection

By a direct detection system we mean one composed of collecting optics, a narrow-band filter, and a photon detector. Conditioned upon the received field the system output is described as in section 3.1.1. This description is of limited utility because the



received field itself usually contains a noise process, and one desires the output statistics unconditioned upon the noise process. In general, it is difficult to obtain these statistics, but it is possible to define a signal-to-noise ratio that is sometimes a useful measure of performance.

Specifically, suppose that the received field consists of a signal of instantaneous envelope power  $P_s$  plus a zero-mean background noise of power density  $P_b$  W/Hz. Also, suppose that the total internal noise of the detector is of mean zero and power density  $P_n$  W/Hz. Subject to some reasonable assumptions, the signal-to-noise ratio per Hz referred to the system output will be

$$\frac{\eta \bar{G}^2 P_s^2}{2h\nu B [\bar{G}^2 (P_s + P_b) + P_n]}, \quad (5)$$

where  $B$  is the signal bandwidth;  $\eta$  is the quantum efficiency, and  $\bar{G}$  and  $\bar{G}^2$  are, respectively, the mean and second moment of the detector gain.<sup>27, 28</sup>

It is clear from (5) that the detector internal noises will influence the signal-to-noise ratio unless

$$P_n \ll \bar{G}^2 (P_s + P_b).$$

At present, it is possible to satisfy this condition for wavelengths shorter than approximately  $1 \mu\text{m}$ . For longer wavelengths, the internal noises are significant, even at detector temperatures of  $4^\circ\text{K}$ . That is, near ideal photon counters are available at wavelengths shorter than  $1 \mu\text{m}$ , but not at longer wavelengths. [Note that we are ignoring the important questions of detector bandwidth and quantum efficiency.] It is also worth noting, for future reference, that the detector areas available at wavelengths shorter than  $1 \mu\text{m}$  are much larger than those available at longer wavelengths.

#### 3.2.4 System Performance

We shall now briefly summarize the state of knowledge about the performance of optical communication systems operating through a free-space channel. The emphasis will be upon important questions that remain to be answered: the corresponding questions for other channels will be discussed eventually.

The performance analysis of heterodyning and homodyning systems is, in essence, the analysis of systems employing known signals in additive white Gaussian noise. Such systems have been studied extensively during the past two decades and are now well understood. Thus the optimum signal processing, the optimum signal sets, and the performance of optimum and many suboptimum systems have all been determined.<sup>29-34</sup>

The performance of direct-detection systems is a subject of recent investigations.<sup>35-37</sup> Most of these investigations have been addressed to problems in which the detector output is a Poisson process when conditioned upon the signal component of

the received field but averaged over the background noise. This situation will prevail for an ideal photon detector, or a detector whose internal noise is Poisson, if there is no background noise. It will also prevail if the predetection noise bandwidth is much greater than both the post detection bandwidth and the signal bandwidth, as it often is for daytime operation with existing optical filters. These investigations have established the optimum post detector processing, optimum signal sets, and system error probability for the optimum and some suboptimum systems.

Some investigations have also been addressed to direct detection of specific signal sets when the noise bandwidth is comparable with the signal bandwidth, that is, when proper predetection filtering is employed.<sup>38-40</sup> As would be expected, the system performance improves as the filter bandwidth approaches the signal bandwidth. Although appropriate filters do not now exist, one can compensate for the filter inadequacies by proper signal design, at least in part.

We shall not attempt a detailed comparison of heterodyne, homodyne, and direct-detection systems. A few general comments may be of value.

Roughly stated, homodyne detection is superior to direct detection for small-alphabet, phase-coherent systems.<sup>40</sup> This is so because simplex signals such as phase reversal keying can be employed. For larger alphabets, direct detection with proper prefiltering appears to be superior (see Appendix C), but such systems have yet to be implemented. The performance of large-alphabet direct-detection systems using wideband filtering can be either superior to or inferior to heterodyning and homodyning systems, depending upon the background noise level and predetection bandwidth. It is important to note that these statements are based upon isolated examples. Much more work must be done before the relative merits of different systems can be fully determined – even for the free-space channel.

### 3.3 LARGE-APERTURE RECEIVERS

For a given collecting area, the optimum receiver for the free-space channel uses diffraction-limited optics. On the other hand, the available and contemplated size of diffraction-limited optics is relatively small,<sup>41</sup> so it is possible that improved system performance could be achieved by employing larger nondiffraction-limited optics.

There are several ways in which such a system could be implemented. If, for example, direct detection were employed, the system could differ from a diffraction-limited system only in the use of a larger detection area, because of the larger point spread function of the optics. Of course, this larger area will result in an increased noise level. A more refined direct-detection system might employ predetection spatial filtering to compensate for some of the distortion introduced by the optics. The objective of this compensation would be to reject spatial modes that contain only noise.

Heterodyning with nondiffraction-limited optics is more difficult than direct detection because, as noted in section 3.2.1, heterodyning systems respond to a single, or at most to a few, spatial modes of the received field. Thus a nondiffraction-limited

heterodyning system will not extract all of the available energy in the aperture, unless the local-oscillator field is properly matched to the signal field that exists at the mixing plane. This matching requires that either the phase distortion or "aberrations" introduced by the optics be known to, or measured by, the receiver, or that the local oscillator transverse that portion of the system that is not diffraction-limited.

Although it may be possible to achieve the required matching by compensation of either the signal field or local-oscillator field, it would appear to be more difficult to implement such a system than it would be to implement the direct-detection system. Under the assumption that both systems can be implemented, their relative performance depends upon the background noise level and detector internal noise.

Transmitted reference systems provide another approach to operation with large apertures that are not diffraction-limited. Such systems can be viewed either as heterodyning systems in which the local oscillator is located at the transmitter, or as direct-detection systems using a particular signaling scheme. For this discussion the former view is more useful, since it clearly demonstrates that the "local-oscillator" field passes through that portion of the system that is not diffraction-limited. One disadvantage of a transmitted reference system is that a strong "local-oscillator" signal is obtained at the expense of signal power. As indicated in Fig. 2, the loss of signal

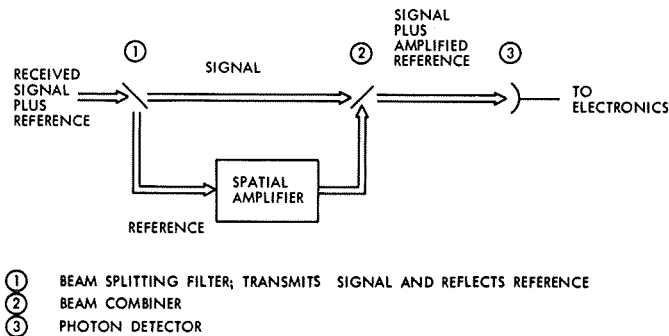


Fig. 2. Heterodyne reception with a transmitted reference.

power can be reduced, at least in principle, by amplifying the reference field while preserving the phase of the spatial modes.<sup>13</sup> The reduction is gained at the expense of increased system complexity and a noisier local-oscillator signal.

The feasibility, relative merits, and ultimate limitations of all three of the schemes above should be explored. At this junction the only firm conclusion is that large-aperture receivers with nondiffraction-limited optics should be given careful attention. It is worth noting that this conclusion will be reached for all of the channels that we consider.

### 3.4 QUANTUM COMMUNICATION THEORY

It is apparent from the preceding discussion that quantum effects comprise an important component of the system "noise," and that the nature of this noise depends in detail

upon the system structure. Although this "noise" can be treated by semiclassical means once the system structure has been specified, fundamental difficulties are encountered in determining the optimum system structure. These difficulties are the subject of quantum communication theory.<sup>38-40, 42-47</sup> We shall now describe the nature of these difficulties, briefly review the approaches that have been explored, and summarize the outstanding unresolved questions. In general, our remarks are applicable to all optical channels, not just the free-space channel.

It is clear from the discussion of heterodyning and direct-detecting systems that optical receivers are subject to a noise whose properties are determined by the receiver's structure. Stated alternatively, the receiver interacts with the channel in a way determined by the detailed nature of the receiver. This interaction poses no fundamental conceptual problems in the analysis of any specific system, but it does raise deep issues in system optimization. Specifically, over what class of receivers is the optimization to be performed? Clearly, we prefer to optimize over the class of all receivers that are consistent with the laws of physics. The problem then becomes that of characterizing the class of such receivers in a manner that is amenable to analysis.

Two approaches to this problem are now available; both are based upon the axioms of quantum theory. The first, and perhaps more traditional, begins with the premise that every receiver can be characterized by a set of "observables" and that the performance of the receiver can be determined according to the techniques of quantum theory (see Appendices D and E). This approach has been rewarding, but it does have some shortcomings. In particular, it appears that receivers exist that cannot be characterized by a set of observables, for example, a system employing optical preamplification.

A second approach to receiver characterization has recently been introduced (Appendix F). It consists, in essence, of relaxing some of the constraints that an "observable" must satisfy. Precisely stated, the first approach characterizes a measurement by a complete set of orthogonal vectors; in the second approach the vectors need only be complete. The second approach does encompass optical preamplification and, perhaps, may encompass all receivers that are consistent with the laws of physics.

Both receiver characterizations are useful, but fundamental problems are encountered in their application. For simplicity, we discuss these problems in the context of the first approach; that is, we assume that every receiver of interest can be described by a set of "observables."

First, it is doubtful that every "observable" corresponds to a receiver — even ideally. Thus, if one attempts to find the best receiver in the class of those characterized by observables, he may be led analytically to an observable that does not characterize a receiver. This situation is further complicated by lack of any synthesis procedure. That is, there is, at present, no way of determining how a given observable might be

"measured." Instead, it is necessary to recognize the observable as that associated with a known measurement. These problems are fundamental and must be resolved before the performance of quantum-limited communication systems can be fully assessed.

In addition to the conceptual problems just discussed, there are many difficult and unresolved problems in the analysis of quantum systems. For example, even for binary signaling, the optimum system performance has been determined only for a very limited number of waveform sets and channels. For larger alphabets little progress has been made, except for a special case that corresponds roughly to orthogonal signaling alphabets (see Appendices C, D, and F).

Although the available results are fragmentary, they do suggest some interesting conclusions and questions. First, in some specific situations the performance associated with the optimum observable is superior to that attained by the heterodyning, homodyning, and direct-detecting systems usually contemplated. Moreover, it appears that the optimum observable can, at least in some cases, be implemented in principle (see Appendix F). On the other hand, in the problems that have been solved, this superiority in performance has been rather small – so small as to be of little practical importance.

It is apparent that our understanding of quantum systems is quite meager. This is due in part to the limited attention that such systems have received, and in part to the conceptual and analytical difficulties involved. In any event, the study of such systems represents an important aspect of optical communication and should be continued. In particular, an effort should be made to determine whether or not there are systems whose performance is substantially superior to that which can be realized by a combination of optical amplifiers, local oscillators, and ideal photon counters.

## IV. Atmospheric Channel

We shall consider optical communication systems involving propagation through the clear turbulent atmosphere. The term "clear atmosphere" emphasizes that only the effects that are due to random refractive index variations are being considered. Haze, smog, clouds, and other atmospheric conditions characterized by concentrations of aerosols causing scattering effects are considered in Section V.

The effects of interest to us are caused by the random variations of the optical refractive index in time and space. These refractive index changes are caused by small temperature variations in the atmosphere, which in turn are caused by turbulent mixing of various thermal layers. Although the typical temperature variations are only 1°K, the resulting refractive index variations of a few parts in  $10^6$  cause extensive effects on optical radiation fields propagating through the atmosphere. The astronomer viewing a star has long been faced with image dancing or wander, image distortion (tremor disk), and scintillation.<sup>48</sup> The image dancing, observed in small telescopes, corresponds to variations in the angle of arrival of the energy. The tremor disk, observed in larger telescopes, corresponds to the deviations from a plane phase front caused by refractive index variations. Scintillation is the term used to describe the variations in intensity of the image. Thus the random refractive index variations make the received field fade in both time and space and also cause extensive phase front distortion.

At first glance, these effects might appear to severely limit communication through the atmosphere at optical frequencies. Indeed, some investigations have shown that systems that are near optimum in the absence of turbulence suffer large performance degradations in its presence. On the other hand, experience at other frequencies and preliminary results at optical frequencies suggest that atmospheric turbulence may not severely limit either rate or reliability if the system is properly designed. One of the major objectives at the Workshop was to determine the issues that must be examined before the potential for atmospheric optical communication can be fully assessed. All of these issues revolve around the ideas of modern communication theory.

A prerequisite for the application of the ideas of modern communication theory is a statistical description of the effects caused by the channel. For a clear atmosphere at optical frequencies, a complete statistical description of the received field, given the transmitted field, is required. A goal of the group studying the turbulent atmospheric channel was to assess the present state of theoretical and experimental knowledge related to this description to establish the areas in which further work would be required before an adequate model could be obtained. A second goal was to use the available channel descriptions to assess avenues of system development that might ultimately lead to efficient systems.

We shall review the effects of atmospheric turbulence upon optical radiation and the areas in which further knowledge is required. Then we shall discuss a channel model that is consistent with available knowledge. Finally, we shall use this model to explore

the avenues along which efficient systems might be developed. The emphasis will be on a ground-to-space or space-to-ground path, but the results for terrestrial paths will be used whenever they are helpful.

#### 4.1 TURBULENT EFFECTS

We turn now to a detailed examination of the effects caused by atmospheric turbulence, that is, by the random refractive index variations.

##### 4.1.1 Theoretical Approaches

Theoretically, the turbulent irregularities in the atmosphere should cause some depolarization of a linearly polarized light wave. The effect has been calculated and estimated to be approximately  $10^{-9}$  per km.<sup>49</sup> Saleh was unable to observe the depolarizing effect of the atmosphere on a 2.6-km path, even with an equipment sensitivity of -42 dB. We conclude, at the present time, that depolarization effects of the atmosphere are of no practical interest in the design of optical communication systems. Consequently, theoretical studies of turbulent effects reduce to solving the wave equation, with the refractive index described as a random field in time and space.

By the use of a frozen atmosphere assumption<sup>50</sup> and Fourier transform techniques, the problem is further reduced to solving the Helmholtz equation

$$\nabla^2 U(\vec{r}) + k^2 [n(\vec{r})]^2 U(\vec{r}) = 0, \quad (6)$$

where  $U(\vec{r})$  is the complex amplitude of one frequency component of the field,  $k = 2\pi/\lambda$ , and  $n(\vec{r})$  is the random refractive index at an instant of time. The time behavior is introduced by assuming that the refractive index field, which is frozen in space, is blown across the optical path by a normal wind component. The temporal behavior is thus specified in terms of the spatial behavior that is already known. The validity of the frozen atmosphere assumption depends on the fact that the refractive index changes are very slow relative to the motion induced by the wind.<sup>50</sup>

Several approximate solution techniques have been tried on Eq. (6) and all have limitations. One of the most popular and controversial of these is the Rytov solution technique used extensively by Tatarski,<sup>50</sup> Chernov,<sup>51</sup> and others. This approximation has recently been under severe attack,<sup>52-55</sup> and must certainly be modified for very long and/or very turbulent paths to account for certain experimental facts. A modification has already been proposed by Tatarski, and others<sup>56</sup> are under active development. While the Rytov approximation technique has many critics and some proven failings, it does have some important virtues. The most important of these is that it has been in agreement with most of the available experimental evidence. Furthermore, while it is known to break down for very turbulent or long paths, it should be applicable for almost any ground-space path. Moreover, it is by far the most satisfactory theory now available.

In the Rytov technique it is assumed that the solution to Eq. (6) is of the form

$$U(\vec{r}) = U_0(\vec{r}) \exp \gamma(\vec{r}), \quad (7)$$

where  $U_0(\vec{r})$  is the solution when  $n(\vec{r}) = \overline{n(\vec{r})}$  and the bar indicates an expectation operation,

$$\gamma(\vec{r}) = \chi(\vec{r}) + j\phi(\vec{r}), \quad (8)$$

with  $\chi(\vec{r})$  the normalized log amplitude, and  $\phi(\vec{r})$  the phase difference between the unperturbed solution and the actual solution. The technique involves substituting Eq. (7) in Eq. (6) and linearizing the resulting equation for  $\gamma(\vec{r})$ , which can then be solved. The importance of  $\gamma(\vec{r})$  in the theory has made it occupy a central role in experimental work.

To study  $\gamma(\vec{r})$ , it is necessary to have a description of the random refractive index  $n(\vec{r})$ . This description is now considered to establish some needed notation. In terms of  $N = (n-1) 10^6$ , the deviation of the refractive index from unity in parts per million, the refractive index fluctuations are proportional to the temperature fluctuations, with the relationship<sup>57</sup>

$$\delta N = \frac{79}{V-1} \frac{p}{T^2} \delta T, \quad (9)$$

where  $p$  is the pressure in mb,  $T$  is the temperature in °K, and  $V = c_p/c_v$ , the ratio of specific heats, is approximately equal to 1.4 for air. Variations in the mean of  $N(\vec{r})$  with  $\vec{r}$  correspond to large-scale or systematic refractive effects that are not being considered here. It follows from the Kolmogorov theory of turbulent flow at large Reynolds numbers, particularly as applied to temperature fluctuations by Obukhov and Corrsin, that the second-moment description of the spatial variations of  $N$  is given by the relation<sup>58-60</sup>

$$D_N(r) = C_N^2 |\vec{r}|^{2/3} \quad (10)$$

for  $r < L_0$ , where  $L_0$  is the outer scale of the turbulence and is of the order of meters. Here  $D_N(r)$  is the refractive index structure function defined by

$$D_N(r) = \overline{[N(\vec{R}+\vec{r}) - N(\vec{R})]^2}, \quad (11)$$

where  $r = |\vec{r}|$  and the bar indicates an expectation operation over the ensemble of atmospheres.  $C_N^2$  is the refractive-index structure constant that depends strongly on meteorological conditions, time of day, and altitude, among other things.  $C_N^2$  has been measured for various conditions on horizontal paths.<sup>61-63</sup> The variation of  $C_N^2$  with altitude has been discussed,<sup>64</sup> but actual measurements are badly needed, as we shall discuss in sections 4.1.2 and 4.1.4.

We shall see that this second-moment characterization of the refractive index variations appears to be sufficient to answer most questions of interest. A less adequate



second-moment model that describes the spatial variation of  $N(r)$  in terms of a Gaussian-shaped correlation function has also been used, for mathematical simplicity, in some theoretical studies.<sup>51</sup> Certain detailed questions involved with the time evolution of the effects caused by the refractive index variations may require a more complete statistical model than the second-moment characterization now being considered.

After this short survey of the over-all theory, we now turn to a more complete examination of the theoretical and experimental results associated with various aspects of the received field.

#### 4.1.2 Intensity Scintillation

Intensity scintillation caused by turbulence has long been noted by astronomers, and it is the most noticeable feature of laser radiation received over an atmosphere path. The turbulence causes the received field to fade in both time and space; that is, the beam exhibits a pattern of light and dark spots that change with time. To first order, the pattern seems to move across the beam.

To examine the effect of this phenomenon on optical communication systems, it is necessary to characterize the random amplitude variations conditioned on the values of the transmitted field. This characterization is considerably simplified if the frequency coherence bandwidth of the turbulent effects exceeds the bandwidth of the signal that is being considered, that is, if the amplitude and phase of each frequency in the transmitted signal undergoes the same turbulent change. This appears to be a reasonably good assumption for most signal bandwidths. While there are not direct measurements, both theoretical and experimental evidence indicate that the envelope and phase variations caused by the turbulence are highly correlated over large bandwidths for ground-space paths.<sup>65</sup> Stellar scintillation data indicate selective color fading over bands of 1000 Å and essentially identical intensity traces for 200 Å bands separated by 833 Å.<sup>66</sup> Pulse transmission studies with a GaAs laser over a 1.8-km path have shown that the coherence bandwidth is greater than  $10^9$  Hz.<sup>68</sup>

With the frequency coherence assumption, it is only necessary to characterize the scintillations that result for a transmitted field containing a single frequency. These scintillations obviously depend on the spatial character of the transmitted field; that is, whether the field is a plane wave, a spherical wave, or a collimated or focused beam.

Statistics: Both theoretical arguments and measurements support the contention that conditionally the received field is a log-normal process. For Gaussian beam and plane or spherical wave propagation with reasonable path lengths, it is easy under the Rytov approximation to use the Central Limit theorem to show that at any instant of time the set of  $\chi(\vec{r})$ 's associated with any set of  $\vec{r}$ 's (at least those in the central portion of the beam for that case) are jointly Gaussian random variables. Here  $\chi(\vec{r})$  is the normalized log amplitude discussed previously. If this result is extrapolated to claim that  $\chi(\vec{r}, t)$ , the quantity that describes the turbulent effects on the field envelope, is a

Gaussian random process of time and space, then the envelope of the received field is a log-normal process.

Experimental measurements of the first-order density of the intensity of the received field have been made over horizontal paths in both the spherical wave and beam cases and they tend to confirm the log-normal result.<sup>62, 68</sup> Note that a log-normal envelope distribution implies a log-normal intensity distribution. There is some evidence that the intensity distribution is not log-normal for very turbulent situations (or long paths).<sup>69</sup> For collimated beams, measurements have tended to be concentrated at the center of the beam. There is some evidence that near the edges of the beam the field amplitude variations may have a more Rayleigh or Rician character.<sup>68</sup> This may be due to the fact that scattering processes tend to cause the dominant effects in these weak intensity regions. While the intensity statistics near the beam edge may be of interest during the acquisition phase, they will generally not be important after the terminals are locked in and tracking one another.

The only intensity statistics measurements associated with a ground-space path are those from GEOS-II. These data are currently being reduced, but preliminary results seem to be in agreement with the log-normal hypothesis. There have been no measurements of the higher order densities in an attempt to verify or disprove the log-normal process assumption. Also, most measurements have been made with small-aperture receivers. As the receiver diameter is enlarged to include many regions with independent fading, one would expect the resulting output signal to show Gaussian fading. There is some experimental evidence<sup>62</sup> and also a theoretical argument<sup>70</sup> to indicate that this effect will either not be observed or will occur only with exceedingly large apertures. It is worth verifying this experimentally as large collectors become available. Also, while one would not expect any effect, measurements of the intensity statistics with large transmitting apertures would be desirable.

If we accept the premise that the received intensity is a log-normal random process, its statistical description reduces to a specification of its mean and covariance. This specification requires only a second-moment description of the refractive index variations. Even if the intensity is not log-normal, the mean and covariance functions are important in the design of efficient communication systems. Thus we now consider them in some detail.

Mean Intensity and Average Beam Divergence: The mean intensity for plane- and spherical-wave propagation is normally calculated by using a conservation-of-energy argument.<sup>71</sup> This amounts to an assumption that  $\bar{I} = I_0$ ; that is, the average intensity is equal to the intensity in the absence of turbulence. We shall show that this assumption, when coupled with the assumption of log-normal statistics, leads to the result that  $\overline{\chi} = -\sigma_{\chi}^2$ ; i.e., the mean of the normalized log amplitude is the negative of its variance.

In optical-communication applications, we are often concerned with the propagation of well-collimated (perhaps even focused) beams. The phase fluctuations induced by turbulent irregularities in the atmosphere cause portions of the beam to propagate in

slightly different directions and hence produce a spreading of the beam with a resultant loss in average antenna gain. The turbulence also causes small changes in the average direction of propagation of the beam, with the result that in the receiving plane the center of the antenna pattern is not fixed in space but rather moves about in time. Thus, for beam propagation, both of these effects will be reflected in the means and variances associated with the intensity profile in the receiving plane. The mean intensity profile can be considered to be the average antenna beam pattern. The intensity variance reflects variations from this average beam pattern, because of beam motion and spreading, as well as scintillation.

A rigorous analysis of the intensity mean and covariance functions associated with beam propagation is quite difficult. All theoretical work in the past has been confined to Gaussian beams. Beckmann has presented an approximate geometrical approach to the beam propagation problem.<sup>72</sup> The Rytov technique has been used by Kon and Tatarski<sup>73</sup> and by Schmeltzer<sup>74</sup> to study Gaussian beam propagation in the turbulent atmosphere. Gebhart<sup>75</sup> has recently used Schmeltzer's results to study the mean intensity. This study is limited, however, to points near the nominal position of the beam center. A technique suggested during the study, relying on atmospheric reciprocity, transforms the calculation of the mean intensity for the case of a uniformly illuminated transmitter aperture into the calculation of the mean signal power associated with an equivalent heterodyne receiver<sup>76</sup> (see Appendices G and H).

The ratio  $G(\vec{r})$  of the average antenna gain in the atmosphere to that with no atmosphere (the normalized mean intensity profile) can be written in terms of the mean and variance of  $\chi(\vec{r})$ , the normalized log amplitude of Eq. (8), as

$$G(\vec{r}) = \frac{\bar{G}_{\text{atmos}}(\vec{r})}{G_{\text{ideal}}(\vec{r})} = \exp\left[2\left(\sigma_{\chi}^2(\vec{r}) + \bar{\chi}(\vec{r})\right)\right]. \quad (12)$$

The result depends on the assumption that  $\chi$  is a Gaussian random variable. Even granting this assumption, the available numerical results for  $\bar{\chi}(\vec{r})$  are only for horizontal (homogeneous) paths and  $\vec{r}$  approximately at the center of the beam.<sup>74, 75</sup> Better data on the variation of the atmospheric structure constant,  $C_N^2$ , with altitude are needed before it is worth while to try to extend these numerical techniques to ground-to-space paths. Even more fundamental, the  $\bar{\chi}$  calculation that is being discussed here applies only to Gaussian beam shapes and is not easily extended.

A study of the average beamwidth for collimated Gaussian beam propagation, which is valid in the near field of the transmitter aperture, has been carried out.<sup>73</sup> The predicted beam radius is

$$R_{\text{eff}} = \text{constant } C_N \ell_o^{1/3} (ka \ln \theta)^{1/2} L \quad (13)$$

where

$$k = \frac{2\pi}{\lambda}$$

$$\theta^2 = \frac{2L^2 N_m^2}{k^2 a^2}$$

$$N_m = 5.91 \ell_o^{-1}$$

L = path length

a/2 = standard deviation of the Gaussian intensity distribution

$\ell_o$  = inner scale of turbulence.

Note that  $R_{\text{eff}}$  grows slightly more rapidly than linearly with path length, because of the  $(\ln \theta)^{1/2}$  term.

There are only a few experimental measurements associated with average beam spread and mean intensity. The separation of beam motion into beam wander and beam spread and the separation of total intensity fluctuations into variations of the mean and deviations from the mean are obviously arbitrary. The separation has normally been accomplished in experimental situations by an arbitrary division according to time scale. Ochs,<sup>77</sup> at ESSA Research Laboratory, has measured beam wander, beam spread, and mean intensity over a 15-km path. An array of 4 photomultipliers was used to provide beam-wander information to the transmitter. An integration time of 0.5 sec is used in determining beam wander. Every 15 min the vertical-beam profile (beam spread) is measured by sweeping the beam past a fifth photomultiplier at a rate of 1.9  $\mu\text{rad}/\text{sec}$  and recording the resulting intensity with a 0.5-sec time constant. The beam spread is measured as the time interval between the 3 dB points divided by the angular sweep velocity. Finally, the peak average intensity is recorded. The results of these measurements are shown in Fig. 3. The values are seen to depend heavily on the time of day. Also, the mean intensity during daylight hours does not vary as the reciprocal of the square of the beamwidth. This may be due to increased extinction by haze, asymmetry of the beam, or hot spots (concentrations of energy within the beam). It would be of value to know the actual cause.

Other beam-pattern experiments have been conducted at ESSA Research Laboratory over horizontal terrestrial paths.<sup>78</sup> These experiments at the HeNe wavelength include both fast- and long-exposure photographs of the received beam pattern over paths varying from 550 m to 145 km. The long-exposure photographs show that the atmosphere limits the extent to which a beam can be focused and causes the half-power beamwidth to grow faster than linearly with path length. In particular, transmitter apertures larger than 11 cm caused no further focusing effect, and the beam diameter followed the relation

$$D = \alpha L^\beta, \tag{14}$$

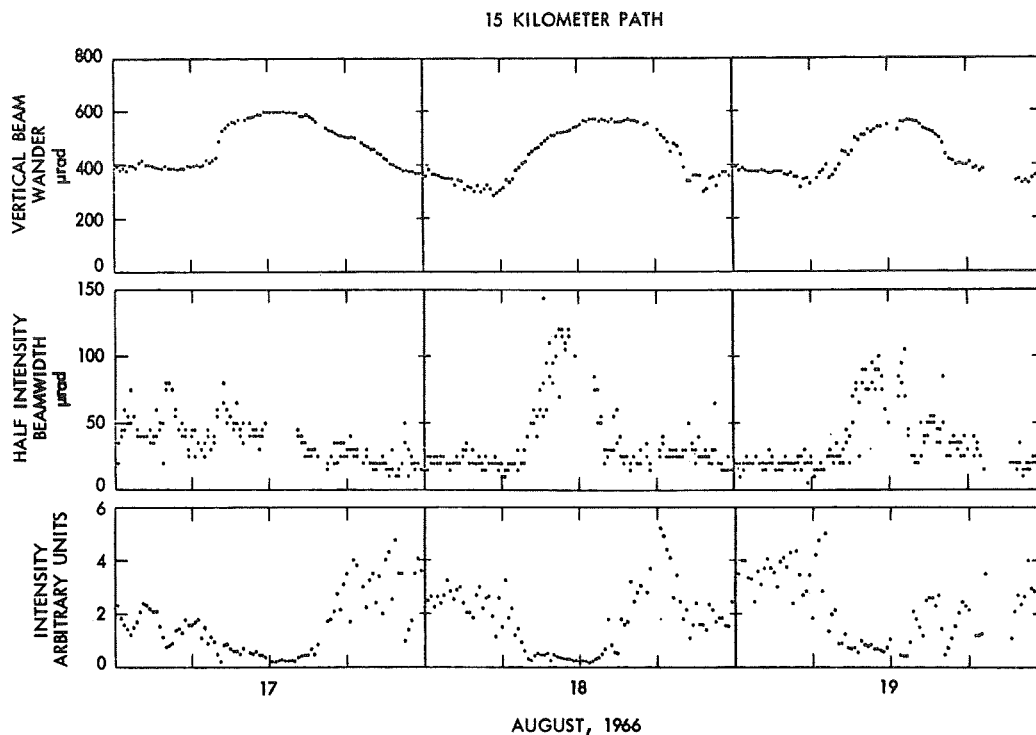


Fig. 3. Measurements of beam wander, beam spread, and mean intensity for a 15-km path. (Measured by Ochs.<sup>77</sup>)

where  $\alpha$  and  $\beta$  are  $4.5 \times 10^{-6}$  and 1.2, respectively. The results here are in rough agreement with the predictions<sup>73</sup> of Eq. (13).

The only average antenna gain data available for laser transmission over a ground-space path is that from the Surveyor VII laser pointing experiment. These results seem to imply that the beam is broadened by a factor of two or three by atmospheric turbulence, but this is at best a crude estimate.

To the extent that the atmosphere is reciprocal, it is reasonable to expect that the average size of star images observed with a terrestrial telescope can be used to estimate the average spreading of a beam transmitted from the Earth along the same slant path. The image size in the focal plane of the telescope is a measure of the angular spectrum required to represent the distorted wavefront of the stellar energy at the telescope objective. This fact, which will be considered again when phase fluctuations are considered, seems to imply that a reasonable estimate of average beam spread can be obtained by convolving the transmitter beamwidth with the average stellar angular spectrum. Image diameters of stars observed with large telescopes are usually within the range of 5-20  $\mu$ rad at night. There are no corresponding data for the daytime hours.

It is clear that more experimental data on average beam spreading and mean

intensity for ground-to-space paths are needed. A possibility for obtaining more data is to use the GEOS-II Argon detector that is now in orbit. Because of the pointing accuracy that is possible, the frequency limitation of the detector at 50 Hz, and the low orbit with the resulting limited observation time, the following kind of experiment looks attractive. An aperture mask would be used to form a pattern of narrow fan beams. These beams would then be moved through space at such a rate that the satellite passes through the pattern slowly, thereby causing the detector to scan the beam pattern. The measurements of the beam pattern and mean intensity at the center of the fan beams could then be compared with the theoretical values in the absence of turbulence. If possible, simultaneous measurements should be made of the angular spread of stars in the vicinity of the satellite to test the hypothesis that beam-spread information can be obtained from stellar observations.

Intensity Variance: Note that with a solution of the form of Eq. (7) the turbulence effects on the field envelope (or intensity) are multiplicative, in the sense that the unperturbed field envelope is multiplied by  $e^{\chi(\vec{r}, t)}$ , where  $\chi(r, t)$  represents the effect of turbulence. Under the assumption that the Rytov approximation is valid, the variance of  $\chi$  has been calculated in the plane-wave, spherical-wave, and Gaussian-beam cases. In the plane-wave case the result<sup>79</sup> is

$$\sigma_{\chi}^2 = 0.56 k^{7/6} \int_0^z C_N^2(s) s^{5/6} ds, \quad (15)$$

while the result of the spherical-wave or collimated Gaussian-beam cases<sup>80</sup> is

$$\sigma_{\chi}^2 = \frac{8.16}{8\pi} \Gamma(-5/6) k^2 \operatorname{Re} \left[ \int_0^z ds C_N^2(s) \left\{ \left[ -\frac{1}{2} \operatorname{Re} g(z, s) \right]^{5/6} - \left[ -\frac{1}{2} g(z, s) \right]^{5/6} \right\} \right], \quad (16a)$$

where  $C_N^2(s)$  is the refractive-index structure constant defined in Eq. (10),  $z$  is the path length,  $a_0$  is the standard deviation of the amplitude of the Gaussian beam, and

$$g(z, s) = \frac{2(z-s)}{ik} \frac{s - ik a_0^2}{z - ik a_0^2}. \quad (16b)$$

The integrations are along the optical path. The result is applicable at positions near the beam center and reduces to the spherical-wave case when  $a_0 \rightarrow 0$ . Indeed, for Gaussian-shaped laser-beam propagation over path lengths such that  $ka_0^2/z \ll 1$ , the resulting variance is in good agreement with the spherical-wave result, which can be written for a homogeneous path as<sup>80</sup>

$$\sigma_{\chi}^2 = 0.124 k^{7/6} z^{11/6} C_N^2. \quad (17)$$

These equations can be used to calculate answers, once  $C_N^2(s)$  is specified along the optical path. These calculations have been carried out for specific choices of  $C_N^2(s)$ ; in particular, results are available for a homogeneous path,<sup>79, 80</sup> and for a particular choice of atmospheric vertical profile.<sup>81, 82</sup> The numerical values of the atmospheric profile results are open to question, because of the lack of a verified atmospheric model and recent changes in the postulated model.<sup>64</sup> In particular, the  $k^{7/6}$  dependence for plane- and spherical-wave propagation,<sup>83, 84</sup> and the  $z^{11/6}$  dependence for plane- and spherical-wave propagation over short enough homogeneous paths<sup>83</sup> seem to be confirmed. The wavelength dependence, of course, is important in frequency selection.

Experimental results over very long and/or very turbulent atmospheric paths point out the inadequacy of the Rytov approximation results for these cases. Several experimenters have noticed that when the turbulence is severe enough and the path sufficiently long, the log-amplitude variance saturates and remains constant as the path length is increased.<sup>83, 85, 86</sup> Tatarski has proposed a modification of the theory to account for this effect.<sup>56</sup> While it is important to better understand this saturation effect and its implications, it is unlikely to occur on any reasonable optical path associated with a ground-space link. To obtain an expression for the intensity variance from the log-amplitude variance, it is necessary to use the log-normal statistic assumption and to know the mean of  $\chi$ . The result is

$$\sigma_I^2 = \bar{I}^2 \left[ e^{4\sigma_\chi^2} - 1 \right] \quad (18)$$

$$\bar{I} = I_0 e^{2 \left[ \bar{\chi} + \sigma_\chi^2 \right]}, \quad (19)$$

with  $I_0$  the intensity in the absence of turbulent effects. The value of  $\bar{\chi}$  has already been discussed. As indicated, for plane- or spherical-wave propagation a conservation-of-energy argument can be used to argue that  $\bar{I} = I_0$ , with the result that  $\bar{\chi} = -\sigma_\chi^2$ .

Stellar measurements have provided some experimental data on the intensity variance at visible wavelengths during the night. In particular, measured values of  $\sigma_I/\bar{I}$  (sometimes called the coefficient of intensity variance) for starlight are approximately 0.4 for small apertures and zenith paths.<sup>48</sup> This corresponds to  $\sigma_\chi$  of approximately 0.2. Intensity scintillation measurements have been made, using a detector on GEOS-II with an Argon laser on the ground. The transmitter has a beam divergence of 0.7 mrad, and the detector aperture has a 1-cm diameter. The value of  $\sigma_I/\bar{I}$  measured over a 160-Hz bandwidth was 0.3, which is roughly in agreement with stellar observations.<sup>84</sup>

The intensity (or log-amplitude) variance depends on both the transmitting and receiving apertures. The variance expressions discussed above were for receiving apertures that are small relative to the intensity correlation distance. As the receiving aperture becomes larger, it includes several independent intensity values and

a law of large numbers argument can be used to show that the variance decreases. For the case of plane-wave propagation along either a horizontal homogeneous path or a path from space to earth using a particular atmospheric model the amount of the reduction has been calculated.<sup>71</sup> The results show that for reasonable log-amplitude variance values ( $<0.6$ ), the intensity variance decreases by a factor equal to the coherence area divided by the receiver area, where the coherence area has a diameter approximately equal to the coherence diameter defined in the next section.

There are a lot of experimental data on the aperture averaging effect for stellar sources.<sup>48, 87</sup> These data show reasonable agreement with theory. Experiments also have been carried out with laser sources and terrestrial paths.<sup>62</sup> These measurements seem to indicate that a beam source has different correlation properties from those attributed to a plane wave and that the received intensity may have larger correlation distances and hence less aperture averaging. These correlation properties will be discussed.

The variation of intensity variance with transmitter aperture is more complicated. Here, we are concerned with beam transmission and a very small (relative to the correlation distance) receiver aperture. It appears that the intensity variance will depend on the aperture illumination, as well as on its size. In the case of a uniformly illuminated aperture the assumption of atmospheric reciprocity permits the calculation of the received-field variance in terms of the output of a heterodyne detector whose aperture area is the same as that of the transmitting aperture area. The variance of the output of a heterodyne detector has been studied.<sup>88</sup> In the case of a Gaussian-shaped aperture illumination Schmeltzer's<sup>74</sup> beam-propagation results have been used to examine the effect of beam size on the intensity scintillation. This has been done for focused and collimated beams on homogeneous paths and for collimated beams on an Earth-to-Space path.<sup>80, 81</sup> The results tend to show that the resulting log-normal variance can be written as a product of the spherical-wave variance and an aperture-dependent factor that is a function of  $ka_0^2/z$ . For a homogeneous path, the aperture-dependent factor does not vary too much for collimated beams, but decreases rapidly for focused beams when  $ka_0^2/z$  becomes larger than one. For the vertical transmission path and a collimated beam, the aperture-dependent factor decreases with increasing values of  $ka_0^2/z$ , and for large values goes approximately as the inverse of the aperture area.

There appear to be no detailed experimental measurements of the transmitter aperture effect. The theoretical variation of scintillation for a collimated Gaussian beam and a vertical path suggest that an interesting experiment would be to vary the size of the beam that is used with a ground-based transmitter and monitor the scintillations at a spacecraft-borne receiver.

Intensity Correlation: As we have indicated, the received field, after propagation through the turbulent atmosphere, exhibits a spatially varying intensity. A particularly important parameter of this variation is its correlation distance, the distance over which the field envelope is correlated. It is possible to define this parameter in many



essentially equivalent ways; here, we shall define it as the value of separation for which the spatial covariance function of the envelope goes to zero for the first time.

Again, most of the theoretical results are in terms of the log-amplitude of the field. The assumption that this quantity is a Gaussian process means that lack of correlation implies statistical independence. This is a particularly important feature that will be exploited in the design of some of the receivers discussed in section 4.3.

The spatial covariance function of the normalized amplitude,  $C_{\chi}(\rho)$ , has been studied by using the Rytov approximation. Note that in general the spatial covariance function would have to be written

$$C_{\chi}(\vec{r}_1, \vec{r}_2) = \overline{[\chi(\vec{r}_1) - \overline{\chi(\vec{r}_1)}][\chi(\vec{r}_2) - \overline{\chi(\vec{r}_2)}]} \quad (20)$$

For many situations of interest, however, the covariance across the receiver plane can be written in terms of the single scalar

$$\rho = |\vec{r}_2 - \vec{r}_1|.$$

The intensity covariance function can be written in terms of  $C_{\chi}(\rho)$  by using the log-normal property. It has the form

$$C_I(\rho) = \bar{I}^2 \{ \exp[4C_{\chi}(\rho)] - 1 \} \quad (21)$$

when the field is homogeneous over the set of  $\rho$  values of interest. Note that, while the intensity and log-amplitude correlation distances are equal, the curves have greatly different shapes, with the shape of the intensity curve dependent on the amount of turbulence present, that is, the log-amplitude variance.

In the case of plane- and spherical-wave propagation over homogeneous paths, it is possible to write  $C_{\chi}(\rho)$  in the form

$$C_{\chi}(\rho) = \sigma_{\chi}^2 f(\rho/(4z/k)^{1/2}), \quad (22)$$

where  $f$  is a function that depends on the propagation mode but is insensitive to the amount of turbulence. The first zero of  $f$  is defined as the correlation distance,  $L_{\chi}$ , and is approximately  $0.76\sqrt{\lambda L}$  for the plane-wave case, and  $1.8\sqrt{\lambda L}$  for the spherical-wave case.<sup>81, 82, 89</sup> Here  $L$  is the propagation path length.

In the beam case, an expression like Eq. (22) is no longer sufficient. A rather complicated integral expression for  $C_{\chi}(\rho)$  is available for the case of Gaussian-beam propagation (and hence spherical-wave propagation in the limit of zero source width) over any path.<sup>74</sup> While this expression is difficult to evaluate, it has been suggested that an expression of the functional form

$$C_{\chi}(\rho) = \sigma_{\chi}^2 g_1(\rho/(4z/k)^{1/2}, \rho/(2z/ka_o)) = \sigma_{\chi}^2 g_2(\rho/(4z/k)^{1/2}, \Omega), \quad (23)$$

where  $\Omega = ka_0^2/z$ , may be appropriate in this case.<sup>82</sup> Note that this expression involves the additional scale length provided by aperture diffraction effects.

Experimental measurements over terrestrial paths tend to confirm these results.<sup>50, 85, 90</sup> When  $\Omega \gg 1$ , even the beam propagation results are in reasonable agreement with the plane- and spherical-wave theory.<sup>68</sup> When  $\Omega$  becomes smaller and less than one, however, there seems to be the possibility of correlation over larger distances of the order of  $2z/ka_0$ . There is some experimental support for this in terms of studies of the intensity variation with aperture diameter.<sup>62</sup> The dependence of the correlation distance on the source size is an important consideration for terrestrial optical communication systems. For a ground-space link it is probably less important because the satellite receiver will normally intercept energy over an area that contains at most a few, and normally only one, coherence areas, that is, areas over which the intensity is essentially constant. Also, the value of  $\Omega$  associated with the down link will normally be large. Neither theoretical nor experimental results are available for the log-amplitude covariance functions on paths such that the intensity has saturated.

If the turbulence is not homogeneous along the propagation path, the results are changed. If the variation is a smoothly decreasing (or increasing) function along the path, as in some models for the variation of  $C_N^2$  with altitude, then the major modification to the results will be to replace the path length with a scale height associated with the variation of  $C_N^2$ . Thus, for the assumption of an exponential variation of  $C_N^2$  with altitude, the correlation distance depends on the parameter  $(4h_0 \sec \theta/k)^{1/2}$ , where  $h_0$  is the scale height of the atmosphere, and  $\theta$  is the zenith angle of the path.<sup>80, 82</sup> On the other hand, anomalies in the path will tend to further distort the covariance function from its homogeneous path form. For spherical-wave or Gaussian beam propagation, the effects of various turbulence regions in changing the shape of the covariance function depend on their position in the path. It has been suggested that this dependence of  $C_\chi(\rho)$  on the turbulence profile along the path be used for indirect measurements of  $C_N^2$  along the path.<sup>91</sup> This would involve measuring  $C_\chi(\rho)$  and then trying to invert an appropriately modified form of an equation of Schmeltzer.<sup>74</sup> Further investigation would be necessary to establish the feasibility of carrying out this inversion procedure on this equation. Such a procedure using stellar or laser sources would have value in monitoring the amount of turbulence at the tropopause, and might even have some value in the detection of clear air turbulence if it were sensitive enough.

Measurements along inhomogeneous terrestrial paths have shown the effects described above. Measurements between two mountain peaks, which emphasized the turbulence at the two ends of the path, showed correlation distances corresponding to the positions of the two effective turbulence regions.<sup>85</sup> About the only data available on spatial intensity correlation for ground-space paths are stellar data. Both direct and indirect measurements seem to indicate that the correlation distance is between 7 cm and 10 cm for a zenith path.<sup>87, 92</sup> Indirect results tend to indicate that the correlation distance is proportional to the square root of the secant of the zenith angle.

Frequency Spectrum of Scintillation: If the frozen-atmosphere assumption that was discussed at the beginning of section 4.1.1 is assumed valid, the spectrum of the log-amplitude (or intensity) fluctuations can be obtained from the spatial covariance function. Indeed, a joint time-space covariance function can be defined as

$$C_{\chi}(\rho, \tau) \triangleq \overline{[\chi(\vec{r}_1, t_1) - \bar{\chi}(\vec{r}_1, t_1)][\chi(\vec{r}_2 + \vec{\rho}, t_1 + \tau) - \bar{\chi}(\vec{r}_2 + \vec{\rho}, t_1 + \tau)]} = C_{\chi} \left( \rho + \frac{\vec{v} \cdot \vec{\rho}}{\rho} \tau \right), \quad (24)$$

where  $\vec{v}$  is the wind velocity vector normal to the optical path. The spectrum of  $\chi$  is obtained by setting  $\rho$  equal to zero in Eq. (24), that is, making the argument  $v\tau$ , and then Fourier transforming. This has been carried out by Tatarski for plane-wave propagation over homogeneous paths and he has obtained a normalized form that is only a function of the variable  $f\sqrt{\lambda L/v}$ , where  $v$  is the normal wind velocity.<sup>50</sup> This result has been tested experimentally by Tatarski and Ryznar.<sup>50, 93</sup>

Intensity scintillation spectra for laser sources and terrestrial paths have been measured by many people.<sup>94, 95, 68, 93</sup> These measurements show that almost all of the power is confined to the first few hundred Hertz. These spectra seem to be of the form  $S_{\chi}(f) = A e^{-af}$ , where  $a$  depends on the atmospheric conditions, for example, wind speed. The spectrum seems to be independent of receiving aperture for collimated beam propagation.<sup>95</sup>

Many scintillation spectrum measurements have been made for stellar sources.<sup>48, 87</sup> In general, for stars near zenith and with small apertures, the spectrum exhibits a break point at approximately 100 Hz and decreases to essentially zero by 500-1000 Hz. The spectra are receiving aperture sensitive, with the high frequencies decreased as the aperture increases.<sup>87, 96</sup> This is due to more averaging of the rapid spatial changes that are responsible for the higher frequencies.

Spectral measurements made with a detector on GEOS-II have shown a peak in the spectrum near 100 Hz. This may be caused by the relative spacecraft-ground motion acting as a pseudowind. The measurements in general confirm the lowpass nature of the scintillation spectrum.

#### 4.1.3 Phase Fluctuations

The phase-front distortions caused by atmospheric turbulence are responsible for the stellar image dancing and distortion that are so familiar to astronomers. The first of these effects is due to changes in the angle of arrival of the received energy, a tilt of the phase front, and the second is due to phase-front distortions that are more complicated than a simple linear tilt. These random spatial variations of the phase are often the limiting factor in system performance. This is true for imaging systems and also for some field measurement communication systems<sup>97, 98, 63, 76</sup> (see section 4.3.3).

The purpose now is to examine the characterization of this phase-front distortion. In particular, a characterization that is an adequate description for the design and

analysis of optical communication systems is required. We have shown in sections 4.1.1 and 4.1.2 that for any instant of time this corresponds to characterizing the spatial random process  $\phi(\vec{r})$  defined by Eqs. (7) and (8), and in general we are concerned with the time-space process  $\phi(\vec{r}, t)$ . For most communication system applications it is the variation of  $\phi(\vec{r}, t)$  with  $\vec{r}$  that is the important effect. Thus the major emphasis will be on the spatial behavior of the phase process.

Statistics: The same Central Limit theorem argument that is used to imply that the set of  $\chi(\vec{r})$ 's associated with any set of  $\vec{r}$ 's are jointly Gaussian random variables (see section 4.1.2) also implies that the corresponding set of  $\phi(\vec{r})$ 's are jointly Gaussian random variables. Again, this result is often extrapolated to claim that  $\phi(\vec{r}, t)$ , the phase fluctuation process, is a Gaussian random process of time and space. Thus, the process is characterized by its mean and covariance functions.

The theoretical form of the phase statistics associated with the field collected over a large aperture is unknown. It is perhaps a reasonable conjecture that the resulting temporal phase process is a Gaussian process. In many communication applications the quantity that appears is  $\phi(\vec{r}_1, t) - \phi(\vec{r}_2, t)$ , the phase difference between two points. If  $\phi(\vec{r}, t)$  is a Gaussian process, then this difference is also a Gaussian process.

The measurement of the absolute phase variation, a direct measurement of  $\phi(\vec{r}, t)$  relative to a known reference phase, requires the use of a stable transmitting laser and a field detector, for example, a heterodyne detector with a very stable local oscillator. Such measurements can be used to study the statistics, mean, variance, and spectrum of  $\phi(\vec{r}, t)$  at a single point (actually over an aperture that is small relative to a phase coherence area) or these same quantities for larger apertures. Only a few measurements of this kind are available and they have been somewhat limited by the available laser stabilities. Because of the difficulties of holding two oscillators stable relative to one another and even the stabilities that can be easily achieved with a single laser, the same laser has been used for both the source and local oscillator and relatively short two-way paths.<sup>99-101</sup> The results thus far are mostly qualitative. No measurements of the probability density of  $\phi(\vec{r}, t)$  for either small or large apertures are now available; again, this is due in part to source stability limitations. Thus there has been no direct experimental verification of the hypothesis that the phase process is Gaussian, and there are no measurements to show what happens to the phase statistics as the receiving aperture is increased.

Measurements of the statistics of the phase difference between two points in the receiving aperture have been made.<sup>102</sup> These measurements do not present the laser stability problems discussed above for the case of absolute phase measurements. These measures were made over a 4.5-km path with an He-Ne laser. They show that the phase differences have a first-order Gaussian density.

Mean Function and Variance: The mean function is a measure of the changes in the average optical path length. In most communication situations the mean of  $\phi(\vec{r}, t)$  is not

a function of  $\vec{r}$  for points in the receiving plane over the range of  $\vec{r}$  of interest. Also the time variation of the mean function is normally slow relative to signaling times and to some extent systematic. Thus  $\overline{\phi(\vec{r}, t)}$  is often unimportant in the design or analysis of communication systems.

An integral expression is available for the mean and variance of the phase function for the case of Gaussian beam propagation.<sup>74</sup> This, of course, also provides a solution for the spherical wave case in the limit of the transmitting aperture shrinking to zero. The evaluation of this integral presents the same difficulties as the evaluation of  $\overline{\chi}$ . Also, the turbulence description of Eq. (10) must be modified before it can be used in these calculations; that is, a suitable description for large and small values of  $\rho$  must be specified. One expects the phase variance to vary inversely as the square of the wavelength and linearly with the quantity  $\int_0^L C_n^2(\vec{r}) dx$ , where the integral is along the path.

Experimental measurements of the phase mean and variance present the same difficulties as the measurement of the absolute phase statistics. At the present time, few measurements of the phase mean are available. The absolute phase data that are available often show a phase drift corresponding to either path-length changes or oscillator instabilities. Even when this drift is removed from the data, short-term (few seconds) measurements indicate a variance of many radians.

Phase Structure Function and Coherence Area: The spatial variation of the phase can be described in terms of the phase-structure function

$$D_\phi(\vec{r}, \vec{r}+\vec{\rho}) = \overline{|\phi(\vec{r}) - \phi(\vec{r}+\vec{\rho})|^2}. \quad (25)$$

This is a particularly convenient description when the refractive index variations are described by Eq. (20). Often the structure function is only a function of  $\vec{\rho} = |\vec{\rho}|$  and can be written as  $D_\phi(\rho)$ .<sup>\*</sup> Note that  $D_\phi(\rho)$  grows as the correlation between the phase values decreases, and when  $D_\phi(\rho)$  is equal to twice the variance, the phases are uncorrelated and with the Gaussian assumption independent.

The phase-structure function has been calculated for plane-wave propagation. In this case the result<sup>79</sup> is

$$D_\phi(\rho) = Ak^2 \rho^{5/3} \int_0^L C_N^2(\vec{r}) dx, \quad (26)$$

---

\*When the phase process is homogeneous in the receiving plane, that is, the mean and variance are not functions of  $\vec{r}$ , at least in the region of interest, the structure function and covariance function are related by

$$D_\phi(\rho) = 2[C_\phi(0) - C_\phi(\rho)],$$

where

$$C_\phi(\rho) = \overline{[\phi(\vec{r}) - \phi(\vec{r})][\phi(\vec{r}+\vec{\rho}) - \phi(\vec{r}+\vec{\rho})]}.$$

where the integration is along the propagation path, and  $A$  is 1.46 for  $\ell_0 \ll \rho \ll \sqrt{\lambda L}$  and 2.91 for  $\sqrt{\lambda L} \ll \rho \ll L_0$ . The parameters  $\ell_0$  and  $L_0$  are associated with the turbulence and are of the order of millimeters to centimeters and meters, respectively. The path length is  $L$ , and  $k = 2\pi/\lambda$ . A series expression for  $D_\phi(\rho)$ , which is good over the whole  $\rho$  range, also exists.<sup>82</sup> The results for homogeneous paths and for a particular atmospheric profile of  $C_N^2$  variation are available in graphical form. They are of the same general form as Eq. (26), except that the exponent is larger for small values of  $\rho/\sqrt{4L/k}$ .

In the spherical-wave case, a series expression for  $D_\phi(\rho)$  has been calculated for homogeneous paths.<sup>89</sup> Again the results are available in graphical form as a function of  $\rho/\sqrt{4L/k}$ .  $D_\phi(\rho)$  has a functional form similar to the plane-wave result.

For Gaussian beam propagation the phase-structure function results are more complex. While an integral expression that is valid for most Gaussian beam situations exists, it has not been evaluated.<sup>74</sup> Tatarski has published a result that is applicable in the near field of the transmitter.<sup>73</sup> The expression has been evaluated for a homogeneous path. In this case if  $\vec{r}$  and  $\vec{r} + \vec{\rho}$  are assumed to lie on a line passing through the center of the beam, the expression for  $D_\phi(\vec{r}, \vec{r} + \vec{\rho})$  can be written

$$D_\phi(\rho, R) = 0.46 k^3 a C_{N_m}^2 \rho^{-2/3} \ln 2 \theta^2 \left[ {}_2F_1\left(\frac{1}{6}, 1, \frac{2R^2}{a^2}\right) - {}_1F_1\left(\frac{1}{6}, 2, \frac{2R^2}{a^2}\right) \right], \quad (27)$$

where

$$k = \frac{2\pi}{\lambda}$$

$$n_m = 5.91 \ell_0^{-1}$$

$$\theta^2 = \frac{2L^2 n_m^2}{k^2 a^2}$$

$$g = \frac{2L}{ka^2}$$

$L \triangleq$  path length

$a/2 \triangleq$  standard deviation of Gaussian intensity distribution

$\ell_0 \triangleq$  inner scale of turbulence<sup>57</sup>

$$R = |\vec{R}| = \left| \vec{r} + \frac{\vec{\rho}}{2} \right|$$

$$\rho = |\vec{\rho}|.$$

This relation is valid for  $\sqrt{\lambda L} \gg \ell_0$ ,  $g \ll 1$ , and  $\theta \gg 1$ . Note that it predicts a  $\rho^2$  variation

rather than the  $\rho^{5/3}$  variation associated with plane- and spherical-wave propagation. Also the path-length and wavelength dependence differs from that for the plane- and spherical-wave cases.

The correlation function between phase variations at the beam center and at a point a distance  $\rho$  from the center has also been studied under the assumption that the refractive-index variations can be specified in terms of a Gaussian-shaped correlation function.<sup>103</sup> This work yields explicit results, but their applicability to the atmosphere is open to question because of the choice of refractive-index correlation function.

For very turbulent situations (or very long paths) it has been suggested that the phase-structure function approaches a constant limit independent of the path length; that is, it no longer increases with increasing path length.<sup>56</sup> The calculation was carried out for plane-wave propagation, and the  $\rho$  and  $k$  dependence was not affected.

Perhaps the most important single parameter associated with the spatial variation of the phase is the coherence area (or coherence diameter). This is the region of the receiving plane over which the variations of  $\phi$  are within some specified limits. If the validity of Eq. (26) is assumed, the coherence diameter would be defined as

$$D_{\text{coh}} = \left[ \frac{\Delta\phi^2}{Ak^2 \int_0^L C_N^2(\vec{r}) dx} \right]^{3/5}, \quad (28)$$

where  $\Delta\phi$  is the allowed rms phase variation over the area. Often the phase coherence is specified in terms of the parameter ( $\Delta\phi^2 \approx 6.88$ )

$$D_{\text{eff}} = \left[ \frac{0.0588 \lambda^2}{\int_0^L C_N^2(\vec{r}) dx} \right]^{3/5} \quad (29)$$

which plays a fundamental role in describing the performance of heterodyne communication systems<sup>63, 76</sup> (see section 4.3.2). Values of this parameter (actually an essentially equivalent parameter) have been calculated for a particular atmospheric model.<sup>76</sup> For visible wavelengths, the values of  $D_{\text{eff}}$  for a space-to-earth path are a few centimeters, while they are approximately thirty times larger at 10.6  $\mu\text{m}$ .

Few direct measurements of  $D_\phi(\rho)$  or the coherence diameter have been carried out. One set of measurements over a 4.5-km horizontal path with a collimated 2-in. diameter He-Ne source have shown that the phase is quite coherent over apertures of 0.5 in. with only a small amount of linear phase-front tilt.<sup>102</sup> One-inch apertures, however, showed higher order phase disturbances. Measurements of the phase difference between two small apertures for separations of 2-18 in. showed little dependence on the separation. These results would lead one to conclude that the phase variations are already essentially uncorrelated for two 2-in. separations. These measurements also indicate a standard deviation of approximately 8.5 rad for the phase difference. Indirect

measurements of  $D_{\text{eff}}$  have been carried out by studying heterodyne receivers operating in the atmosphere.<sup>63</sup> These results seem to be in general agreement with those that would be predicted by Eq. (26).

Phase Front Representation: The form of the spatial variation of the phase has been studied both theoretically and experimentally. One representation for the entire received process, including both the amplitude and phase variations, is in terms of a plane-wave decomposition, that is, an angular spectrum. (This is discussed in section 4.2.)

Decomposition of the phase-front variations into various geometrical shapes has also been considered.<sup>104</sup> The combination of these shapes (linear tilt, spherical deformation, hyperbolic deformation) in essence form the first few terms of a two-dimensional Taylor series expansion of  $\phi(\vec{r})$ . This study showed that if the phase-structure function had the form of Eq. (26), the linear phase-front tilt would be the most significant form of distortion for small apertures. Indeed, the value of  $D_{\text{eff}}$  can be increased by a factor of 3.4 by using phase-front tracking to remove the effects of this linear phase variation (see section 4.3.2). Corrections for the higher order distortion terms (spherical and hyperbolic) were not nearly as effective in terms of compensating for the phase distortion.

The power pattern of the energy collected over a circular aperture has been considered for both tracking and nontracking operation, under the assumption that Eq. (26) is valid.<sup>105</sup> This study determined the average power pattern and compared the beamwidths and gains for tracking and nontracking apertures. The tracking beamwidth is approximately 30% narrower for apertures such that the rms phase difference between the center and the edge of the aperture is a few radians.

Several experimental studies of the phase-front distortion over various apertures have been carried out by using interferometry.<sup>101, 102, 106, 107</sup> Most of these experiments have been over relatively short paths, although one experiment, which involved the use of a Fresnel biprism to form an interference pattern between the field collected over two parts of the aperture, was carried out over a 4.5-km path.<sup>102</sup> The results of these experiments are largely qualitative. They do tend to confirm the importance of the linear phase-front tilt for small apertures and the importance of other phase distortions as the aperture gets larger.

Phase-Time Behavior: The "frozen atmosphere" assumption (see section 4.1.1) ties together the temporal and spatial phase changes. Thus the spatial variations of  $\phi(\vec{r})$  already considered can be used to predict the spectrum of the phase. The frequency spreading, caused by the atmosphere, of the output of a large-aperture field-measurement (heterodyne) receiver has been studied.<sup>88</sup> This result includes effects attributable to the amplitude fading, as well as the phase effects. If the phase effects tend to predominate, then this calculation is associated with the bandwidth of the spectrum of the phase process. The results show that the atmospheric frequency spreading decreases with receiving aperture.

Measurements have been made of the phase spectrum over paths up to several



kilometers.<sup>100, 101</sup> These measurements show that the phase spectrum is at most a few hundred cycles wide, being only slightly wider than the amplitude spectrum.<sup>100</sup> The temporal structure function of the phase difference between two small apertures has been measured.<sup>102</sup> The results are in agreement with the frozen-atmosphere assumption and indicate bandwidths of hundreds of Hz.

Angle of Arrival: The variation of the average angle of arrival of the energy incident on the receiving aperture is the quantity of interest here. This is obviously just the linear phase-front tilt of the field over the aperture, which has been discussed above. This quantity is sufficiently important in its own right in such applications as acquisition, phase-front tracking receivers, and adaptive transmitters that it will now be considered in more detail.

Under the assumption of plane-wave propagation, the variations of the average angle of arrival of the energy over a receiving aperture of diameter  $D$  has the mean-square value<sup>87</sup>

$$\overline{(\Delta\alpha)^2} = 1.027 AD^{-1/3} \int_0^L C_N^2(\vec{r}) dx, \quad (30)$$

where the phase-structure function is assumed to have the form of Eq. (26). Here  $\Delta\alpha$  is actually the variation of the propagation direction, from the normal to the aperture, along one of two orthogonal axes. Based on the phase statistics one would expect  $\Delta\alpha$  to be a Gaussian random process with zero mean and a spectral width similar to that for the total phase spectrum. The total variation in the angle of arrival can with good accuracy be described as composed of two orthogonal components like  $\Delta\alpha$ . It is thus a Rayleigh random process with twice the variance of  $\Delta\alpha$ . The phase-structure function for Gaussian beam propagation, which is given by Eq. (27), can be used to study the angle of arrival in the beam case. Tatarski computes the effective angle of arrival over the entire received beam, but the limitation to path lengths in the near field of the transmitter effectively restricts the receiver size to be similar to the transmitter size.<sup>73</sup> The result is

$$\overline{(\Delta\alpha)_{\text{eff}}^2} = 0.1 C_N^2 k a \ell_o^{2/3} \ln 2\theta^2, \quad (31)$$

where the parameters and conditions for validity are specified below Eq. (27). This result is different from Eq. (30) both in its path length and its wavelength dependence. In this case the variance goes as the log of the path length.

Measurements of the average angle of arrival have been made at the Marshall Space Flight Center.<sup>108, 109</sup> These measurements were made with a 6328 Å He-Ne laser over a range of 3200 m. At the transmitter, the beam diameter was 3 mm and beam divergence was 2 mrad. The angle of arrival was measured at the focal plane of a variable-aperture telescope with an image dissector angle error sensor. These measurements show that each component of the angle of arrival is essentially a Gaussian process. The

rms deviation in angle of arrival was measured as a function of receiver aperture diameter. The results were: 6.3  $\mu$ rad rms using a 5-in. receiver; 4.6  $\mu$ rad rms using a 12-in. receiver. The frequency response of the angle sensor was 150 Hz. When a spectral analysis was performed on the received fluctuations it was found that the spectral power density fell to 10% of the DC spectral power density at 10 Hz and to 1% at 100 Hz.

Measurements were also made in a double pass of the laser beam to a corner reflector and back, over 6,400 m. There was no significant difference in the power spectrum, and the rms angle deviations appeared to be approximately doubled. Note that these results are in approximate agreement with Eq. (30).

Stellar measurements can be used to study the variations in angle of arrival for a Space-to-Earth path.<sup>48, 87</sup> These results indicate that the rms variation varies linearly with the secant of the zenith angle and has a magnitude of 1 arc second for average seeing conditions. For bad seeing it may be as large as 10 arc seconds.

#### 4.1.4 Beam Time Behavior

The mean and variance of the intensity, for collimated beam propagation, have been considered at various positions in the receiving plane. The mean calculation corresponds to a calculation of the ensemble average antenna gain, which is the time average antenna gain if the process is assumed to be ergodic. The variance calculated at any point includes the effects of both beam wander, changes in the position of the center of the received beam, and beam profile changes, as well as the scintillation effects. In some situations, particularly those associated with the acquisition phase of the communication problem and the use of adaptive transmitters, it is important to consider separately the time behavior of the variations in beam center and profile. Thus important questions are the following.

1. What are the statistics associated with the beamwidth when the beamwidth is measured relative to the instantaneous beam center?
2. What are the statistics associated with the position of the center of the beam?
3. What is the time scale of the changes in beam center position and beam profile?
4. To what extent can the average antenna gain be increased and the variations in antenna gain decreased by compensating the transmitted wavefront for the distortion it will suffer (for example, by using a "self-steering" phased-array transmitter tracking a beacon)?

Few theoretical or experimental results that are applicable to these questions are available. This is due in part to the difficulty of separating the various intensity effects in both analysis and measurement. The variation of the beam center is directly related to the average direction of propagation over the beam profile. The angle-of-arrival results from section 4.1.3 can be used to study this motion at a satellite by considering the angle of arrival of the beam in a plane at the "top" of the atmosphere. Some feeling for the importance of the beam motion in causing antenna gain variation can be obtained under the assumption of instantaneous atmospheric reciprocity (see

Appendices G and H). Then the gain of a uniformly illuminated transmitter could be evaluated in terms of the intensity variations of an equivalent heterodyne receiver. One would then draw the following conclusions.

1. The gain is a log-normal process (if the heterodyne output is shown to be log-normal (see section 4.3.2)).

2. The transmitter aperture size should be limited to the diameter over which the received field is phase-coherent; that is, essentially uniform phase, if the gain fluctuations are to be small.

3. At  $10.6 \mu\text{m}$ , and possibly at  $0.5 \mu\text{m}$ , the transmitting aperture size can be increased by roughly a factor of three, without increasing the gain fluctuations, by tracking the instantaneous phase-front tilt of a beacon and locking the transmitter to this signal. (Note that the applicability of phase-front tracking of a beacon to transmitter performance depends on the point-ahead angle required and on whether the atmosphere is reciprocal for two paths with this angular separation.)

Item 2 can be interpreted to mean that beam wander is unimportant in this case and item 3 would be a simple case of transmitter adaptation to increase the antenna gain (make the beam width smaller) and still keep the effects of beam wander on gain variation small. These results, however, are far from a complete answer, in that they are only applicable to the uniform transmitter illumination case, they need experimental verification, and the interpretation is somewhat subjective.

#### 4.1.5 Areas Requiring Further Investigation

We shall list those areas of the channel characterization problem that seem to need further theoretical or experimental work and in some cases suggest experiments. The list should not be considered as complete, but rather as indicating ideas and areas that are important for optical communication work, but have not been sufficiently studied at the present time. Some ideas may have been neglected in the press of time. Also, future communication system work may point out other channel characterization ideas that need to be studied and may indicate that some of those now thought to be well understood require further work. This may be particularly true when trying to analyze some of the communication systems discussed in section 4.3. Most of the ideas mentioned here have been considered in our channel characterization discussions. The emphasis here is, in general, on the Earth-Space channel.

Theory: Further theoretical work is needed in the following areas.

##### 1. Turbulence Process

Examine further the applicability of the Rytov solution to the turbulent optical propagation problem. Specifically, effort should be concentrated on determining whether this approximation is adequate for providing an adequate statistical model for the communication situations of interest. If necessary, alternative techniques applicable to beam propagation should be developed.

Examine the variation of the turbulence structure constant with atmospheric conditions and altitude. In particular, try to determine conditions that will lead to very turbulent layers at particular altitudes.

Examine the communication implications, if any, of a more complete description of the actual three-dimensional structure of the turbulence (temperature fluctuations). The temperature fluctuations are non-Gaussian and skew.

## 2. Mean Intensity and Average Beam Divergence

Examine the average beam intensity pattern associated with uniformly illuminated apertures.

Determine the average beam divergence and mean intensity for an Earth-to-Space path for both a Gaussian beam and a uniformly illuminated aperture. Consider the implications of the results in terms of optimal energy usage. This work will require a realistic theoretical model of the variation of  $C_N^2$  with altitude.

Examine the applicability of stellar measurements in determining average beam divergence on a ground-to-space path.

## 3. Intensity

Compare the intensity variance throughout the beam profile for both Gaussian beams and uniformly illuminated apertures and examine the communication implications. Look also for any theoretical evidence of a breakdown of the log-normal model off beam center.

Examine the variation with transmitter aperture size and illumination of the intensity covariance function. This is particularly needed for uniform aperture illumination.

Study the feasibility of using intensity covariance function measurements, or some other indirect measurement, to determine the  $C_N^2$  altitude profile.

Try to obtain a better understanding of the implications of the intensity variance saturation effect with particular emphasis on determining when it will be important in communication applications.

Examine the spatial continuity of the temporal intensity variations within the beam. This may have implications for the control systems associated with tracking receivers.

## 4. Phase

Examine the dependence of the angle-of-arrival characteristics – variance, statistics, and receiving aperture dependence – on the transmitting aperture size and illumination.

Determine the dependence of the phase-structure function, particularly the phase-coherence diameter, on the transmitting aperture size and illumination.

## 5. Beam Time Behavior

Examine the magnitude, statistics, and time scale of any rapid variations in the beam antenna pattern, including beam wander.

Examine the relationship between temporal variations in the beam antenna pattern and temporal variations in the intensity at a fixed point in the receiving plane. In particular, consider the possibility of adaptive techniques at the transmitter to reduce temporal variations in the antenna pattern and potentially permit the use of increased antenna gain while keeping the intensity variations small.

Experiments: Many experiments and measurements are needed to verify and assist theoretical work. In many cases an Earth-Space path is desirable, at least ultimately, for the experiments. Ground-based experiments will be very useful, however, in establishing experimental techniques and better defining experimental goals. The experimental work should include the following areas.

### 1. Turbulence

Measure the height distribution of  $C_N^2$ , the refractive index structure constant. The measurements should include direct temperature measurements, probably with balloons, and, if feasible, the development of some remote sensing technique.

### 2. Mean Intensity and Average Beam Divergence

Make measurements of the beam spread and the variations in beam spread. Quantities of interest are the beam divergence, the statistics, magnitudes, and time scale of any variations in beam spread, and the relation between beam spread and the mean intensity at the center of the beam.

Measure the beam intensity profile and its dependence on aperture illumination.

Make daylight measurements of stellar image sizes if it can be established that the measurements can be used to predict the average beam divergence.

Measure the beam divergence on a ground-to-space path, using the Argon detector on GEOS-II. (This experiment is discussed in section 4.1.2.) Simultaneous stellar image measurements should be made to determine their applicability to the beam-divergence problem.

### 3. Intensity

Measure the distribution and variance of the intensity variations throughout the beam profile.

Measure the intensity statistics and variance for large apertures as a function of receiving aperture size and transmitter aperture size and illumination.

Obtain the statistics associated with stellar scintillation, if the required results are not available in published works.

Measure the intensity covariance function for stellar sources in the visible and infrared spectral regions.

Measure the dependence of the intensity covariance function on the transmitting aperture size and illumination.

Measure the intensity time-space covariance function to determine the applicability

of the frozen-atmosphere hypothesis, with particular emphasis upon determining the conditions under which it breaks down.

Measure the amplitude frequency coherence bandwidth. Such a measurement could be made directly with two lasers operating at different wavelengths if one can be tuned over an appreciable frequency range. Pulse measurements could also be used in a more indirect manner.

#### 4. Phase

Measure the statistical distribution and variance of the absolute phase fluctuations caused by the turbulence.

Measure the statistics of the phase-difference process associated with two small spatially separated apertures to check the Gaussian hypothesis.

Measure the phase-structure function to determine its spatial separation dependence as a function of transmitter aperture size and illumination.

Make measurements to determine whether the phase-structure function saturates with respect to its path-length dependence; that is, Does the linear path length dependence persist even for very turbulent and/or long paths? If there is saturation, measure the phase-structure function.

Measure the phase-coherence area, suitably defined, for ground-space paths.

Measure the spectrum of the phase difference between two small apertures as a function of their spatial separation. Also measure the time-space structure function to check the frozen-atmosphere assumption relative to the phase variations.

Measure the spectrum associated with the absolute phase variations caused by the turbulence.

Measure the frequency coherence bandwidth of the phase process.

#### 5. Beam Time Behavior

Measure the temporal behavior of the beam antenna pattern. Quantities of interest are the statistics, magnitudes, and time scale of variations in the beam position (beam wander, and the beam divergence or profile). The relationship between these quantities and the intensity variations at the nominal beam center is of interest.

Measure the beam wander to determine if the temporal variations are spatially continuous.

#### 6. Phase and Intensity

Measure the correlation between phase and amplitude variations. In particular, make measurements to determine under what conditions, if any  $[\chi(\vec{r}_1) + \chi(\vec{r}_2)]$  and  $[\phi(\vec{r}_1) - \phi(\vec{r}_2)]$  are independent quantities.

#### 7. Adaptive Systems

Make measurements to test the concept of atmospheric reciprocity, including its

angular dependence. The measurements should be made for various aperture sizes. The possibility of using the Argon detector on GEOS-II in conjunction with a simultaneous stellar measurement should be considered.

Make measurements of the increase in antenna gain that can be achieved by servoing the transmitter to follow the linear phase-front variations measured by a receiver tracking a beacon. In this case, the receiver and transmitter must share a common aperture.

#### 4.2 CHANNEL MODELS AND REPRESENTATIONS

We shall now attempt to unify the results of section 4.1 by considering a formal mathematical model of the signal portion of the field at a receiving plane, given a particular transmitted field. Various approximate representations will then be discussed. A statistical channel model has previously been proposed, and it still seems to be useful.<sup>110, 111</sup> Its validity, while predicated on the correctness of the log-normal statistics, is not directly dependent on the Rytov approximation.

For simplicity, assume that the transmitted field is linearly polarized and is not spatially modulated. The polarization assumption can be modified by introducing a second orthogonal polarization component. These components can be considered separately because the turbulence causes negligible depolarization. It will be seen that any time-varying spatial modulation could be introduced in the same way that the antenna pattern (a fixed form of spatial modulation) is handled, that is, by considering it as a time-varying antenna pattern.

The problem is now reduced to a scalar field problem and the transmitted field is described simply as a function of time, as if it were an infinite plane wave. It is convenient to represent this field and also the received field in terms of their complex envelopes.<sup>29</sup> That is, the transmitted and received (signal portion) fields are assumed to be of the form

$$\epsilon_t(t) = \text{Re} \left[ E_t(t) e^{j2\pi\nu t} \right] \quad (32)$$

$$\epsilon_s(t, \vec{r}) = \text{Re} \left[ E_s(t, \vec{r}) e^{j2\pi\nu t} \right], \quad (33)$$

where  $E_t(t)$  and  $E_s(t, \vec{r})$  are complex envelopes, and  $\nu$  is the optical carrier frequency. The received field is now described in terms of  $E_t(t)$  by the relation

$$E_s(t, \vec{r}) = E_t(t) e^{\gamma(t, \vec{r})}, \quad (34)$$

where

$$\gamma(t, \vec{r}) = \chi(t, \vec{r}) + j\phi(t, \vec{r}). \quad (35)$$

The random process  $\gamma(t, \vec{r})$  is assumed to be a complex Gaussian random process. It

represents the effects of the antenna pattern, as well as those of the turbulence. As a complex Gaussian process,  $\gamma(t, \vec{r})$  is completely described statistically in terms of its mean and covariance functions. The covariance function is described in terms of the covariance functions of  $\chi(t, \vec{r})$  and  $\phi(t, \vec{r})$  and the cross covariance function of  $\chi$  and  $\phi$ . The antenna pattern of the transmitted radiation obviously affects the means and covariances in a fundamental way and, even for the same turbulence conditions, they will change greatly as the actual transmitted field changes in form from a plane wave to a spherical wave or a beam.

The adequacy of the suggested model rests heavily on the assumption that the coherence bandwidth of the turbulent channel exceeds the bandwidth of  $E_t(t)$ . (In section 4.1 we indicated that this is a reasonable assumption for many signal bandwidths.) Furthermore, the log-normal process assumption is being fully exploited. This has not been adequately checked, but at the present time it seems to be the only reasonable alternative. While this model provides the desired complete statistical description of the received field, it obviously places a very heavy burden on the mean and covariance calculations. In many cases it may be very difficult to obtain analytical expressions for these functions, and hence certain detailed analyses may not be possible. On the other hand, the model does contain, in a very compact form, the relevant features of the received field; consequently, it is useful in problem formulation and in understanding the over-all character of the problem.

To complete the description of the total received field, it is necessary to describe the additive background radiation arising from natural and artificial sources. When the receiver field of view does not include the sun, most of the natural radiation in the wavelength region from the visible up to  $3 \mu\text{m}$  is due to reflected or scattered solar radiation. The noise magnitude depends on the solar altitude and receiver attitude. In the visible range on a clear day, the background noise for a receiver on the Earth with a field of view centered at the zenith is approximately  $10^{-18} \text{ W/cm}^2 \text{ - ster - Hz}$ .<sup>112</sup> In many situations the important background radiation above  $3 \mu\text{m}$  is the thermal radiation of the terrestrial environment whose spectral shape is reasonably modeled by  $\approx 280^\circ\text{K}$  black-body radiation.<sup>112,113</sup> Around  $10 \mu\text{m}$  the background level is  $10^{-16} \text{ W/cm}^2 \text{ - ster - Hz}$ . There is some variation in the background level above  $3 \mu\text{m}$  with atmospheric conditions (clear or overcast), but little dependence on the time of day. At the visible wavelengths, the amount of background is clearly time-dependent and greatly reduced at night. In some applications direct sunlight, reflections from planets or the moon, or stellar radiation may be important. Information on these sources is available.<sup>10</sup>

Normally, it will be appropriate to model the background radiation field as an unpolarized field whose components in the plane of the receiver are independent. Moreover, the complex envelope of each polarization component can be considered to be a zero-mean complex Gaussian random process with independent components that



are stationary in time and space.\* Furthermore, it is reasonable to assume that the time-space covariance function of the real and imaginary parts are identical, and that the spatial covariance function is only a few wavelengths wide. Finally, the time spectra of the real and imaginary parts can reasonably be considered to be flat over most bandwidths of interest.

At this point, the channel characteristics can be summarized. The model corresponds to a fading channel with little time or frequency dispersion. It differs from the more familiar Rayleigh fading channel, in that the amplitude variations have a log-normal rather than a Rayleigh distribution, and the phase variations have a Gaussian rather than a uniform distribution. The variance associated with the phase process may be large.

The model discussed above presents severe and largely unsolved analysis problems. Using it to obtain optimum systems is very difficult. One reason for this difficulty is the unsolved problem of obtaining a suitable mathematical representation for the sum of log-normal and Gaussian processes. Also, the very compactness of the model may tend to obscure certain features of the received process that good suboptimal communication systems might exploit. Thus, while the model is very useful, there is a great need for reasonable approximations and alternative representations that retain the relevant statistical details but increase the feasibility of both optimum analysis and the design of good suboptimum systems. This need for a sufficient understanding to make reasonable approximations is the motivation for the detailed consideration of the various turbulence effects that we have carried out in section 4.1.

The problem of formulating meaningful and useful approximations to this model or obtaining useful alternative representations is certainly not a closed issue. Several possible approaches either have been used or look promising. One approximation that can be used to reduce the problem to manageable form is to partition the receiving aperture into coherence areas and to consider the field amplitude and/or phase as constant (coherent) within each area and statistically independent from area to area. At some cost in complexity, the areas may differ for the amplitude and the phase. This reduces the problem to the consideration of a set of independent phase normal processes, that is, processes defined as the exponentiation of complex Gaussian processes. If the time variation is considered in the same way, defining the process to be constant over a coherence time and perhaps independent from coherence time to coherence time, then the problem is reduced to the consideration of a set of phase normal random variables. Such a description obviously involves a lot of approximation, but it does seem to retain the essence of the statistical description and it has proved useful in analyzing communication systems.<sup>111</sup>

---

\*In some situations, it is appropriate to consider the background field as a constant intensity. The applicability of this approach depends on the relative bandwidths of the optical predetection filter and the final electronic receiver processing.

A process model that looks potentially useful and should be given further theoretical and experimental study is a representation in terms of the angular spectrum of the received process. This might take the form of trying to represent the field over the receiving aperture in terms of a set of plane waves characterized by random direction vectors and Gaussian-shaped intensity distributions with independent amplitudes. A representation of this kind would be useful in the analysis and synthesis of systems.

An interesting aspect of the representation problem is the relation between aperture plane representations and representations in the focal plane of a lens. This obviously provides a tie between the ideas discussed above. The number of degrees of freedom should be preserved by the linear lens transformation. It appears that, at least for a large lens, the degrees of freedom, which are represented as coherence areas in the aperture plane, are represented by Airy disk-sized areas in the focal plane. The number of Airy disk-sized areas in the focal plane with significant energy would then be equal to the number of coherence areas in the aperture plane. This relationship should be examined theoretically and experimentally.

The present discussion suggests the need for additional theoretical and experimental work with respect to the channel representation problem. Theoretical work of interest should include the following investigations.

1. Study of the ergodicity question for the received field process and determination of the communication implications of non-ergodicity.

2. Study of the relationship between aperture and focal plane representations of the received field, particularly with reference to the form in which the degrees of freedom appear.

Study of the representation of the received process in terms of an angular spectrum of plane waves, perhaps with nonuniform intensities. This study should include an examination of the implications and validity of such a representation both as an exact model and as a reasonable approximation. Particular emphasis should be placed on the statistics associated with the representation.

Experimental work should include measurements to determine the applicability of an angular spectrum representation for the process and the statistics associated with such a representation. Such measurements might be made by performing simultaneous measurements of the amplitude and phase processes associated with an array of detectors appropriately located in the focal plane of a receiving lens.

#### 4.3 COMMUNICATION SYSTEMS

Here we consider communication systems for use with the turbulent atmospheric channel. The emphasis is almost completely on digital communication systems. While a fundamental quantum detection viewpoint would be pursued, ideally, it seems more fruitful at this time to concentrate on systems that use direct (energy) detection or some form of heterodyne (field) detection. The use of optical preamplifiers before the final detection operation will also be considered. Several novel communication techniques

that can be used with any detection scheme, and with analog, as well as digital, systems, are also discussed. All of the systems and techniques considered here have promise in some applications and merit further investigation.

#### 4.3.1 Direct-Detection Techniques

We are concerned here with atmospheric systems that use direct (incoherent or energy) detection. This type of detection process was discussed in section 3.2.3 in the context of the free-space channel. Since many of the results discussed there are directly applicable to the turbulent-channel case, the major emphasis here will be on those additional features that arise because of the turbulent channel.

A fairly general model for the direct-detection problem will be introduced. It will provide a framework within which to discuss various important features of the problem. Various promising receiver structures, the signal design problem, and useful experiments will then be considered.

A Model for Direct-Detection Systems: The model shown in Fig. 4 is useful in discussions of direct detection techniques for the turbulent channel. For simplicity, a scalar field, linearly polarized, has been assumed. As indicated in Fig. 4, D signals, represented by the components of the vector  $\underline{r}(t)$ , are available to be processed. Each

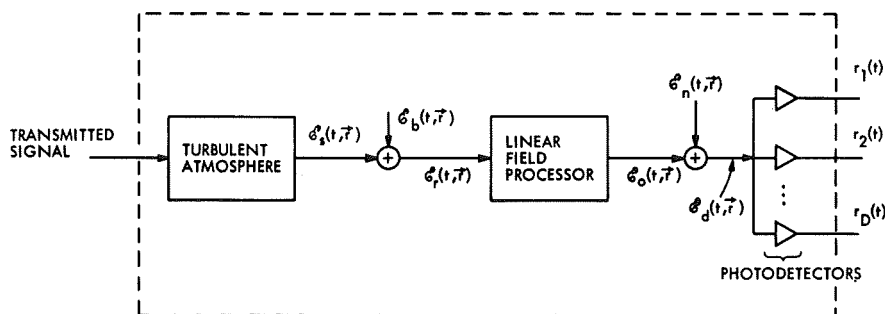


Fig. 4. General model for structured atmospheric channel receivers.

component is the sum of a conditional Poisson process, representing the photodetector output as described in section 3.1, and a signal-independent Gaussian noise process representing the noise associated with the detector and its associated circuitry, for example, the electronic preamplifier.\* The inclusion of several photodetectors allows the possibility of using separate detectors for various spatial regions or modes of the received field. The input to the system, the portion in dashed lines, is the transmitted signal. The field  $\epsilon_s(t, \vec{r})$  represents the signal portion of the field at the receiving plane

\*In general, for nonideal devices the photodetector output is a conditional filtered Poisson process and there is also an additive signal-independent shot process (dark current). (See section 3.1.1.)

after it has propagated through the turbulent channel. The relationship between the transmitted field and  $\epsilon_s(t, \vec{r})$  was discussed in section 4.2. The field  $\epsilon_b(t, \vec{r})$  represents the background noise.

The model allows for linear processing on the field before the energy detection operation; the additive noise source  $\epsilon_n(t, \vec{r})$  accounts for the possibility that this processing introduces noise. The field processing may, for example, be time filtering with an optical bandpass filter, spatial filtering with some optical system, amplification with an optical preamplifier, or some combination of these operations. This processing might be used to decompose the field into various spatial modes. The processing noise  $\epsilon_n(t, \vec{r})$  is especially needed to model adequately the optical preamplifier case (see section 3.1.2).

As discussed in section 3.1, the process  $\underline{r}(t)$  depends on the intensity (instantaneous power) of the field at the detectors,  $\epsilon_d(t, \vec{r})$ , and on the Gaussian signal-independent noise associated with electronic operations on  $r(t)$ . The direct detection technique is most attractive when the additional noise caused by the electronic operations is negligible compared with the signal-dependent noise. This is often the case when the detector has internal gain, for example, a photomultiplier, and thus with the present technology is most attractive in the visible and near infrared wavelength region. On the other hand, the situation may also occur if the linear field operation includes sufficient amplification — a technique most applicable in the longer wavelength region.

Assuming, for the present, that the detectors are noiseless (the additive Gaussian noise can be ignored), consider the noise caused by the detection operation on  $\epsilon_d(t, \vec{r})$ . Several situations are of interest. One is when the background field  $\epsilon_b(t, \vec{r})$  is sufficiently large relative to the signal field  $\epsilon_s(t, \vec{r})$  that it dominates  $\epsilon_o(t, \vec{r})$ , with  $\epsilon_n(t, \vec{r}) = 0$ . In this case the "shot noise" in  $\underline{r}(t)$  will depend essentially on the filtered version of  $\epsilon_b(t, \vec{r})$  and will be signal-independent. If the rate is high enough relative to subsequent processing time constants, it may be appropriate to consider this shot noise as a Gaussian process. Also, if the input optical filter bandwidth is large relative to the bandwidth associated with the subsequent processing, it may be appropriate to consider the background intensity as a constant. Another case is when an optical preamplifier is used. Normally, for this case,  $\epsilon_n(t, \vec{r})$  is the dominant input noise to the detectors. Finally, the signal field, especially after the linear processing, which will normally limit the field of view and contain an optical bandpass filter, may be larger than any noise field. In either of the last situations, the "shot-noise" processes  $\underline{r}(t)$  must normally be considered in detail in any analysis.

As noted in section 3.2, the forms of the optimal processors are available for known signals in the presence of a noise whose bandwidth greatly exceeds the post detection bandwidth, that is, where the background noise can be treated as a fixed intensity.<sup>114-116, 36</sup> Some performance results for the case in which the background intensity is large relative to the signal intensity are also available.<sup>116</sup> In the problem of interest here, however, the signal fields are random, and it seems that only a few special situations are understood. Before discussing specific direct-detection systems, we

first enumerate the questions that would ideally be asked. For digital signaling with a probability-of-error criterion the relevant issues, in terms of the model are these.

1. The optimum receiver structure, that is, processing on  $\underline{r}(t)$ , and the resulting performance.
2. Optimization of the linear field processing within some set of allowed operations to minimize the probability of error. For example, only passive devices may be allowed for this processing.
3. Optimization for a given receiving aperture of the number of photodetectors and of any spatial field processing performed by the detectors.
4. Optimization of the signal set with respect to the minimum probability-of-error criterion.

It is necessary, in general, to consider the joint optimization problem.

Analog communication and parameter estimation problems are also of interest. Among the important problems that should be considered are the FM communication problem and the synchronization problem (time estimation) for digital communication systems. FM systems look attractive because of the relatively small effect that the turbulence has on the frequency. Some of the more promising digital systems, such as PPM, require accurate timing, and thus good synchronization, even though one terminal may be moving relative to the other. Of course, optimizations of the type specified for digital signaling are the desired result.

Several receiver structures, suggested by the available analysis results, by the channel properties discussed in sections 4.1 and 4.2, and by previous experience with channels that have similar properties, will now be considered. In general, they will be considered in order of increasing complexity and presumably performance. As indicated above, few theoretical results are available, but the structures considered are all sufficiently promising in terms of simplicity or performance to warrant further study. Specifically, their performance and the trade-off between complexity and performance should be established. Of course, other promising implementations may also exist. The systems considered here reflect the interests and experience of those at the Workshop.

The property of the received field that most directly affects the design and performance of the receivers considered here is the turbulence-induced temporal and spatial variations in the magnitude of the complex envelope rather than its phase variations. Thus the receivers that we now consider will try to combat the effects of these envelope magnitude variations.

Single Detector (Photon Bucket): The simplest receiver consists of a single detector followed by electrical processing of the detector output. Such a receiver has been considered in section 3.2.3. For the case of a known signal and background noise which is modeled as a fixed intensity, the detector output processing is known to be essentially a weighted counting operation and the error performance has been evaluated

for certain signal sets.<sup>117-118</sup> When the signal is a random process, as it is for the atmospheric channel, much less is known about the optimum receiver and its performance. The following theoretical results provide some insight into the characteristics of a good system of this type.

The maximum-likelihood receiver for the problem of detecting a spatial Gaussian process that does not vary over the transmitted pulse length in the presence of an additive Gaussian noise field that is spatially and temporally white has been considered in the absence of any real detector effects.<sup>119</sup> The receiver computes a quadratic

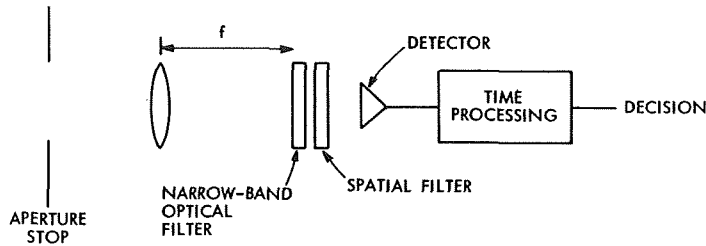


Fig. 5. Single detector (photon bucket) receiver for atmospheric channel.

functional of the received data. By extrapolating from this analysis, we conjecture that a receiver of the form shown in Fig. 5 will provide good performance for the detection of Gaussian fields, even when the real detector statistics are considered. In Fig. 5, the optical filter approximates, for the pulse signal, the temporal matched-filter operation that should be performed on the field. As it is now possible to build optical filters with a  $1/20 \text{ \AA}$  bandwidth, at least at  $6328 \text{ \AA}$ , the optical filter may be a satisfactory approximation to a filter matched to the pulse. Such filters have a field of view limited to approximately a degree, but this should not present a problem for the turbulent channel. The spatial filter has a transmission function that is roughly analogous to that of a field stop, which limits the receiver field of view. In the Gaussian case, the spatial transmission of this filter is

$$t(\vec{\rho}) = \left( \frac{T\tilde{R}(\vec{\rho})}{N_o + T\tilde{R}(\vec{\rho})} \right)^{1/2}$$

when the covariance function of the signal field,  $R(\vec{\rho})$ , is narrow relative to the aperture diameter and  $T$  is the pulse length. Here the tilde denotes a Fourier transform. The exact time processing on the detector output, of course, depends on the signals used, but it seems reasonable to expect that it will be some form of weighted counting.

Some theoretical results, again with real device characteristics ignored, seem to indicate that if the signal is spatially phase-normal rather than Gaussian, that is, it has log-normal amplitude and Gaussian phase processes, more complicated processing

involving additional quadratic functionals is required.<sup>120</sup> These results, which assume that the aperture is large relative to the correlation distance, would call for two quadratic functionals when only phase variations are important, and three when both phase and amplitude variations must be considered. It is possible that the form of these results is due in part to the analytical approximations that were required to make some progress with this very difficult problem. Next, an ad hoc receiving structure based on the Gaussian detection results discussed above will be considered.

On an intuitive basis, if only a single detector is used, it would seem that the receiver shown in Fig. 5 should also be a good choice in the turbulent channel case, that is, when the signal field is phase-normal. While no theoretical results are available to specify the filter for this case, it is reasonable to expect that the filter should be matched to the turbulence so that only the detector area where appreciable signal energy is expected is exposed. In terms of the channel models of section 4.2, this filter could be designed by considering the plane wave or angular spectrum representation of the field over the aperture or, for sufficiently large apertures, by using the focal plane degrees of freedom model previously suggested. Note that these representations provide an indication of the physical size of the detector. Also note that "poor" receiving optics, where the quality is judged relative to the field perturbations caused by the turbulence, will cause an unnecessary increase in the size of this field stop and hence admit more background noise. The time processing, of course, is in general different in detail from that for the Gaussian field, although it is still probably some form of weighted counting.

No general performance results are available for direct detection with the turbulent channel. Some results are available for binary on-off signaling when the aperture is limited to a single amplitude coherence area, and for orthogonal signaling with optimum quantum detection.<sup>121,122</sup> Also, the variation of error probability with log-amplitude variance has been considered for a PCM/PL system operating with a constant background intensity.<sup>118</sup> It has been assumed by Peters and Arguello<sup>118</sup> that the intensity fluctuations for large apertures are log-normally distributed, which still requires further confirmation. The same assumptions could be used to get other special case results. Some experimental direct-detection measurements for a turbulent channel have been made<sup>123</sup> and others are in progress.<sup>124</sup>

Array of Detectors: The potential of using more than one detector can be visualized by using the aperture plane representation of the received field in terms of a set of coherence areas. For the present case, the field amplitude is the important quantity, and it is the amplitude coherence diameter that is important. The use of a detector array permits a more optimal use of the spatial diversity available in the signal field than may be possible with the system of Fig. 5. One possible form for the array receiver is shown in Fig. 6. The two lenses and spatial filter form a field of view filter to limit the background noise. The spatial filter field stop, as in the case of the single detector, would be adjusted to the angular spread of the signal field. The output of this lens system is an image of the aperture with reduced background noise. As in the

single-detector case, the optical bandpass filter acts as a matched filter on the field and further reduces the noise.

The array of detectors is so chosen that each individual detector diameter is equal to the amplitude correlation distance. [Note that additional optics could be used to reduce the required detector size (minify the image).] The output of each detector is individually processed before the outputs are combined to make a decision. For digital signaling with an alphabet of size  $M$ , the time processing associated with each detector

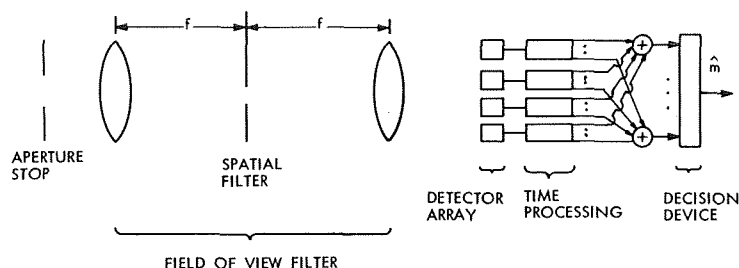


Fig. 6. Direct-detection array receiver for atmospheric channel.

channel will have  $M$  outputs, where each is associated with one of the possible transmitted signals. The outputs corresponding to each particular signal are added together to form the decision variables. The exact processing on each diversity channel is not known. In general, it will depend on both the signal set and the variance of the log-normal amplitude fluctuations. Intuitively, the reason that individual detector processing permits more efficient use of the spatial diversity is that it can perform an operation on the data to suppress the effect of those detectors, and hence aperture areas, that are dominated by noise; that is, in forming the decision variables the outputs of the high intensity signal areas are weighted more heavily than the outputs from the areas with severe fading. Thus the processing attempts to estimate the fade associated with each diversity path using the data from a single baud time and uses these estimates in forming the decision variables.

The performance of this type of receiver has not yet been studied. A particularly important result that should be obtained is the gain in performance, over that of a single detector, which can be attained by the use of a detector array with individual processing. It is not clear how substantial the gain will be.

Channel Measurement Receiver: A channel measurement receiver is an extension of the detector-array receiver. In the array receiver the decision procedure weights the contributions of various portions of the collecting aperture according to the gain associated with that portion. A more efficient operation would, of course, be possible if the gains were actually known. This suggests that the receiver should try to estimate the log-normal channel amplitude space-time process and then use this estimate in carrying out the signal processing. This approach has proved to be a useful



technique for other channels. It is potentially very attractive for the turbulent channel, because of the very slow (1 msec) fading rates that allow long time averages to be used in the channel measurement procedure with a resultant decrease in measurement noise. A pilot tone can be used to facilitate the channel estimation, for example, an unmodulated signal at a nearby wavelength, but it is also possible ordinarily to use the information signal directly for this estimation. The last procedure is often more convenient, and for some channels more efficient in terms of total signal power.

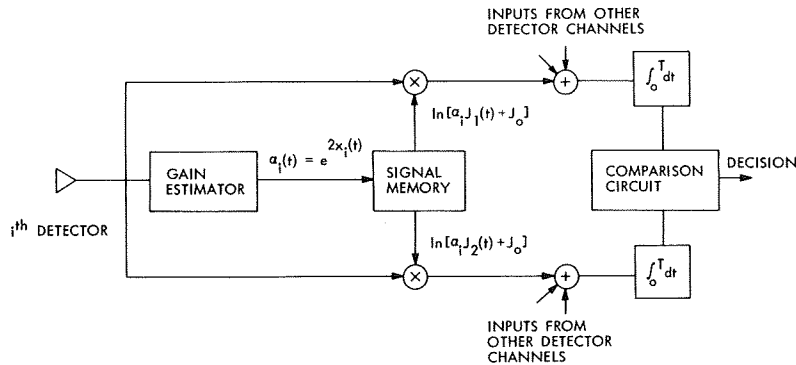


Fig. 7. Direct-detection channel measurement receiver for atmospheric channel. (Only a single diversity element is shown and binary signaling is assumed.)  $J_1(t)$  and  $J_2(t)$  are the intensities associated with the two messages.  $J_0$  is the background intensity.

As an example of the ideas involved, consider the receiver shown in Fig. 7. It is assumed that binary signaling is being used, but the modification for the M-ary case is evident. Only a single one of the spatial channels is shown in detail; there would in general be as many channels as in the array receiver just considered. The detector would be preceded by the field of view and wavelength filters shown in Fig. 6. It should be emphasized that even the limited structure shown here is an ad hoc configuration based on conjecture rather than analysis. The heart of the system is the gain estimator box that uses its knowledge of the possible signals and the background level to form a noisy estimate of  $e^{2\chi_1(t)}$ , the channel intensity gain. The key to its operation is an averaging or filtering of the signal over a time corresponding to the channel amplitude coherence time. This gain estimate is then used in the matched-filter (weighted counting) receiver that would be optimum if the gain were exactly known.\*

Channel measurement receivers look promising and certainly warrant further investigation. The most desirable form of the receiver and its performance should be determined. Emphasis must be given to the effect of signal sets on performance:

\*These ideas are based upon an exposition which P. E. Green submitted at the Workshop (see Appendix I).

Signal Design: The signals for a direct-detection system must be designed to permit discrimination between the signal states on the basis of energy alone. Any form of amplitude modulation is certainly applicable, as is polarization modulation and frequency shift keying. In the last cases, predetection processing is required to convert the information into an energy form. For binary signaling with polarization modulation, this prior processing might take the form of a Wollaston prism to spatially separate the two polarizations; for frequency-shift keying, a set of bandpass optical filters would probably be used.

For direct detection with real devices the performance in general depends on the signal shape, as well as on its energy. Direct-detection performance for digital signaling with the free-space channel is in general improved by the use of pulsed operation with high peak intensities. None of the characteristics of the turbulent channel would seem to change this situation. The signal set that minimizes the probability of error is not, in general, known. On the other hand, characteristics of the turbulent channel are such that most of the signal sets that have been considered for the free-space channel can still be used effectively. Pulse position modulation which is attractive for the free-space channel should still be very attractive for the turbulent channel. This reflects the fact that the time spreading of the channel is very small, and that the frequency coherence bandwidth is very large. The fact that the turbulence causes almost no depolarization makes polarization modulation attractive. The very small frequency spreading caused by the atmosphere makes frequency-shift keying look attractive. This technique is limited by present modulation technology.

Experiments: Several of the more important experiments for direct-detection receiving systems await further theoretical work. This is particularly true of experiments associated with arrays of detectors and with channel measurement. A better understanding of the optimum system implementation and of the design trade-offs is required before the ideas can be meaningfully tested on the turbulent channel. It should be emphasized, however, that such experiments will be necessary to assess the many approximations involved in the system models. Furthermore, the information about the intensity spatial covariance function provided by some of the experiments discussed in section 4.1 will be required for an analysis of these systems.

On the other hand, some meaningful experiments could be carried out immediately with the single-detector receiver. These include performance measurements (probability-of-error measurements) for various turbulence and background conditions to study the optimization of the field stop (spatial filter) and the effects of aperture size. Also, measurements of the intensity distribution for large apertures are needed to verify the extent to which it remains log-normal. This information is needed for the theoretical performance calculations. The experimental situation should, of course, simulate the space-ground path and include a study of the effects of transmitter aperture size and type of aperture illumination. In making error performance measurements, care should be taken to optimize the processing of the detector output signal with

respect to the digital signal set that is being used. Signal sets should not be limited to binary. For example, M-ary PPM should be considered.

#### 4.3.2 Heterodyne Detection Techniques

We now consider turbulent-channel communication systems that use heterodyne (field) detection. This type of detection was discussed in section 3.2.1 for the free-space channel and some of those results are directly applicable to the turbulent channel. The emphasis here will be on the features that arise because of the turbulence. Thus we focus on techniques to overcome the effects of the phase-front distortion and amplitude fading.

We shall consider heterodyne detection within a fairly general framework. This will permit an examination of some of the relevant features. Subsequently, we shall consider important and/or promising receiver structures, signal design, and experiments that would be of interest.

Model for Heterodyne Detection Systems: The model of Fig. 4 can be applied to heterodyne detection by replacing the field  $\epsilon_n(t, \vec{r})$  by a heterodyning field  $\epsilon_h(t, \vec{r})$ . The actual heterodyning field is sometimes introduced ahead of the linear field processing, but this causes no fundamental change because this processing is known and linear. Although the complex envelope of the heterodyning field often has no spatial or time variations, the additional generality provided by these variations is sometimes required.

As discussed in section 3.2.1, it is sometimes appropriate to approximate the vector process  $\underline{r}(t)$  by a term that depends on the field  $\epsilon_o(t, \vec{r})$  (and  $\epsilon_h(t, \vec{r})$ ) and an additive Gaussian noise term that depends only on  $\epsilon_h(t, \vec{r})$ .<sup>125</sup> In this case, all of the detection and communication theory results for random signals in additive, signal-independent, Gaussian noise become applicable. The situation is still complicated, however, by the fact that the random signal process is log-normal (or log-normal derived).

Another source of noise may arise when heterodyne detection is used. This is the noise associated with the heterodyne field  $\epsilon_h(t, \vec{r})$ . If this field is generated by a laser local oscillator, it is usually reasonable to assume that it contains no undesired variations. If, however, this field is obtained by amplifying part of the transmitted signal, this noise will be very significant.

The phase-front distortions caused by the turbulence are extremely important when heterodyne detection is used because the detector output corresponds to an integral of the field over the detection surface. If the field is not spatially coherent over the detector surface, the varying field phase will cause large variations in the detector output signal strength and severely affect performance. Thus a primary concern in designing heterodyne receiver systems is attempting to minimize the effects of the phase-front distortions and still use large receiving apertures to increase the collection of signal power, and hence error performance or communication rate. While the phase-front distortions will be most influential in determining the form of heterodyne receivers, the

amplitude scintillations will be very important in determining their performance. Thus both effects will be important in the examination of the receivers that follow.

We shall now consider several heterodyne receiving structures, beginning with the optimization and performance of the simplest heterodyne receiver and preceding to more complex and powerful systems. The objective is to survey those ideas that look most promising either from the point of view of ease of implementation, or performance, or some trade-off between these two factors. These receivers, of course, try to exploit the channel characteristics discussed in section 4.1. In some cases, the receiver structure is based on analysis, while in others it is suggested by the channel characteristics and previous experience with similar channels.

Single Detector: We consider, first, the receiver that uses a single photodetector. This detector would be preceded by optics to limit the field of view to the angular spread expected in the signal portion of the field. If the receiving aperture is sufficiently small, so that the phase is essentially constant over it, the detector output will correspond to a log-normal signal in independent additive white Gaussian noise. The error performance of such a system, for binary orthogonal signaling and baud times less than the channel amplitude coherence time, has been calculated numerically and also bounded.<sup>121, 126</sup> These results are a function of the parameter  $\sigma_\chi^2$ , the variance of the normalized log amplitude of the received field. As would be expected, if the value of  $\sigma_\chi^2$  is nonzero, more average signal power is required to achieve the same performance as in the nonfading,  $\sigma_\chi^2 = 0$ , case. The situation just considered is essentially true for any reasonable receiver aperture on a spacecraft, because the field there normally has large amplitude and phase-coherence areas.

For the heterodyne receiver on the ground, as the receiver aperture increases, signal variations caused by the phase variations of the received field start to become important. This effect has been studied both theoretically and experimentally.<sup>63, 76, 123, 127, 128</sup> These studies show that the average signal power saturates as the receiver aperture increases, and that for any aperture size, it will not exceed the power that would be associated with an aperture of diameter  $D_{\text{eff}}$  (or  $r_o$ )<sup>63, 76</sup> (see section 4.1.3). The curve of average signal power versus diameter has a knee at  $r_o$  and is approximately 3 dB below the saturated value for an actual receiving diameter of  $r_o$ . For a homogeneous path,  $r_o$  is<sup>63, 76</sup>

$$r_o = 1.2 \times 10^{-8} \lambda_\mu^{6/5} L^{-3/5} C_N^{-6/5},$$

where  $\lambda_\mu$  is in microns, and  $L$  is the path length. It has also been evaluated for an inhomogeneous space-to-ground path using an assumed model for the refractive index variations.<sup>76</sup> Again, it depends on wavelength to the 6/5 power. In summary, the average signal power for a single heterodyne detector is limited, and a larger receiving aperture can be effectively used at 10.6  $\mu\text{m}$  than in the visible wavelength range. Studies have shown that performance is actually degraded when the receiving aperture diam-

eter approaches or exceeds  $r_o$ .<sup>88,121</sup> The signal variance increases rapidly with aperture size beyond  $r_o$ ; the error performance obtained if the aperture is assumed to contain two independent phases is significantly degraded. To the extent that the resulting signal is accurately modeled as a log-normal process, the signal variance results can be used with the previously mentioned results for the detection of log-normal variables in white Gaussian noise to evaluate the performance of these larger aperture receivers. While some theoretical and experimental results are available,<sup>70,123</sup> the use of the log-normal model for the received signal as the aperture size increases should receive further investigation. These considerations tend to indicate that the aperture diameter should be  $r_o$  or less, and that good error performance may require that it be limited to perhaps  $r_o/2$ .

The results just presented are for the case of a uniform local-oscillator field. The effective aperture area can be increased by using a local oscillator whose intensity is matched to the average phase properties of the atmosphere, that is,  $\epsilon_h(t, \vec{r})$  now actually varies with  $\vec{r}$ . This is equivalent to antenna beam shaping and the gain that would be expected is in general small (perhaps 10%). This technique for improving system performance has been studied.<sup>129</sup> (This technique is to be distinguished from instantaneous phase front matching by the heterodyning signal where the gain may be extremely large. We consider this in section 4.3.3.)

Single Detector with Angle-of-Arrival Tracking: The fact that the predominant phase-front distortion for small apertures is a linear tilt has been discussed in section 4.1.3. Thus, one way to increase the effective aperture for a heterodyne receiver, over that which can be used with a single static detector, is to remove the effect of this instantaneous phase-front tilt. This can be accomplished by tracking (measuring) the instantaneous angle of arrival and matching the heterodyning field to it. The range of values and frequencies associated with this angle of arrival have been considered in section 4.1.3. Possible techniques for accomplishing the tracking and introducing the local oscillator have been considered.

The performance improvement that can be attained with tracking heterodyne receivers has been considered theoretically.<sup>104,105,130</sup> These studies consider the resulting power pattern<sup>130</sup> and the increase in receiver aperture or average signal power that can be obtained by tracking.<sup>104,105</sup> The results show that the effective aperture diameter is increased by a factor greater than three, and that the resulting average signal power is increased by more than a factor of 10. Thus, while the static receiver is limited to a diameter such as  $r_o/2$  by the signal fluctuations, it is likely that a diameter of the order of  $1.5 r_o$  can be used in the tracking receiver. This last statement is based on the conjecture that the signal fluctuations associated with a tracking receiver are of the same order of magnitude as those associated with a static receiver whose aperture is approximately one third as large.\* No theoretical or experimental results concerning the signal fluctuations of a tracking receiver are available at the present

---

\* This conjecture was suggested by D. L. Fried

time. To the extent that the signal fluctuations are reasonably modeled as a log-normal process and that their variance can be predicted from the static receiver case, the performance of the tracking receiver can be obtained from results discussed in conjunction with the static heterodyne receiver. These conjectures would lead to the conclusion that the same error performance can be obtained with a tracking receiver as with a static receiver, with a reduction of more than 10 dB in signal power.

Array of Detectors: It may often be desirable to use receiving apertures that are larger than those that can be effectively used by a single detector receiver. One technique to increase the aperture size that can be used with heterodyne detection is the use of an array of detectors with separate processing on each detector output. Such a receiver would be modeled by Fig. 4 when  $\epsilon_n(t, \vec{r})$  is replaced by the heterodyning field,  $\epsilon_h(t, \vec{r})$ , and each of the detector outputs  $r_i(t)$  results from a phase-coherent portion of the signal field. Following the discussion of section 4.2, the detectors may be located in the image plane of preceding optics, which limits the field of view to that occupied by the signal energy, or they may be located in the focal plane of the optics. In the former case each detector area must be limited to the size of the imaged phase-coherence area (which is assumed to be less than or equal to the amplitude-coherence area), while in the latter case the detector size is determined (at least for large apertures) by the diffraction pattern of the receiving optics. In either case, the total number of detectors is essentially the same.

The optimum receiver must process the vector output  $\underline{r}(t)$  in such a way as to make a minimum probability of error decision in the digital communication case. This requires that the output of each detector, each component of  $\underline{r}(t)$ , be separately processed. The optimum use of spatial diversity and the requirement for separate processing of each detector output has already been discussed in section 4.3.1 in conjunction with direct-detection receivers, but the individual detector processing is more important, indeed crucial, in the heterodyning case. The form of the processing that is required has been studied under the assumptions of large phase variance, equal-amplitude and phase-coherence areas, statistical independence of the amplitude, and phase values associated with the various coherence areas, and decision times short relative to the channel coherence time.<sup>111</sup> The resulting receiver, which is min-max relative to phase, requires at each detector output a set of matched filters followed by envelope detectors and a nonlinear operation. The receiver is shown in Fig. 8. There must in general be a filter matched to each signal in the allowed signaling set (correlation operations could also be used). The processed outputs associated with the same signal are added together after the nonlinear operation to form the decision variables. The receiver chooses the message corresponding to the largest sum. The required nonlinear operation has the form of a biased squaring operation which depends on  $\sigma_\chi^2$ , the variance of the normalized log amplitude of the received field, and the signal-energy-to-noise spectral density ratio per diversity path.<sup>122</sup>

The performance of this diversity receiver, with the same assumptions, has also been studied.<sup>111</sup> It has been shown that the probability of error,  $\text{Pr}[\epsilon]$ , for

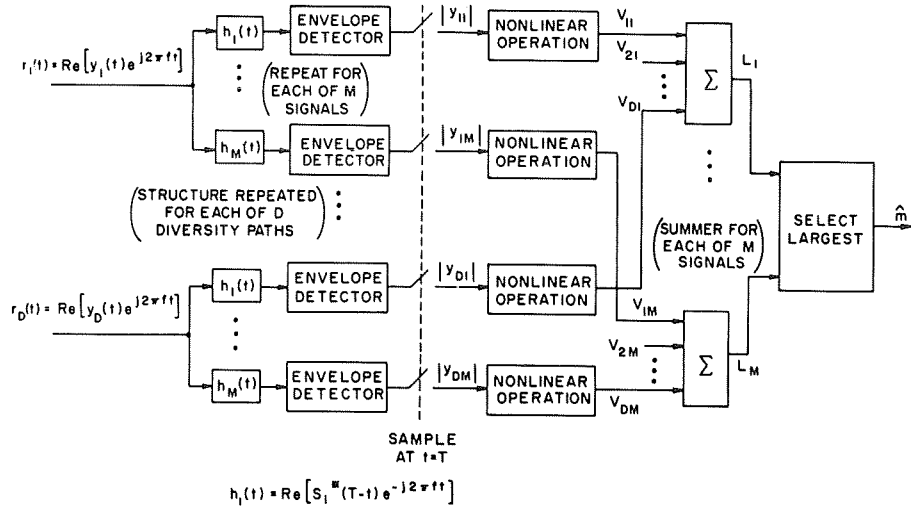


Fig. 8. Min-Max phase receiver for heterodyne detection with an array of detectors. (The  $S_1(t)$  are the complex envelopes of the signal waveforms.)

a signaling alphabet of  $M$  signals, which are orthogonal at the receiver, can be written in the form

$$\tilde{\mathfrak{B}} 2^{-KE} \leq \text{Pr}[\epsilon] \leq 2^{-KE},$$

where  $K$  is the number of information bits per transmitted waveform, and  $E$  is called the "system reliability." The exponent  $E$  depends on  $\beta$ , the average received energy-to-noise spectral density ratio per information bit,  $a_p$ , the averaged received energy-to-noise spectral density ratio per diversity path, and  $\sigma_\chi^2$ , the variance of the normalized log amplitude of the received field. The coefficient  $\tilde{\mathfrak{B}}$  also depends on  $K$ ,  $\beta$ , and  $a_p$ .

A reasonable understanding of system performance can be obtained by using the exponent  $KE$  as a performance measure. It can be shown that, for fixed values of  $K$  and  $a_p$ ,  $KE$  increases at least linearly with increases in the number of coherence areas contained in the receiving aperture. Thus the error exponent increases at least linearly with increasing receiver area. The fact that  $KE$  depends on  $a_p$  makes diversity control desirable. While the spatial diversity is not completely controllable by the system designer, some control can be exercised by varying the transmitter power or antenna gain; also time diversity can be used to augment the available spatial diversity if this is necessary. If appropriate values of  $a_p$  can be achieved, the minimum value of  $\beta$  at which arbitrarily reliable communication is possible is  $\ln 2$ ; that is, the turbulent channel, when used with orthogonal waveforms, has the same channel capacity as an equivalent nonfading white Gaussian noise channel.

Numerical techniques have been used to study, for large values of  $\beta$ , the KE vs  $a_p$  curves as a function of  $\sigma_\chi^2$ . The results show that a broad range of  $a_p$  values will yield relatively good performance, but that too much diversity can greatly degrade the performance. Furthermore, the results show that, even for the optimum value of  $a_p$  which is a function of  $\sigma_\chi^2$ , the error probability for a fixed alphabet size (fixed K) and a receiver of fixed area is increased by the turbulence. Thus, for reasonable turbulence levels and near optimum  $a_p$  values, 3-4 dB of additional signal power at the visible frequencies, and somewhat less in the 10.6- $\mu\text{m}$  portion of the spectrum, are required to achieve the error probability that would result on a nonfading channel. Similar numerical techniques could be used to study the behavior of KE for other ranges of  $\beta$ .

The receiver structure and performance results obtained above should also be applicable to the nonheterodyning case in which the background noise is very large; that is, it is so large that it dominates the signal and the filtered shot process at the detector output looks Gaussian. Similar results are probably also applicable when an imaging optical preamplifier is used in the receiver front end (See sec. 4.3.3, "Avenues for System Development," p. 65).

Detector Array with Angle-of-Arrival Tracking: One way to reduce the required number of separate detectors (and resultant processing), for a given receiving aperture, would be to track the instantaneous angle of arrival of the received field. This type of operation has previously been discussed for a single detector and is extended here to the array case. If the total receiving aperture is not too large, so that the major phase distortion over the entire aperture is a tilt, the number of detectors required in the array can be reduced by tracking the phase-front tilt with the entire array. The individual detectors can now be approximately three times larger in diameter without significantly degrading the performance. The relatively large phase-coherence diameters at 10.6  $\mu\text{m}$  might make this system attractive at that wavelength. If the receiving aperture is too large ( $>3$  or  $5 r_0$  or some other criterion), the phase distortion over the entire aperture is poorly approximated by a linear tilt and the idea of tracking with the entire array is no longer attractive. In this case the tracking would have to be done by individual detectors (or subarrays). Note that phase-front tracking may reduce the required number of detectors by a factor of approximately ten.

It should be emphasized that the over-all system that is being discussed here is essentially a cascade of a tracking system(s) and an array-processing system. Because of this, the comments on phase-front tracking and array processing made in the two preceding sections are directly applicable. Also, as with the case of a single tracking detector, no direct performance calculations are available, but the performance can perhaps be inferred from the static-array case; the conjectures required for this to be the case are similar to those for the single tracking detector.

Array with Channel Measurement: If the time-space variations of the amplitude and phase of the received field were known, the optimum receiver would use this knowledge to help it make its decisions. For the receivers discussed above no attempt has



been made to obtain this type of channel knowledge with the exception of the angle-of-arrival tracking receivers, which do measure and exploit a knowledge of the field angle of arrival, although not directly in the decision process. Thus a logical extension of the array receivers discussed above is a channel-measurement receiver which attempts to measure both the amplitude and phase associated with each coherence area in the receiving aperture, and uses this measurement in its decision process. Such a procedure is attractive for the turbulent channel because the amplitude and phase changes associated with each spatial diversity path are very narrow-band processes ( $<1000$  Hz). The narrow-band property means that accurate measurement of these processes should be feasible. The gain of this receiver over the simple array receiver is that the decisions can be based on channel measurements made over times long relative to the basic decision time for the signal set that is being used. If this decision time approaches the channel coherence time ( $\approx 1$  msec), then the measurement receiver will not provide any gain; on the other hand, if high rates and reasonable signal set sizes are to be used, the decision time will be short relative to the channel coherence time. This type of receiver also exploits more fully the instantaneous properties of the phase process than was the case with the array receiver which assumed phase independence from coherence area to coherence area; this latter assumption is clearly not true and leads to a suboptimal receiver. It should also be noted that the angle-of-arrival tracking and channel measurement ideas can be used simultaneously; again this would reduce the number of detectors required.

One implementation of the type of channel-measurement receiver described above is considered in Appendix I. The required processing is developed as an extension of the array receiver previously described. There are many open questions about the exact receiver structure and the signaling schemes that are appropriate for this application. In particular, the signals should be such that the channel measurement can be effectively carried out without dependence on the signal sequence that is actually transmitted. As pointed out in Appendix I, the relative stability between the transmitting and local-oscillator lasers will be a crucial parameter, at least in the system implementation discussed there. This requirement may, however, be somewhat relaxed by the use of one or more phase-lock loops, and other realizations of the basic measurement idea such as those discussed in section 4.3.3 essentially avoid this particular difficulty.

The performance of such channel measurement receivers is unknown at the present time. The evaluation of this performance will involve both questions of how well the channel parameters can be measured and how well the receiver can make its decisions when using noisy channel estimates.

Signal Design: The signal design question for heterodyne reception essentially reduces to the signal design question for radiofrequency channels in which the channel is slowly fading and the noise is additive, white, and Gaussian. The only differences are that the fading is log-normal and the phase Gaussian. If, however, a large phase variance is assumed, the channel behaves as though the phase had a uniform phase

density. As well as the normal modulation techniques, polarization modulation is also readily available at optical frequencies. If it is used, care must be taken in the receiving optics and heterodyning system to ensure that no incidental polarization changes are introduced.

Experiments: As with direct detection, the form and advisability of some heterodyne system experiments must await further theoretical work, for example, the measurement receiver. There are several experiments, however, which can and should be performed soon. These are concerned in part with the statistics of the resulting fading signal after heterodyning. In particular, the fading statistics associated with a single detector, as a function of receiving aperture diameter, should be measured. This information will assist the evaluation of the performance of such receivers when the receiving aperture exceeds a coherence diameter. Similar measurements of the fading statistics should be made for a phase-front tracking receiver. These measurements are needed before the static and tracking receivers can be compared with any confidence. It should be emphasized that the statistical measurements, probability density, and perhaps joint density which are being discussed here would provide much more information than the few signal-to-noise ratio measurements available, at present. Also measurements should be made to evaluate the performance of heterodyning arrays.

#### 4.3.3 Avenues for System Development

Several means of improving optical system performance, over that previously contemplated, have already been discussed, for example, proper use of the spatial diversity within an aperture, and estimation of channel parameters for use in tracking or adaptive receivers. Another means of improving system performance may be to employ an adaptive transmitter, that is, a transmitter that measures the state of the channel and bases its operating strategy upon these measurements.

Several kinds of transmitter adaptation are of interest. For example, one might attempt to estimate the path loss and adjust the transmitted data rate to ensure a given level of performance. Or, one might attempt to estimate the spatial phase degradations that the atmosphere will introduce, and predistort the transmitted phase front to compensate for these distortions.<sup>131</sup> The latter adaptation, at present, is somewhat speculative and falls in the category of advanced research; the former adaptation has been used on some nonoptical channels and is better understood. We now briefly discuss it further. [These ideas were introduced at the Workshop by P. E. Green.]

Variable-Rate Communication: Variable-rate communication is of interest because the amplitude fluctuations on the atmospheric channel are log-normal and slowly varying. The log-normal distribution is very skewed and values that far exceed the mean are not improbable. Astronomers find that with small apertures sudden intensity increases of up to 400% above the mean level, with durations of a hundredth of a second, are not unusual.<sup>48</sup> The objective of variable-rate communication is to exploit these periods of high channel gain.

To be specific, we consider a ground-to-space communication system. The space receiver will normally observe only a single coherence area, so it is reasonable to try to match the transmission rate at the earth terminal to the instantaneous channel gain. The improvement is large for Rayleigh fading\* and should be even greater for the more asymmetrical log-normal form of fading. Of course, the use of a variable-rate system requires that the channel gain be monitored. For the ground-to-space channel, under the assumption of instantaneous atmosphere reciprocity, this gain can be monitored by measuring the gain of a signal coming from the spacecraft. (The concept of instantaneous atmosphere reciprocity is considered in Appendix G.) The effect of the point-ahead angle requirement may limit the utility of the technique at visible frequencies. It does not appear to be a problem at the far infrared ( $10.6 \mu$ ), however, even at synchronous altitudes (see D. L. Fried, Appendix H).

The variable-rate communication system, while an attractive idea, must be examined further to establish its potential and applicability. In addition to the reciprocity concept, the system-performance improvement, in terms of rate gain or power savings, and the associated queuing times, with the resultant memory requirements, require investigation.

Transmitted Reference Systems: Transmitted reference techniques provide another promising approach to the problem of channel measurement and receiver adaptation. Briefly stated, one first observes that, if the spatial phase of the received field were known, a near optimum receiver would heterodyne the received field with a local-oscillator field possessing the same spatial phase. One then observes that the required "local-oscillator" field can be obtained by transmitting a reference, or pilot tone, along with the signal. Precisely stated, if the reference and signal fall within a coherence bandwidth, they will both experience the same spatial phase distortion in the channel and hence the received "pilot tone" will have the desired spatial phase and may be used as a "local oscillator."

The difficulty with the transmitted reference scheme is that a strong local oscillator can be attained only at the expense of signal power. An appealing variation on the scheme employs an optical amplifier to bring the level of the received pilot tone to the desired value. This amplifier must, however, preserve the spatial phase of the pilot tone; that is, it must be an "image amplifier." Such a system is discussed further in Appendix I. Its feasibility should be further investigated. More generally, the question of optimum quantum receivers for transmitted reference systems should be explored.

---

\*For example, to achieve a  $10^{-6}$  error probability one would have to use 47 dB or more average transmitter power with a fixed rate system than with a variable-rate system whose bit rate was varied to maintain constant received signal-to-noise ratio.<sup>132</sup>

## V. Scattering Channels

One objective of the Workshop was to determine the questions that must be answered and the data that must be obtained to assess the feasibility of optical communication through clouds, fog, and haze. Since this possibility has not been widely recognized, the questions that emerged can only be viewed as an initial attack on the problem. If the answers to these questions are encouraging, more comprehensive investigations will be in order.

The common characteristic of channels containing clouds, fog, and haze is that they involve an appreciable amount of scattering of the incident radiation. Although many of our comments apply to all of these channels, we shall frame our discussion in the context of water clouds.

That the utility of an optical communication link is reduced by the presence of clouds is well known. Indeed a system that is designed to operate efficiently in their absence will often be inoperable in their presence. Preliminary results suggest, however, that useful communication from space to ground through clouds may be possible with a properly designed system. In particular, the effect of clouds upon the optical field at a receiver is grossly similar to the effect of the turbulent atmosphere; in both situations the field is distorted in time, frequency, and space. The magnitudes of the effects differ appreciably, however. For example, the turbulent atmosphere may spread the beam-width of a highly collimated pencil beam to tens of microradians – a thick water cloud may spread it to tenths of a radian. Also, the coherence bandwidth of the atmosphere appears to be at least tens of GHz – that of a thick cloud may be of the order of MHz. Finally, the fading statistics of the field received through the clear turbulent atmosphere are probably log-normal; for a thick cloud they are probably Gaussian.

### 5.1 COMMUNICATION CONSIDERATIONS

The differences just discussed are important; but the gross similarity between cloud and atmospheric channels suggests that the potential for communication through clouds is primarily limited by the ratio,  $\bar{P}_r/N_o$ , of the total average collected signal power to the noise power per Hz that would be collected if the receiving aperture were diffraction-limited.<sup>111, 133</sup> This conclusion is perhaps even better established for the cloud channel than for the turbulent atmosphere because channels with Gaussian statistics have been more thoroughly investigated than have those with log-normal statistics.<sup>134, 34, 135</sup>

In particular, if large bandwidths and high peak powers can be employed, the system capacity is just  $1.44\bar{P}_r/N_o$  bits/sec; that is, the system error probability can be made arbitrarily small (at the expense of system complexity) for any information rate less than  $1.44\bar{P}_r/N_o$  bits/sec. This capacity is, of course, just that which would be obtained in the absence of a cloud for the same ratio of total average collected received power to noise power density per Hz per diffraction-limited antenna.

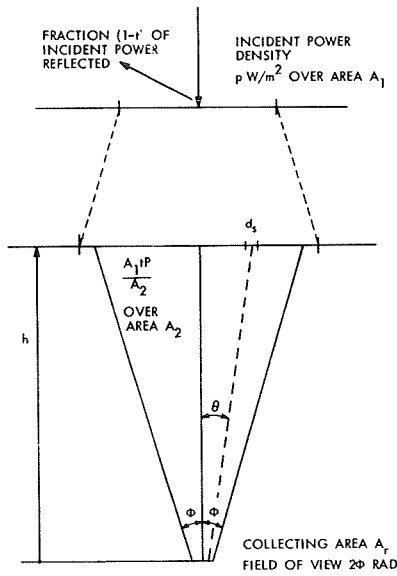


Fig. 9. Model for propagation through a thick cloud.

we ignore the radiation emerging from the sides. Furthermore, we suppose that the cloud is so thick that its lower surface behaves as a Lambertian radiator that is "uniformly" bright over some circular area  $A_2$ . The power density from this area is then  $\frac{A_1 t}{A_2} p \text{ W/m}^2$ .

Now suppose that the receiving optics has a circular collector of area  $A_r$  parallel to the cloud surface and a field of view with cone half-angle  $\phi$  which is completely filled by the area  $A_2$ . The power collected from an incremental surface area  $ds$  of the cloud will then be

$$dP = \frac{1}{2\pi} \left( \frac{\cos \theta}{2} \right)^2 \frac{A_1 p t A_r \cos \theta}{A_2} ds, \quad (36)$$

where  $\theta$  is the zenith angle to the element. If the axis of the optical system passes through the center of  $A_2$ , the total collected power will be

$$\bar{P}_r = \frac{A_1 A_r p t}{A_2} [1 - \cos \phi], \quad (37a)$$

provided that

$$\tan \phi \leq \frac{1}{h} \sqrt{\frac{A_2}{\pi}}. \quad (37b)$$

Given the preceding conclusions, the next question is, To what extent does the presence of a cloud reduce the power that can be collected by a given receiving aperture? The answer, at present, is not precisely known, but in the absence of absorption the amount of power collected by a receiver located "near" the cloud is often at least a few per cent of that which could be collected in the absence of the cloud. To see how this can occur, we consider the particular case of a thick nonabsorbing cloud.

As illustrated in Fig. 9, we suppose that a beam of power density  $p \text{ watts/m}^2$  and area  $A_1$  is normally incident upon the upper surface of a thick planar cloud. A fraction,  $1-t$ , of the power is reflected and, in the absence of absorption, the remaining fraction,  $t$ , traverses the cloud emerging from the bottom or sides. For simplicity,

For purposes of comparison, the power collected in the absence of the cloud would be

$$\tilde{P}_r = pA_r$$

if the transmitted beam did not diverge appreciably in propagating the distance  $h$ . Note that the ratio of the powers is

$$\frac{P_r}{\tilde{P}_r} = \frac{A_1 t}{A_2} [1 - \cos \phi].$$

In the particular case of an infinite plane wave incident upon the cloud (so that  $A_1 = A_2 = \infty$ ) and a receiver with a  $2\pi$  sr field of view (so that  $\phi = 90^\circ$ ),  $P_r/\tilde{P}_r$  is approximately equal to  $t$ . Since  $t$  often exceeds 10%, we conclude that, subject to our many assumptions, the presence of the cloud would reduce the collected power by at most a factor of 10. Although such a loss is certainly significant, it may be acceptable in some circumstances.

Of course, the loss increases, for a given field of view, as the distance from cloud to receiver increases and also as absorption in the cloud increases. Moreover, it varies with wavelength and cloud type. Thus the crudity of the preceding arguments cannot be overemphasized; however, they do suggest that the possibility of optical communication through clouds should be further investigated.

Our ignorance of the channel characteristics is so great that it would be premature to discuss system structure in any detail, but a few general observations are in order. Since very small coherence areas are expected for clouds located near the receiver, an efficient receiver must properly exploit the "diversity" contained in its aperture just as a system must for the turbulent atmosphere. In fact, the entire discussion of receivers for the atmosphere is relevant to clouds, although the reduced coherence bandwidth (increased time dispersion) of clouds may limit the utility of transmitted reference systems. Also, because of this dispersion, special signaling techniques may be necessary when operation at high information rates is required.

A practical difference between the use of atmospheric and cloud channels relates to the field of view requirements and the size of the coherence areas. As noted previously, the receiver field of view may have to be much larger for a cloud channel than for an atmospheric channel — perhaps tens of degrees. Also, the coherence areas of a cloud channel will often be much smaller than those of an atmospheric channel. Consequently, the already difficult problems of implementing a near-optimum atmospheric system become even more pronounced for a cloud system — at least when the cloud is near the receiver.

In summary, efficient systems for operating through clouds may be feasible, but they will probably require the development of new devices. These devices are similar in kind to those needed for atmospheric systems, but the requirements are more

severe. Based upon crude extrapolations of available data, it appears that such systems may yield adequate performance in some situations. A thorough investigation of the channel characteristics is, however, a prerequisite to any serious contemplation of its use. Some of the channel characteristics to be investigated will now be considered.

## 5.2 CHANNEL CHARACTERIZATION

To fully assess the communication potential of scattering channels, it is necessary to know the statistical properties of the field in the receiving aperture for the modulation scheme of interest. Although it has not been verified, the multiple scattering that occurs in a thick cloud will probably result in a received field that behaves as a zero-mean Gaussian random process. Assuming this to be true, the channel characterization problem reduces to that of determining the correlation function of the received field. Although complicated by the spatial fading that occurs, this problem is similar to that encountered with other fading dispersive channels; hence, we shall not pursue the general approach to the problem here.<sup>134, 34, 135-138</sup>

A complete determination of the correlation function referred to above is desirable, but experience with other channels has shown that its detailed behavior does not greatly influence the communication performance levels that can be realized. In fact, a set of approximately four parameters usually provides sufficient information to grossly estimate the performance levels that can be achieved. Were it not for the spatial distortion of the wavefront, three parameters would suffice: the total average received power collected by the aperture, the average time dispersion suffered by a short pulse, and the average frequency dispersion suffered by an unmodulated carrier. In the presence of spatial distortion, an estimate of the number of coherent areas contained in the receiving aperture is also required. Stated alternatively, one must know the number of essentially independent plane waves that are required to characterize the field over the receiving aperture, that is, the angular spectrum. These parameters, of course, are just those discussed in connection with the turbulent atmosphere.

Clearly, the complete correlation function of the field, on the one hand, and the four-parameter description, on the other, represent extremes of channel characterization. Between them lie a variety of compromises in completeness and complexity. One characterization that may be particularly useful in the development of efficient systems consists in specifying the absolute propagation delay and Doppler shift, as well as the time and frequency dispersion, associated with each of the plane waves that are contained in the angular spectrum of the received field.

The development of adequate characterizations for cloud channels will involve an application of scattering theory. We now briefly review the state of that theory, the information about the cloud that is required to apply it, and the physical properties of clouds, fog, and haze.

### 5.2.1 Scattering Theory

The scattering and absorption of a plane, monochromatic wave by a homogeneous sphere can be described exactly by using a theory developed by G. Mie.<sup>139</sup> The theory yields absorption and scattering cross sections for arbitrary polarization of the incident beam, as well as the distribution of the light scattered out of the beam, that is, the scattering function. Efficient computer programs for evaluating the relevant expressions have been written. The cross sections depend on the refractive index of the particle with respect to the surrounding medium, the wavelength of the incident light, and the particle diameter. The more difficult case of many-layered spherical droplets with different refractive indices has been solved by several investigators.<sup>140, 144</sup> The resulting equations can also be programmed, with additional computer time being the main price paid for the additional complexity. For more realistic models of atmospheric particles (that is, irregular, possibly many-layered, particles), the scattering properties are not calculable, but they can be measured.

The single-particle properties can be integrated for a particle size distribution (if known) to arrive at the monochromatic volume scattering properties of the medium. Thus for an optically thin medium with plane-parallel boundaries in which only single scattering is important, the reflected and transmitted intensities are simple functions of the medium's optical thickness and geometry. The optical thickness of a medium along an incremental path length  $ds$  is  $d\tau_E = \beta_{\text{ext}} ds$ , where  $\beta_{\text{ext}}$  is the volume extinction coefficient, the sum of the volume scattering and absorption coefficients ( $\beta_{\text{ext}} = \beta_{\text{sc}} + \beta_{\text{abs}}$ ).

For optically thick media, in which multiple scattering must be considered, the transfer problem is extremely complicated. Except for a few simple cases, solutions to the exact equations of radiative transfer are difficult to obtain. The problem becomes even more complicated when polarization is taken into account because the light beam must be expressed in terms of the four-component Stokes vector notation and the scattering phase function becomes a four-by-four matrix. In general, polarization should be considered, since an initially unpolarized beam will be partially polarized after a single scattering event. Its importance in communication applications is still not clear.

At least two alternative means of approaching the multiple-scattering problem exist. One means consists in modeling the scattering events by a random-walk, or Monte Carlo, calculation. This method can be applied to nonhomogeneous media and is not limited to plane-parallel boundaries. Polarization can be included in the model quite easily; also, nonstationary problems can be solved. Several authors<sup>142, 143</sup> have used the Monte Carlo method to calculate multiple-scattering effects in haze and clouds. Their results for multiple scattering in a plane-parallel, Rayleigh (or molecular) atmosphere compare very well with the exact results of Chandrasekhar and of Coulson, Dave, and Sekera.<sup>144</sup>

Another means of approaching the multiple-scattering problem is to model a thick cloud as a concatenation of thin clouds within which only single scattering occurs.<sup>138</sup>



This approach involves approximations that preclude high-precision results, but it has the advantage of yielding closed-form results that may be suitable for communication system analysis and design. Further correlation of the results obtained from this approach with those obtained by Monte Carlo techniques and by experiment will be required to establish their suitability for communication applications.

In summary, several techniques for predicting the transmitted (and reflected) intensities from haze, fog, or clouds are at hand. Also, the scattering properties of spherical particles can be described exactly if we know the particle size distribution, and the real ( $n$ ) and imaginary ( $n'$ ) parts of the index of refraction. More information on cloud particle composition, size distribution, and particle shape may be required to apply these results to real clouds. Alternatively, it may be worth while to try to assess the significance of these unknown parameters by examining extreme cases covered by existing models. If the results indicate a lack of sensitivity to them, our present knowledge of cloud characteristics may suffice. This state of knowledge will now be briefly summarized.

### 5.2.2 Physical Characteristics of Clouds, Fog, and Haze

Clouds are assemblies of ice crystals and/or water droplets with sizes ranging from 1 to 50  $\mu\text{m}$ . The liquid water content of clouds varies from 0.1-1.0  $\text{gm}/\text{m}^3$  for clouds above 2 km, to as much as 10  $\text{gm}/\text{m}^3$  for low-level clouds. Detailed knowledge of clouds is incomplete, but certain general characteristics are known. At altitudes where the air temperature is less than  $-40^\circ\text{C}$ , clouds in thermal equilibrium with the surrounding air are composed of ice crystals rather than water droplets. Samples show that the ice crystals in cirrus and stratus clouds are predominantly prismatic columns with hexagonal bases averaging 40  $\mu\text{m}$  in width and 100  $\mu\text{m}$  in length. The particle density in stratus and cirrus clouds is small (0.1-30 particles/ $\text{cm}^3$ ). There is no information about the properties of high-altitude cumulus clouds. Low-altitude cumulus cloud samples indicate that they contain many more particles (100 particles/ $\text{cm}^3$ ) and that the droplets are smaller, ranging in diameter from 5.0  $\mu\text{m}$  to 30  $\mu\text{m}$ . On occasion, these lower clouds can be composed both of ice crystals and water droplets. The cloud structure is far from homogeneous, and the cloud properties can vary considerably from point to point in the cloud.

Haze and fog, being more accessible to ground-based instruments, have been studied extensively for many years, with a large variety of instruments and techniques. According to Rozenberg<sup>145, 146</sup> haze and fog are characterized by smaller particle sizes. Haze is composed of finely dispersed formations of particles with diameters that are less than 1.0  $\mu\text{m}$  and with an average particle size that is close to 0.25  $\mu\text{m}$ . Dry haze is observed at relatively low humidity (<80%), foggy haze at moderate to high humidity (>85%), and true fog is observed only at saturation. Roughly stated, for dry haze the visibility (or meteorological range) is greater than 3 km; for foggy haze, visibility is between 1 km and 3 km, and for true fog the visibility is less than 1 km.

Haze and foggy haze are relatively stable atmospheric states. The transition from haze to foggy haze is practically continuous, and is usually accompanied by the disappearance of relatively fine particles and the appearance of larger particles (presumably by particle growth and agglomeration).

### 5.2.3 Complex Refractive Index for Water and Ice Particles

The complex refractive indices of water and ice play an important role in scattering theory. The real and imaginary parts of this index are denoted  $n$  and  $n'$ , respectively.

The values of  $n$  and  $n'$  are known "reasonably" well for liquid water.<sup>147</sup> Recent experimental results at  $10.6 \mu\text{m}$  suggest, however, that the tabulated values may not be sufficiently accurate for path loss estimates at that wavelength, because of the strong influence of  $n'$  on the results of theoretical calculations. Direct measurement of  $n'$  is difficult, but an approximate value can be obtained by measuring the intensity of the radiation scattered at several well-chosen angles by an artificial water cloud (see section 5.3). The uncertainties in the published values of  $n$  and  $n'$  for ice are probably greater than those for water, so theoretical calculations of scattering by ice clouds are questionable.

There is some experimental evidence that relatively thin ice clouds absorb appreciably.<sup>148</sup> Results of experiments on the reflection and transmission of ice clouds and frost layers indicate that near  $2.8 \mu\text{m}$  the reflectance is low and the transmittance is greater than at neighboring shorter or longer wavelengths. This phenomenon, known as the Christiansen filter effect, is caused by a decrease in  $n$  on the short wavelength side of a strong absorption band. Transmission measurements on artificial ice clouds would determine if the transmittance near  $2.8 \mu\text{m}$  is great enough to be of value in a communication system. The possibility of increased transmittance on the short wavelength side of the strong absorption band near  $13 \mu\text{m}$  should also be investigated. Since the Christiansen filter effect requires a rather sharp absorption band and the liquid water absorption bands do not appear to be as sharp as the bands of ice, a useful Christiansen effect in a water-droplet cloud is less probable than in an ice cloud.

## 5.3 PROPOSED EXPERIMENTS

We now describe experiments that would provide some of the required information just discussed. The proposed experiments are relatively simple and are designed with two objectives in mind. The first objective is to obtain enough data on the scattering and absorption properties of real clouds to determine whether optical communication through clouds is feasible, and if so, to determine the order of magnitude of the attenuation to be expected from various types of clouds. The second objective is to obtain some of the information about  $n$ ,  $n'$ , and cloud composition that is now lacking. No advances in the state of the art are required to perform the experiments.

### 5.3.1 Laboratory Experiments on Artificial Clouds (Ice and Water)

The scattering phase function for various sized single spherical particles of water and ice should be measured at several wavelengths and all scattering angles. This will provide information on  $n$  and  $n'$  and will give a check on the theory for single particles. Additional experiments should be performed on single ice crystals of different sizes and shapes in order to obtain information about their scattering properties relative to spheres.

The extinction and scattering at several angles by artificial clouds of water droplets and ice should be measured at several wavelengths in the atmospheric windows, particularly in regions where one might expect the Christiansen filter effect. The results will provide much needed information on  $n$  and  $n'$  and a check on the theory for a wide range of optical thicknesses.

### 5.3.2 Natural Cloud Experiments

One experiment of this kind is designed to measure the insertion loss through typical clouds. Specifically, a heliostat, or other sun-tracking device, should be used to form an image of the sun on a monochromator, or other bandpass filter. The major atmospheric windows from the visible to 25  $\mu\text{m}$  should be included, and the field of view of the instrument should be smaller than the sun. The radiant intensity in the direction of the sun would be measured through a clear sky and through a variety of cloud cover. Measurements should also be made while viewing clouds within  $10^\circ$  or  $15^\circ$  of the sun.

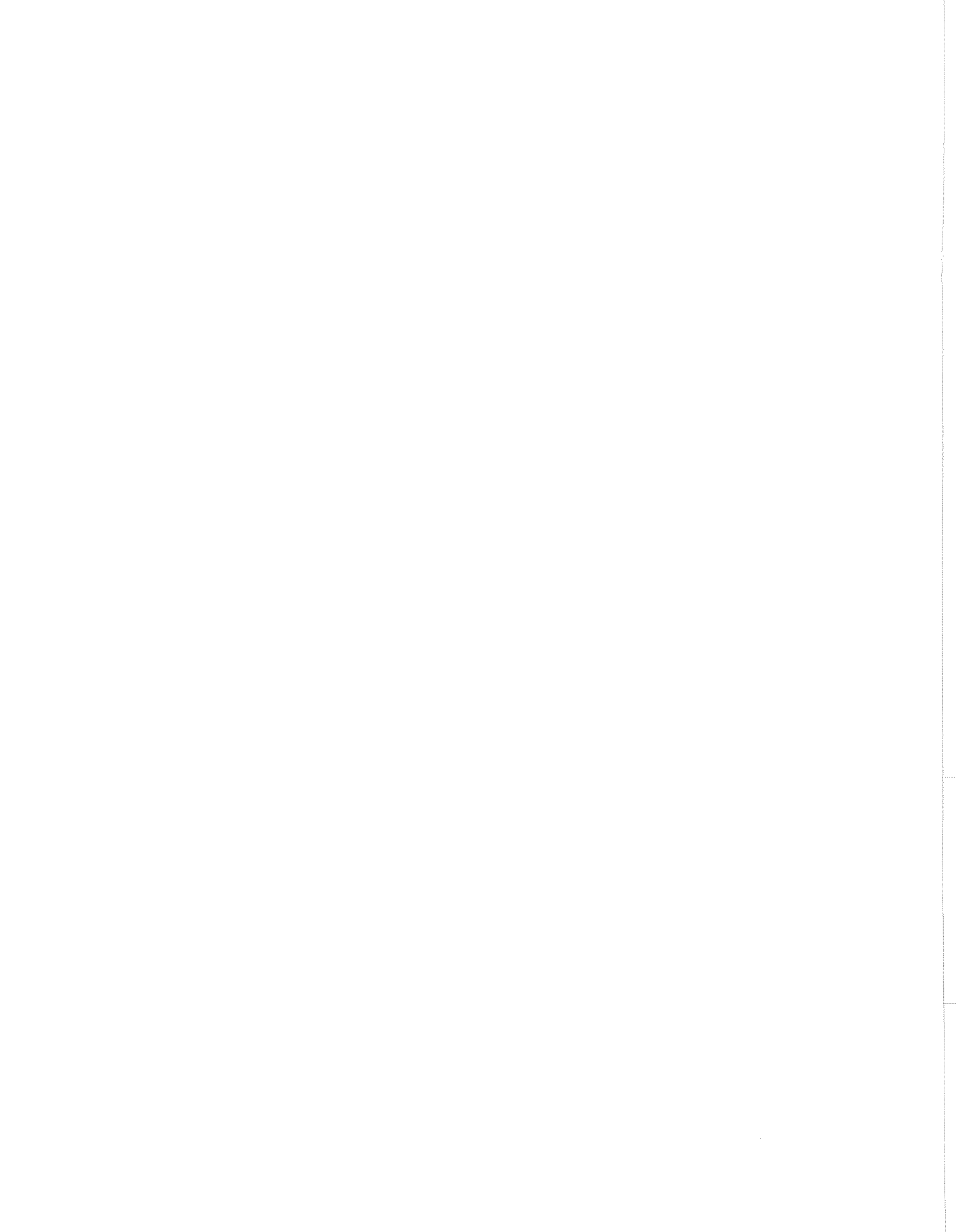
The experiment will provide information on the attenuation coefficients at different wavelengths. Although the cloud composition will not be known accurately, it may be possible to establish approximate cloud conditions under which optical communications are feasible. By making some realistic assumptions about cloud composition, rough checks on the theory and on the values of  $n$  and  $n'$  can be obtained.

Another useful experiment would illuminate a cloud with a narrow beam. The energy source could be a visible wavelength laser, mounted in an airplane or on a mountain top. Receivers on the ground would record the signals received through the cloud. Given this configuration, or its converse, the following measurements suggest themselves.

1. With a cw source, the average energy distribution on the underside of the cloud should be mapped, using a scanning energy-detecting receiver and several receiver fields of view. The fluctuations in energy should also be examined.

2. Again with a cw source, using two energy-detecting receivers, the correlation of the energy from two small areas on the underside of the cloud should be studied as a function of the spacing between these areas. For this measurement, the area should be as small as possible; that is, the receiver field of view should be as small as is consistent with usable signals. For thick clouds the expectation is that the two energies will be uncorrelated if the areas do not overlap.

3. With the energy source pulsed, and two energy-detecting receivers, the time-dispersion effects of multiple scattering through the cloud should be measured. As a very rough estimate, one would expect multipath dispersion through a 1-km thick cloud to be between 0.3  $\mu$ sec and 3  $\mu$ sec. On this scale, a source pulse of 100 nsec duration with receiver time resolution of 200 nsec would be adequate.



## References

1. L. Mandel, in Progress in Optics, Vol. II, E. Wolf (ed.). Interscience Publishers, New York, 1963.
2. E. C. G. Sudarshan, "Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams," Phys. Rev. Letters, Vol. 10, No. 7, pp. 277-279, April 1, 1963.
3. P. L. Kelley and W. H. Kleiner, "Theory of Electromagnetic Field Measurement and Photoelectron Counting," Phys. Rev., Vol. 136, No. 2A, pp. 316-334, October 19, 1964.
4. R. J. Glauber, in Quantum Optics, C. Dewitt et al. (eds.). Gordon and Breach Science Publishers, Inc., New York, 1965, pp. 65-185.
5. P. L. Kelley, et al. (eds.), Physics of Quantum Electronics. McGraw-Hill Book Co., New York, 1966.
6. R. Serber and C. H. Townes (eds.), Symposium on Quantum Electronics. Columbia University Press, New York, 1960.
7. J. R. Singer (ed.), Advances in Quantum Electronics. Columbia University Press, New York, 1961.
8. "Quantum Electronics Issue," Proc. IEEE, Vol. 51, January 1963.
9. N. Bloembergen, Nonlinear Optics. W. P. Benjamin, New York, 1965.
10. M. Ross, Laser Receivers. John Wiley and Sons, Inc., New York, 1966.
11. L. K. Anderson and B. J. McMurtry, "High-Speed Photodetectors," Appl. Opt., Vol. 5, p. 1573, October 1966.
12. K. M. van Vliet, "Noise Limitations in Solid State Photodetectors," Appl. Opt., Vol. 6, No. 7, pp. 123-127, July 1967.
13. A. Yariv, Quantum Electronics. John Wiley and Sons, Inc., New York, 1967, p. 412.
14. E. J. Reinbolt and J. L. Randall, "How Good Are Lasers for Deep-Space Communications?," Astronautics and Aeronautics, NASA.
15. M. E. Lasser, "Detection of Coherent Optical Radiation," IEEE Spectrum, pp. 56-89, April 1966.
16. E. Brookner and M. Kolker, "Deep-Space Optical Communications," IEEE Spectrum, pp. 116-118, January 1967.
17. F. E. Goodwin and T. A. Nussmeier, "Optical Heterodyne Communications Experiments at 10.6  $\mu$ ," IEEE J. Quantum Electronics, Vol. QE-4, No. 10, pp. 612-617, October 1968.
18. J. R. Kerr, "Microwave-Bandwidth Optical Receiver Systems," Proc. IEEE Vol. 55, No. 10, October 1967.
19. R. F. Lucy, K. Lang, C. J. Peters, and K. Duval, "Optical Superheterodyne Receiver," Appl. Opt., Vol. 6, No. 8, pp. 1333-1342, August 1967.
20. W. K. Pratt, "Binary Detection in an Optical Polarization Modulation Communication Channel," IEEE Trans. on Communication Technology, Vol. COM-14, No. 5, pp. 664-665, October 1966.
21. W. S. Read and R. G. Turner, "Tracking Heterodyne Detection," Appl. Opt., Vol. 4, No. 12, pp. 1570-1573, December 1965.
22. M. C. Teich, "Infrared Heterodyne Detection," Proc. IEEE, Vol. 56, No. 1, January 1968.
23. R. T. Denton and T. S. Kinsel, "Terminals for a High-Speed Optical Pulse Code Modulation Communication System: I. 224-Mbit/s Single Channel," Proc. IEEE, Vol. 56, No. 2, pp. 140-145, February 1968.

24. T. S. Kinsel and R. T. Denton, "Terminals for a High-Speed Optical Pulse Code Modulation Communication System: II. Optical Multiplexing and Demultiplexing," Proc. IEEE, Vol. 56, No. 2, pp. 146-154, February 1968.
25. M. Ross, op. cit., p. 24.
26. A. E. Siegman, "The Antenna Properties of Optical Microwave Receivers," Appl. Opt., Vol. 5, No. 10, p. 1588, October 1966.
27. M. Ross, op. cit., Chap. 4.
28. B. M. Oliver, "Thermal and Quantum Noise," Proc. IEEE, Vol. 53, No. 5, pp. 106-111, May 1965.
29. C. W. Helstrom, Statistical Theory of Signal Detection. Pergamon Press, New York, 1960.
30. J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering. John Wiley and Sons, Inc., New York, 1965.
31. D. J. Sakrison, Communication Theory: Transmission of Waveforms and Digital Information. John Wiley and Sons, Inc., New York, 1968.
32. A. J. Viterbi, Principles of Coherent Communication. McGraw-Hill Book Co., New York, 1966.
33. H. L. Van Trees, Detection, Estimation and Modulation Theory, Part I. John Wiley and Sons, Inc., New York, 1968.
34. M. Schwartz, W. R. Bennett, and S. Stein, Communication Systems and Techniques. McGraw-Hill Book Co., New York, 1966.
35. B. Reiffen and H. Sherman, "An Optimum Demodulator for Poisson Processes; Photon Source Detectors," Proc. IEEE, Vol. 51, No. 10, pp. 1316-1320, October 1963.
36. J. W. Goodman, "Optimum Processing of Photoelectron Counts," IEEE J. Quantum Electronics, Vol. QE-1, No. 4, pp. 80-81, July 1965.
37. S. Karp and R. M. Gagliardi, "M-ary Poisson Detection and Optical Communications," NASA TN D-4623, June 1968.
38. C. W. Helstrom, "Detection and Quantum Mechanics," Inform. Contr., Vol. 10, pp. 254-291, March 1967.
39. C. W. Helstrom, "Detection Theory and Quantum Mechanics (II)," Inform. Contr., Vol. 13, p. 156, August 1968.
40. Jane Win-Shi Liu, "Reliability of Quantum Mechanical Communication Systems," Technical Report 468, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., December 31, 1968.
41. J. Dimeff, W. D. Gunter, Jr., and R. J. Hruby, "Spectral Dependence of Deep-Space Communications Capability," IEEE Spectrum September 1967.
42. T. E. Stern, "Some Quantum Effects in Information Channels," IEEE Trans. on Information Theory, Vol. IT-4, No. 1, p. 435, September 1960.
43. J. P. Gordon, "Quantum Effects in Communication Systems," Proc. IRE, Vol. 50, No. 11, p. 1898, November 1962.
44. C. W. Helstrom, "Quantum Limitations on the Detection of Coherent and Incoherent Signals," IEEE Trans. on Information Theory, Vol. IT-11, No. 4, pp. 482-490, October 1965.
45. J. I. Bowen, "The Effect of Attenuation on the Capacity of a Photon Channel," IEEE Trans. on Information Theory, Vol. IT-14, No. 1, pp. 32-46, January 1968.
46. C. W. Helstrom, "The Minimum Variance of Estimates in Quantum Signal Detection," IEEE Trans. on Information Theory, Vol. IT-14, No. 2, pp. 234-242, March 1968.

47. J. P. Gordon, "Noise at Optical Frequencies; Information Theory," in Quantum Electronics and Coherent Light, P. Miles (ed.). Academic Press, New York, 1964, pp. 156-182.
48. J. R. Meyer-Arendt and C. B. Emmanuel, "Optical Scintillation; A Survey of the Literature," Technical Note No. 225, National Bureau of Standards, Washington, D. C., April 5, 1965.
49. A. A. A. Saleh, "An Investigation of Laser Wave Depolarization Due to Atmospheric Transmission," IEEE J. Quantum Electronics, Vol. QE-3, No. 11, pp. 540-543, 1967.
50. V. I. Tatarski, Wave Propagation in a Turbulent Medium. McGraw-Hill Book Co., New York, 1961, Chap. 12.
51. L. A. Chernov, Wave Propagation in a Random Medium (translated by R. A. Silverman). McGraw-Hill Book Co., New York, 1960.
52. J. W. Strohbehn, "Comments on Rytov's Method," J. Opt. Soc. Am., Vol. 58, No. 1, pp. 139-140, January 1968.
53. W. P. Brown, Jr., "Validity of Rytov Approximation," J. Opt. Soc. Am., Vol. 57, No. 12, pp. 1539-1543, December 1967.
54. D. A. DeWolf, "Saturation of Irradiance Fluctuations Due to Turbulent Atmosphere," J. Opt. Soc. Am., Vol. 58, No. 4, pp. 461-466, April 1968.
55. G. R. Heidbreder, "Multiple Scattering and the Method of Rytov," J. Opt. Soc. Am., Vol. 57, No. 12, pp. 1477-1479, December 1967.
56. V. I. Tatarski, "On Strong Fluctuations of Light Wave Parameters in a Turbulent Medium," Soviet Phys. - JETP, Vol. 22, No. 5, pp. 1083-1088, May 1966.
57. V. I. Tatarski, op. cit., Chap. 3, ref. 4.
58. J. L. Lumley and H. A. Panofsky, The Structure of Atmospheric Turbulence. John Wiley and Sons, Inc., New York, 1964.
59. A. M. Obukhov, "Structure of the Temperature Field in a Turbulent Flow," Izv. Akad. Nauk, SSSR, Ser. Geograf. Geofiz., Vol. 13, p. 58, 1949.
60. S. Corrsin, "On the Spectrum of Isotropic Temperature Fluctuations in an Isotropic Turbulence," J. Appl. Phys., Vol. 22, pp. 469-473, 1951.
61. V. I. Tatarski, op. cit., Chap. 10, ref. 4.
62. D. L. Fried, G. E. Mevers, and M. P. Keister, Jr., "Measurements of Laser-Beam Scintillation in the Atmosphere," J. Opt. Soc. Am., Vol. 57, No. 6, pp. 787-797, June 1967.
63. I. Goldstein, P. A. Miles, and A. Chabot, "Heterodyne Measurements of Light Propagation through Atmospheric Turbulence," Proc. IEEE, Vol. 53, No. 9, pp. 1172-1180, September 1965.
64. R. E. Hufnagel, "An Improved Model Turbulent Atmosphere," NAS Summer Study, Restoration of Atmospherically Degraded Images, 1966, Appendix 3.
65. R. B. Muchmore and A. D. Wheelon, "Frequency Correlation of Line-of-Sight Signal Scintillations," IEEE Trans. on Antennas and Propagation, Vol. AP-11, No. 1, pp. 46-51, January 1963.
66. A. H. Mikesell, A. A. Hoag, and J. S. Hall, "The Scintillation of Starlight," J. Opt. Soc. Am., Vol. 41, No. 10, pp. 689-695, October 1951.
67. B. G. Chatterton, "Optical Communication Employing Semiconductor Lasers," Technical Report 392, Lincoln Laboratory, M. I. T., Lexington, Mass., June 9, 1965.
68. J. E. Ehrenberg, "A Study of the Effects of Atmospheric Turbulence on Intensity Properties at  $6328 \text{ \AA}$  and  $1.15 \mu$ ," S.M. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, November 1967.



69. D. H. Hohn, "Effects of Atmospheric Turbulence on the Transmission of a Laser Beam at  $6328 \text{ \AA}$ , I-Distribution of Intensity," *Appl. Opt.*, Vol. 5, No. 9, pp. 1427-1431, September 1966.
70. R. L. Mitchell, "Permanence of the Log-Normal Distribution," *J. Opt. Soc. Am.*, Vol. 58, No. 9, pp. 1267-1272, September 1968.
71. D. L. Fried, "Aperture Averaging of Scintillation," *J. Opt. Soc. Am.*, Vol. 57, No. 2, pp. 169-175, February 1967.
72. P. Beckman, "Signal Degeneration in Laser Beams Propagated through a Turbulent Atmosphere," *Radio Science, J. of Research NBS/USNC-URSI*, Vol. 69D, No. 4, April 1965.
73. A. I. Kon and V. I. Tatarski, "Fluctuations of the Parameters of a Spatially Bounded Light Beam in a Turbulent Atmosphere," *Izvest. Vyss. Uch. Zav., Ser. Radiophys.*, Vol. 8, No. 5, pp. 870-875, 1965.
74. R. A. Schmeltzer, "Means, Variances, and Covariances for Laser Beam Propagation through a Random Medium," *Quart. Appl. Math.*, Vol. 24, pp. 339-354, 1967.
75. F. G. Gebhart, "The Log-Amplitude Mean Value for Laser Beam Propagation in the Atmosphere, with Applications for Optical Communications," Technical Report 2156-8, Electro Science Laboratory, The Ohio State University, Columbus, Ohio, 16 May 1968.
76. D. L. Fried, "Optical Heterodyne Detection of an Atmospherically Distorted Wavefront," *Proc. IEEE*, Vol. 55, No. 1, pp. 57-67, January 1967.
77. G. R. Ochs, ESSA Research Laboratory, to be published.
78. W. R. Hinchman and A. L. Buck, "Fluctuations in a Laser Beam over 9- and 90-Mile Paths," *Proc. IEEE*, Vol. 52, No. 3, pp. 305-306, March 1964.
79. V. I. Tatarski, *op. cit.*, Chap. 8, ref. 4.
80. D. L. Fried, "Laser Beam Scintillation in the Atmosphere," *J. Opt. Soc. Am.*, Vol. 57, No. 2, pp. 181-185, February 1967.
81. D. L. Fried, "Scintillation of a Ground-to-Space Laser Illuminator," *J. Opt. Soc. Am.*, Vol. 57, No. 8, pp. 980-983, August 1957.
82. D. L. Fried and J. D. Cloud, "Propagation of an Infinite Plane Wave in a Randomly Inhomogeneous Medium," *J. Opt. Soc. Am.*, Vol. 56, No. 12, pp. 1667-1676, December 1966.
83. D. L. Fried, private communication.
84. H. Plotkin, private communication.
85. R. S. Lawrence, "Experimental Results in Optical Waves," Notes for CU/ESSA Electromagnetic Propagation Course, 1968.
86. M. E. Gracheva and A. S. Gurvich, "Strong Fluctuations in the Intensity of Light Propagated through the Atmosphere near the Earth," *Izvest. Vuz. Radiofiz.*, Vol. 8, No. 4, pp. 717-724, 1965.  
M. E. Gracheva, "Investigation of Statistical Properties of Strong Fluctuations of Light Intensity in Atmospheric Layer near the Earth," *Izvest. Vuz. Radiofiz.*, Vol. 10, No. 6, pp. 775-787, 1967 (in Russian).
87. V. I. Tatarski, *op. cit.*, Chap. 13, ref. 4.
88. D. L. Fried, "Atmospheric Modulation Noise in an Optical Heterodyne Receiver," *IEEE J. Quantum Electronics*, Vol. QE-3, No. 6, p. 213, June 1967.
89. D. L. Fried, "Propagation of a Spherical Wave in a Turbulent Medium," *J. Opt. Soc. Am.*, Vol. 57, No. 2, pp. 176-180, February 1967.
90. G. E. Mevers, D. L. Fried, and M. P. Keister, Jr., "Experimental Measurements of the Character of Intensity Fluctuations of a Laser Beam Propagating in

- the Atmosphere," Technical Memorandum No. 252, IS-TM-65/146, Electro-Optical Laboratory, Space and Information Systems Division, North American Aviation, Inc., 1965.
91. A. Peskoff, "Theory of Remote Sensing of Clear-Air Turbulence Profiles," J. Opt. Soc. Am., Vol. 58, No. 8, pp. 1032-1040, August 1958.
  92. G. Keller, "Relation between the Structure of Stellar Shadow Band Patterns and Stellar Scintillation," J. Opt. Soc. Am., Vol. 45, pp. 845-851, October 1955.
  93. E. Ryznar, "Dependency of Optical Scintillation Frequency on Wind Speed," Appl. Opt., Vol. 4, No. 11, pp. 1416-1418, November 1965.
  94. D. C. Hogg, "On the Spectrum of Optical Waves Propagated through the Atmosphere," Bell System Tech. J., Vol. 42, pp. 2967-2969, November 1963.
  95. M. Subramanian and J. A. Collinson, "Modulation of Laser Beams by Atmospheric Turbulence," Bell System Tech. J., Vol. 44, pp. 543-546, March 1965.
  96. W. M. Protheroe, "Preliminary Report on Stellar Scintillation," Contrib. Perkins Observatory, Ser. 2, No. 4, p. 127, 1955; see comments in ref. 41.
  97. R. E. Hufnagel and N. Stanely, "Modulation Transfer Function Associated with Image Transmission through Turbulent Media," J. Opt. Soc. Am., Vol. 54, No. 1, pp. 52-61, January 1964.
  98. D. L. Fried, "Optical Resolution through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures," J. Opt. Soc. Am., Vol. 56, No. 10, pp. 1372-1379, October 1966.
  99. A. L. Buck, "Effects of the Atmosphere on Laser Beam Propagation," Appl. Opt., Vol. 6, No. 4, pp. 703-708, April 1967.
  100. E. Gehrels, manuscript in preparation.
  101. J. C. Owens, manuscript in preparation.
  102. J. C. Moldon, "Imaging of Objects Viewed through a Turbulent Atmosphere," Technical Report 469, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., March 3, 1969.
  103. Y. Kinoshita, T. Asakura, and M. Suzuki, "Autocorrelation of Gaussian-Beam Fluctuation Caused by a Random Medium," J. Opt. Soc. Am., Vol. 58, No. 8, pp. 1040-1047, August 1968.
  104. D. L. Fried, "Statistics of a Geometric Representation of Wavefront Distortion," J. Opt. Soc. Am., Vol. 55, No. 11, pp. 1427-1435, November 1965.
  105. G. R. Heidbreder and R. L. Mitchell, "Effect of a Turbulent Medium on the Power Pattern of a Wavefront-Tracking Circular Aperture," J. Opt. Soc. Am., Vol. 56, No. 12, pp. 1677-1684, December 1966.
  106. M. A. Tamny, "Heterodyne Measurements of Atmospheric Turbulence at  $6328 \text{ \AA}$ ," S.M. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, May 1968.
  107. R. Bruce Herrick and R. Meyer-Arendt, "Interferometry through the Turbulent Atmosphere at an Optical Path Difference of 354 m," Appl. Opt., Vol. 5, No. 6, pp. 981-984, June 1966.
  108. Robert L. Kurtz and James L. Hayes, "Experimental Measurements of Optical Angular Deviation Caused by Atmospheric Turbulence and Refraction," NASA TN D-3439, May 1966 (George C. Marshall Space Flight Center, Huntsville, Ala.).
  109. C. L. Wyman, "Angular Fluctuations of a Laser Beam Propagated through the Atmosphere," M.S. Thesis, Department of Electrical Engineering, University of Alabama, 1968.
  110. E. V. Hoversten, "The Atmosphere as an Optical Communication Channel," 1967 IEEE International Convention Record, Vol. 15, Pt. 11, pp. 137-145.

111. R. S. Kennedy and E. V. Hoversten, "On the Reliability of the Atmosphere as an Optical Communication Channel," *IEEE Trans. on Information Theory*, Vol. IT-14, No. 5, pp. 716-725, September 1968.
112. F. Moller, "Optics of the Lower Atmosphere," *Appl. Optics*, Vol. 3, No. 2, pp. 157-166, February 1964.
113. R. Sloan, J. H. Shaw, and D. Williams, "Infrared Emission Spectrum of the Atmosphere," *J. Opt. Soc. Am.*, Vol. 45, No. 6, pp. 455-460, June 1955.
114. B. Reiffen and H. Sherman, "An Optimum Demodulator for Poisson Processes; Photon Source Detectors," *Proc. IEEE*, Vol. 51, No. 10, pp. 1316-1320, October 1963.
115. I. Bar David, "Communication under Poissonian Regime," submitted for publication to *IEEE Trans. on Information Theory*.
116. C. W. Helstrom, "The Detection and Resolution of Optical Signals," *IEEE Trans. on Information Theory*, Vol. IT-10, No. 4, pp. 275-287, October 1964.
117. S. Karp and R. M. Gagliardi, "M-ary Poisson Detection and Optical Communications," NASA TN D-4623, June 1968.
118. W. N. Peters and R. J. Arguello, "Fading and Polarization Noise of a PCM/PL System," *IEEE J. Quantum Electronics*, Vol. QE-3, No. 11, pp. 532-539, November 1967.
119. R. O. Harger, "Maximum Likelihood and Optimized Coherent Heterodyne Receivers for Strongly Scattered Gaussian Fields," to appear in *Optica Acta*.
120. R. O. Harger, private communication.
121. D. A. Ohori, "Detection of Signals in the Turbulent Atmospheric Optical Channel," S.M. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, February 1968.
122. S. J. Halme, "On Optimum Reception through a Turbulent Atmosphere," Quarterly Progress Report No. 88, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., January 15, 1968, pp. 247-254.
123. R. F. Lucy, K. Lang, C. J. Peters, and K. Duval, "Optical Superheterodyne Receiver," *Appl. Opt.*, Vol. 6, No. 8, pp. 1333-1342, August 1967.
124. J. Randall, private communication.
125. C. W. Helstrom, "Detectability of Coherent Optical Signals in a Heterodyne Receiver," *J. Opt. Soc. Am.*, Vol. 57, No. 3, pp. 353-361, March 1967.
126. D. L. Fried and R. A. Schmeltzer, "The Effect of Atmospheric Scintillation on an Optical Channel - Laser Radar and Binary Communications," *Appl. Opt.*, Vol. 6, No. 10, pp. 1729-1737, October 1967.
127. S. Gardner, "Some Effects of Atmospheric Turbulence on Optical Heterodyne Communications," 1964 IEEE Convention Record, Pt. 6, pp. 337-342.
128. R. D. Rosner, "Heterodyne Detection of an Optical Signal after One-Way Propagation on an Atmospheric Path," *Proc. IEEE*, Vol. 56, No. 1, pp. 126-128, January 1968.
129. J. P. Moreland, "Optical Heterodyne Detection of a Randomly Distorted Signal Beam," Ph.D. Thesis, Ohio State University, 1969.
130. D. M. Chase, "Power Loss in Propagation through a Turbulent Medium for an Optical-Heterodyne System with Angle Tracking," *J. Opt. Soc. Am.*, Vol. 56, No. 1, pp. 33-44, January 1966.
131. J. H. Shapiro, "Improved Optical Communication through Atmospheric Turbulence Using State Knowledge," Quarterly Progress Report No. 92, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., January 15, 1969, pp. 315-321.
132. G. F. Montgomery, "Intermittent Communication with Fluctuating Signal," *Proc. IEEE*, Vol. 45, No. 12, pp. 1678-1684, December 1957.

133. J. N. Pierce, "Ultimate Performance of M-ary Transmissions on Fading Channels," *IEEE Trans. on Information Theory*, Vol. IT-12, No. 1, pp. 2-5, January 1966.
134. J. V. Evans and T. Hagfors, Radar Astronomy. McGraw-Hill Book Co., New York, 1968.
135. R. S. Kennedy, Fading Dispersive Communication Channels. John Wiley and Sons, Inc., New York, 1969.
136. P. Bello, "Characterization of Randomly Time-Variant Linear Channels," *IEEE Trans. on Communication Systems*, Vol. COM-4, pp. 360-393, December 1963.
137. R. G. Gallager, "Characterization and Measurement of Time- and Frequency-Spread Channels," Technical Report 352, Lincoln Laboratory, M. I. T., Lexington, Mass., April 30, 1964. (Defense Documentation Center AD 443715.)
138. H. M. Heggstad, "Optical Communication through a Multiple Scattering Media," Technical Report 472, Research Laboratory of Electronics, M. I. T., Cambridge, Mass., November 22, 1968; also Technical Report 454, Lincoln Laboratory, M. I. T.
139. H. C. Van de Hulst, Light Scattering by Small Particles. John Wiley and Sons, Inc., New York, 1957.
140. P. J. Wyatt, "Light Scattering from Objects with Spherical Symmetry," *Proc. Interdisciplinary Conference on Electromagnetic Scattering*, Milton Kerker (ed.). Pergamon Press, New York, 1963, p. 25.
141. S. Levine and M. Kerker, "Scattering of E. M. Waves from Two Concentric Spheres when Outer Shell Has a Variable Refractive Index," *Proc. Interdisciplinary Conference on Electromagnetic Scattering*, *op. cit.*, p. 37.
142. S. D. Kastner, "Polarization in Multiple Scattering Using Random Stokes Vectors," *J. Quant. Spectros. Radiat. Trans.*, Vol. 6, pp. 317-324, 1966.
143. M. B. Wells, D. G. Collins, and K. Cunningham, Light Transport in the Atmosphere, Monte Carlo Studies, Ecom. Report No. 00240-1, Vol. I, Ecom. Fort Monmouth, New Jersey, 1966.
144. K. L. Coulson, J. V. Dave, and Z. Sekera, Tables Related to Radiation Emerging from a Planetary Atmosphere with Rayleigh Scattering. University of California Press, Berkeley, Calif., 1960.
145. G. V. Rozenberg and G. I. Gorchakov, "Correlation of the Optical Properties of Finely Dispersed Hazes," *Izvestiya Atmosphere and Oceanic Physics*, Vol. 3, No. 6, June 1967.
146. G. V. Rozenberg, "The Properties of an Atmospheric Aerosol from Optical Data," *Izvestiya Atmosphere and Oceanic Physics*, Vol. 3, No. 9, 1967.
147. W. M. Irvine and J. B. Pollack, "Infra-red Optical Properties of Water and Ice Sphere," *Icarus*, Vol. 8, p. 324, 1968.
148. H. H. Blau and R. Espinola, "Spectral Properties of Clouds from 2.5  $\mu$  to 3.5  $\mu$ ," *Appl. Opt.*, Vol. 7, p. 10, 1968.



# Appendix A

## Abstracts of Briefing Sessions

### NASA Optical Communication Program

H. L. Anderton

Two of the responsibilities of NASA Headquarters are (i) to develop a coordinated program using the best talent in the field centers, and (ii) to obtain financial support from the United States Congress for the program. To get the needed financial support requires a careful presentation of the technical work of the field centers in language that makes sense to the nontechnical members of Congress and still gives an accurate picture of the work that is to be done. The condensation of material that results from these attempts sometimes leads to extremely simplified statements that might lead a dedicated scientist to wonder whether the work was being adequately covered and, especially, whether the proper caveats were being included. (Some examples of Congressional presentations were shown.)

In spite of cutbacks in the over-all NASA budget for 1969 and subsequent years, there are indications that the optical communication technology area will be given increasing emphasis and funding. The rationale for this is the great promise afforded by the laser for increased data rates in space communication and the need for a large Space telescope for astronomical observations. (Typical funding figures were shown.)

### High-Capacity Optical Communication System

T. Kinsel

The large potential bandwidths available at optical frequencies suggest the possibility of an optical communication system capable of operating at very high data rates. Active optical devices that are available at present have relatively limited bandwidths, however. The purpose of this session is to describe a design of the transmitting and receiving terminals that circumvent the usual limitation of available optical devices.

The basic single-channel system operates in the binary pulse code modulation (PCM) mode with the binary states represented by the presence and absence of optical energy. A phase-locked laser driven by an external clock frequency is used as the source and an optical modulator, operating as a gate, transmits or rejects individual optical pulses. The performance of an experimental system utilizing a helium-neon laser operating at a bit rate of 224 M bit/sec, with the modulator constructed of lithium tantalate, is described.

A first stage for increasing the information capacity of the system uses time domain by time division multiplexing. Since the optical pulses from a phase-locked laser are generally quite narrow relative to their period, a number of independent channels can

be interleaved in time -4 if a Helium-Neon laser is used, and 32 if an Nd:YalG laser is used. At the present time, higher ultimate data rates are obtainable if this process is done in the optical domain, by using a device called a sinusoidal polarization switch. This device is constructed with an electro-optic element driven by a sinusoidal timing signal, and systematically varies the states of polarization of a series of pulses which is coupled through it. This device is used in both multiplexing and demultiplexing, and the design of a 24-channel time-multiplexed system will be sketched. Although the data rate of such a system is approximately 6.7 G bit/sec, each active optical device is operating within its bandwidth. The performance of a two-channel time-multiplexed system using a helium-neon laser will be described.

A further stage of multiplexing may be done in the frequency domain, by using techniques of frequency division multiplexing. Here, independent transmitting terminals, which may be time-multiplexed, would be constructed at different carrier frequencies, by using different laser sources or by parametric oscillation. The outputs would be combined into a single polychromatic beam.

A final stage of multiplexing may be done in the spatial domain, by using a multi-mode beam waveguide. One way of doing this would be to launch a number of Gaussian beams into the waveguide with different launching parameters. At the receiver plane these would be resolved as separate beams.

Combining all three stages of multiplexing simultaneously indicates the possibility of constructing, with available or nearly available devices, an optical communication system operating at  $10^{14}$  bits/sec.

### Direct-Detection Communication Systems

J. R. Kerr

Direct-detection optical communication systems will be signal-shot-noise limited when large (bandwidth  $\times$  SNR) products are required, provided that detectors-with-gain are available and background levels are not too high. Thus, with the development of microwave-bandwidth photomultipliers and avalanche photodiodes, direct detection has a clear advantage over heterodyne systems at visible and near-infrared wavelengths. This is especially true when system complexity and atmospheric effects are considered. For instance, lasers operating in the FM mode comprise ideal transmitters for such systems, and single-frequency operation is unnecessary. Alternatively, mode-locked lasers represent simple pulsed sources for optical PCM applications. Studies show that the most attractive system alternatives, at present, are (i) direct detection with doubled-YAG or Argon laser transmitters, and (ii) heterodyne detection at 10.6- $\mu$  wavelengths with stabilized CO<sub>2</sub> lasers. We must recognize, however, that (SNR bandwidth) characterizations of these alternatives may be inadequate from a communication theory standpoint, and more sophisticated configurations may evolve.

Optical Heterodyne Communications Experiments at 10.6  $\mu\text{m}$ 

F. E. Goodwin

A 10.6- $\mu\text{m}$  optical heterodyne communication system that is in operation over a 30-km (18 mile) path was described. Television signals have been transmitted over the beam by the use of optical FM, achieved with internal electro-optic phase modulation of the transmitter oscillator. A frequency deviation of 3 MHz is thus obtained for a driving voltage of 1 kV, as compared with the fundamental frequency noise of the system of  $\sim 1$  kHz. For clear-day operation, average carrier-to-noise ratios at the receiver approach 60 dB, and demodulated video signal-to-noise ratios average 50 dB.

Optical propagation experiments reveal amplitude scintillation noise in excess of 20 dB, represented by fluctuations of +10 dB and -10 dB from the average carrier level. Other observations show that the 10.6- $\mu\text{m}$  beam has a remarkable ability to penetrate fog. Data related to scintillation, visibility, temperature, humidity, and optical absorption were presented.

Gas Laser Sources

W. B. Bridges

In this presentation the salient characteristics and state of the art of Helium-Neon, Argon-ion, and carbon dioxide lasers are reviewed. A comparison is made on the basis of a figure of merit for ground use of the form  $(\text{Power})/(\text{Wavelength})^2$  and  $((\text{Power}) \times (\text{efficiency})^2)/((\text{Wavelength})^2 \times (\text{length}))$  for spacecraft use. The best performance figures for the three laser types in multimode, TEM<sub>00</sub>, and single-frequency operation are given for laboratory, commercial, mil-spec and space-qualified units. Specific examples of lasers under development at the Hughes Aircraft Company are described: a space-qualified Helium-Neon laser, an airborne Argon-ion laser, and a high-power pulsed ground-based CO<sub>2</sub> laser oscillator-amplifier system.

Solid-State Lasers

J. E. Geusic

The solid-state laser is now quite competitive with its gaseous counterparts. Continuous operation at output powers over 100 W has been achieved with Nd:YAlG at efficiencies as high as 1.5%. Laser efficiency and life (1000-5000 h) is limited by the tungsten-iodine pump lamps in present use, and it is expected that employment of luminescent semiconductor pumps now under development will raise the efficiency by a factor of at least five and permit even longer life operation. The solid-state Nd:YAlG laser is quite versatile in operation because, in addition to cw operation, it can be operated in a continuously pumped repetitively Q-switched mode (peak power >1 kW, average power >1 W, pulsewidth 0.1  $\mu\text{sec}$ , pulse repetition rate  $10^3$ - $10^4$  p. p. s) or a mode-locked mode



(pulsewidth  $30 \times 10^{-12}$  sec). The Nd:YAIG laser output, which is at  $1.06 \mu$ , can be converted with 100% efficiency to  $0.53 \mu$  by utilizing the nonlinear barium sodium niobate ( $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ ) inside of the laser cavity. The use of barium sodium niobate with the YAIG laser provides a practical solid-state green coherent source. Wavelength tunable coherent sources are now under development, with barium sodium niobate used as the nonlinear element of a parametric oscillator. Such a continuous, tunable oscillator has been built in the near infrared, with the doubled YAIG green source as pump, and oscillation from  $0.95 \mu$  to  $1.15 \mu$  has been demonstrated. Continuous parametric oscillators covering the entire visible and near infrared are now practically achievable.

### Optical Modulators

J. E. Geusic

Various physical phenomena can be utilized for modulating an optical beam. The electro-optic effect, the magneto-optic effect and the acousto-optic effect are three phenomena that have been seriously considered for high-speed optical modulation. Of these the electro-optic effect has been studied and developed most extensively, and there are good theoretical reasons to believe that in principle it has the best potential. For example, in comparing the electro-optic effect and the acousto-optic effect it is evident that the electro-optic effect has greater potential, since the velocity of light is approximately  $10^5$  times greater than the velocity of sound.

The analysis of electro-optic modulator devices is quite straightforward. One result of this analysis is that (in an optimally designed modulator) electrical power that is required to modulate light is proportional to the modulation bandwidth  $B$ , proportional to the wavelength cubed, and inversely proportional to the square of the electro-optic coefficient (under the assumption of a linear effect). For most electro-optic materials, the electro-optic coefficient is only a weak function of wavelength over the visible to near infrared, and one general conclusion is that it is easier with most materials to build modulators in the visible range than it is at longer wavelengths. At present, the best electro-optic materials are the niobates ( $\text{LiNbO}_3$ ,  $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ -----). With these materials, modulation bandwidths of approximately 500 MHz have been achieved in the visible, with the expenditure of  $\sim 1$  mW/MHz of bandwidth. The electrical  $Q$ 's of the materials are generally 100 or less, and this imposes a practical maximum in the power that can be delivered to the modulator, for the finite  $Q$ 's cause heating effects that produce optical nonuniformities. At present, modulation bandwidths of several GHz in the visible are not practically achievable with present materials, and in the infrared, particularly at  $10 \mu$ , even 0.5-GHz bandwidths are unpractical. New and better electro-optic materials are definitely needed and the improvements can be in the form of higher electro-optic coefficients and/or higher electrical  $Q$ 's.

Tracking in Optical Communication Systems

E. C. Fraser

Three different problems are involved in the operation of an optical communication system: pointing – aiming on the basis of a priori information before acquisition is accomplished; following – maintaining the sensitive axis aligned with the direction of the incoming signal; and data collection – for orbit or trajectory determination. To accomplish these tasks, several fundamentally different system configurations have been considered. An open-loop system points each of the terminals on the basis of externally supplied pointing information. Such a system is therefore always operating in an acquisition mode and, therefore, cannot take advantage of techniques such as beam narrowing following lock-on. In a closed-loop system, once acquisition has been achieved by open-loop pointing, each terminal determines the pointing error of the other terminal's transmitter, and relays this information back to institute pointing corrections. Pointing errors can be detected by such techniques as receiver arrays or scanning of a spatially encoded beam. The principal difficulty with closed-loop systems is their inability to respond to rapid disturbances or medium fluctuations that results from the long transport delay associated with transmission paths of interest (approximately 1/3 sec round trip to a synchronous satellite and approximately 16 min round trip to a spacecraft in the vicinity of Mars). In a cooperative system each terminal's transmitter is used as a beacon at which the other terminal can aim. Thus all tracking by each terminal is done only at that terminal, and maximum tracking performance can be achieved. It appears, therefore, that for most applications a cooperative-system configuration will be most appropriate.

Modern control theory dictates that for optimum tracking performance of each

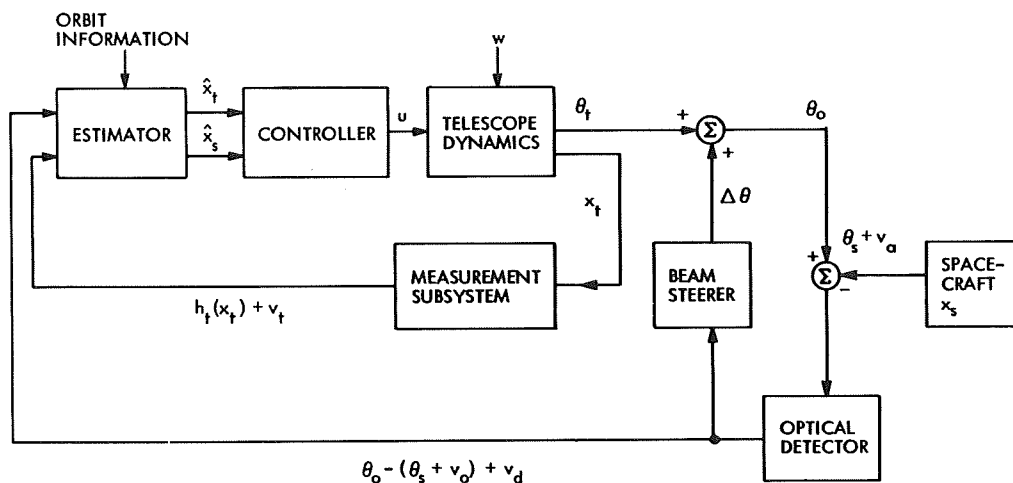


Fig. A-1. Optimum tracking control system.

terminal the tracking control system should take the form shown in Fig. A-1. The purpose of the estimator is to develop, using algorithms evolving from statistical estimation theory, minimum variance (least noisy) estimates of the system state (both target and telescope) from all measurement data obtained from the system. These state estimates are then used by the controller to generate optimum (in a minimum mean square following error sense) control signals for driving the telescope. The high-speed beam steering loop shown following the telescope dynamics must be provided in those cases in which the field of view (or beamwidth) is smaller than components of the path disturbances that are too fast for the main telescope to follow. This situation arises, for example, if one were to attempt tracking through the atmosphere with a fractional second of arc field of view.

The primary feature of the estimator-controller system configuration is that it resolves the dilemma of tracking loop bandwidth selection by separating the tasks of tracking-signal processing and telescope-pointing control. To remove line-of-sight disturbances, the estimator can act as a very lowpass filter, accepting only those signal components pertinent to the target trajectory. On the other hand, the controller can independently realize a high-gain wide-bandwidth tracking loop so as to achieve crisp transient response and immunity to wind torques and other load disturbances.

The implementation of the complete estimator-controller system requires considerable computation capability and complex instrumentation. In many applications the system performance requirements are not sufficiently stringent to require this degree of sophistication. For these cases there are some simplifying approximations that result in systems of considerably reduced complexity. In fact, the conventional autotrack system configuration can be derived from the estimator-controller formulation for the appropriate set of systems specifications and approximations.

#### Précis of Presentation on Effects of Atmospheric Turbulence on Optical Communication

D. L. Fried

The following presentation covers a large part of our present quantitative knowledge concerning optical effects of atmospheric turbulence, particularly as related to communication equipment in the space-to-ground and ground-to-space mode of operation. The quantitative knowledge that we have is based almost entirely on theoretical calculations with a minimum of experimental backup. The theory generally takes the form of generalized solutions for instantaneous optical instrument performance in terms of instantaneous propagation conditions, from which the statistics of instrument performance are calculated in terms of the statistics of propagation. This is coupled with a solution for instantaneous optical instrument performance in terms of instantaneous propagation conditions, from which the statistics of instrument performance are calculated in terms of statistics of propagation. This is coupled with a solution for the

instantaneous propagation results in terms of instantaneous turbulence conditions, from which the statistics of propagation are computed in terms of the statistics of turbulence. The statistics of turbulence, of significance in optical propagation, are derived from the Kolmogorov theory of the inertial subrange, which asserts that the refractive-index structure function, that is, the mean-square difference of refractive index at two points at distance  $r$  apart, is proportional to  $r^{2/3}$ . From this, various propagation statistics are computed, such as the log-amplitude variance, the phase and wave-structure functions, the mutual coherence, etc. Fairly good evidence exists for the validity of the Kolmogorov theory in the inertial subrange in the atmosphere. Data for substantiating this for the lower atmosphere were presented in a series of slides.

A variety of calculations for log-amplitude covariance and wave-structure function for infinite plane-wave and spherical-wave sources was presented in a series of slides, all based on the Kolmogorov hypothesis for atmospheric turbulence. The results are all derived by using the Rytov approximation. The most pronounced feature of the log-amplitude covariance functions is that they go to zero in a distance of the order of the square root of wavelength times path length. For vertical propagation down through the atmosphere, based on our limited knowledge of the vertical distribution of turbulence in the atmosphere, the equivalent path length is approximately 3200 m.

Next, our understanding concerning the performance of a laser transmitter operating in a ground-to-space mode was presented. Calculations, thus far, make available estimates of scintillation of transmitter antenna gain as a function of diameter. Mean transmitter antenna gain is available from some work by F. Gebhart to be published shortly in the Journal of the Optical Society of America. (Not included in this presentation, but a significant result derived during the period of the conference, is the fact that transmitter antenna gain on an instantaneous basis can be shown to be identical to heterodyne receiver gain operating over the same path. Since heterodyne receiver gain is a well-studied quantity, this provides us with a substantial body of knowledge applicable to transmitter antenna gain.)

Very simple results for aperture averaging were presented in a single slide. The results show that aperture averaging may be considered as the consequence of a set of independently scintillating regions on the collection aperture, each region having an area of approximately wavelength times path length.

The problem of general interpretation of wavefront distortion was discussed from the point of view of whether or not any of the distortion features have recognizable and potentially useful aspects. It was pointed out that a large part of the wavefront distortion can be interpreted as a linear variation of complex phase. To the extent that the complex phase is mostly ordinary phase with very little log amplitude, this corresponds to a tilted wavefront, which can be "tracked out." In the absence of log-amplitude variance, tracking the tilt can increase coherent aperture diameter by a factor of approximately 8 dB. In the presence of strong log-amplitude effects, however, the gain is almost negligible.

Results were presented for long- and short-exposure imaging through the atmosphere. Pointed out were the possible gain of short-exposure imaging and the general tendency of resolution to increase uniformly with diameter until a critical diameter is reached, at which point the resolution saturates, that is, stops increasing with increasing diameter. The significance of these results is principally in terms of their relation to the optical heterodyne receiver, imaging resolution being directly proportional to antenna gain and thus to achievable signal-to-noise ratio in a heterodyne receiver. Equivalent results showing the same saturation effect were presented for the heterodyne receiver average signal-to-noise ratio. The knee of the curve, corresponding to the onset of saturation, is located at a diameter denoted by  $r_0$ . Graphs presenting values of  $r_0$  for various propagation paths were shown. Results concerning atmospheric modulation effects in heterodyne receivers were presented, indicating the very rapid, very large increase in modulation noise as the receiver diameter is increased beyond  $r_0$ . Doppler-shift effects were evaluated. It was shown that under all conceivable conditions, Doppler shifts were entirely negligible.

Finally, a series of curves relating binary channel performance in the presence of scintillation to performance in the absence of scintillation. The calculations for these curves are based on the fairly well-established fact that atmospheric scintillation almost always appears to show a log-normal distribution. Significant loss of effective signal-to-noise ratio was found for fairly moderate levels of scintillation.

#### Effect of Atmospheric Turbulence:

#### Experimental Results from the ESSA Research Laboratories

R. S. Lawrence

The optical effects of turbulence are measured with mobile equipment so that a variety of paths can be instrumented. Paths in the Boulder (Colorado) area ranging in length from 25 m to 150 km have been used. Whenever practical, small, high-speed resistance thermometers are used to provide simultaneous direct measurements of the magnitude of the refractive-index variations. The thermometers have a thermal capacity much lower than that of  $1 \text{ mm}^3$  of air and a frequency response exceeding 1 kHz.

Beam spread observed on paths 5-15 km in length is normally  $\sim 10 \mu\text{rad}$  at night, and often exceeds  $100 \mu\text{rad}$  at midday. For short paths, where the turbulence is uniformly distributed, intensity scintillations display the log-amplitude variance and covariance expected from Tatarski's theory. On longer paths, where the height above ground is greatest near the center of the path, log-amplitude variance is reduced and the covariance function is distorted in qualitative agreement with the concept of reduced turbulence at midpath. Quantitative comparisons with the results from airborne temperature sensors are being made to test the usefulness of the covariance function as a measure of the distribution of turbulence along the path.

The saturation effect, where the intensity of scintillations fails to increase with

increasing strength of turbulence, has been demonstrated to begin at a log-amplitude variance of 0.6 on a 1-km path, 2 m above flat ground.

Temporal variation of phase over paths shorter than 100 m is measured with a reversible fringe-counting Michelson interferometer. Resulting power spectra contain only approximately 1% as much bandwidth as the corresponding intensity fluctuations. Measurements of modulation phase over paths of several kilometers show that the usual rms fluctuation of 1-sec averages of phase path relative to the 30-sec moving average is found to be  $3 \times 10^{-9}$  of the total path.

### Statistical Communication Theory at Optical Frequencies

E. V. Hoversten

This talk had several purposes: (i) to discuss the role of statistical communication theory and indicate the modeling questions that must be answered before the theory can be applied, (ii) to discuss the differences between the communication models and problems at optical and microwave frequencies, and (iii) to illustrate the ideas in the first two topics in terms of the turbulence atmosphere channel and discuss some specific results for this channel.

The abilities to establish ultimate performance limits and to provide insight into how to build 'good' systems, whose performance approaches these limits, were cited as the contributions of statistical communication theory. The types of questions that the theory can answer, for example, optimum receiver structure, error performance and capacity, and the need, if the full power of the theory is to be realized, for complete statistical models, particularly of the channel, background noise, and receiver processing, were discussed. The differences between optical and microwave communication that were considered included the importance of the spatial aspect at optical frequencies, the background noise at optical frequencies, the different noises associated with the detector processing, that is, the quantum aspect at optical frequencies, the different channel statistics encountered, and the differences in the operations that can be performed with available devices.

An approximate statistical model was discussed for the turbulent atmospheric channel and the ideas above were illustrated in terms of this discussion. The form of the optimum processing for heterodyne detection with a large receiving aperture was considered. This led to a discussion of diversity receivers and the optimum use of spatial diversity. The discussion was in terms of digital signals whose duration was short relative to the channel coherence time. The error performance that could be achieved with the above-mentioned processing for M-ary orthogonal signaling sets was discussed. It was indicated that the error probability had the form

$$\tilde{B}_2^{-KE} \leq \Pr[\epsilon] \leq 2^{-KE},$$

where K is the number of information digits per transmitted waveform, and E is the

system reliability. The dependence of  $E$  on the aperture area, turbulence, and energy-to-noise ratio per diversity path was discussed. The results which were considered are discussed in section 4.3.2 of this report.

### Optical Propagation through Clouds

A. Holland

The light reflected or transmitted by fog, haze or clouds depends on the number and size of the water droplets or particles present, on the absorption and scattering cross sections of the particles and on the angle of incidence of the incoming radiation. The scattering or absorption cross sections of individual droplets of water can be calculated, exactly by using classical Mie theory if the particle diameter and refractive index are known. These "monochromatic" single-particle properties can be integrated over a particle size distribution (if known) to give the scattering properties of a volume element of the medium. For most cases of interest these monochromatic properties can be used to represent the medium under moderately broadband illumination (under the assumption that the complex index of refraction does not vary rapidly with wavelength). Then for optically thin media, in which only single scattering is important, the reflected and transmitted intensities are simple functions of the optical thickness and transfer geometry.

On the other hand, for optically thick media for which multiple scattering must be considered, the transfer problem is extremely complicated. With the exception of a few special cases, solutions to the exact equation of transfer are difficult to obtain. It is possible, however, to model multiple-scattering problems by using a random-walk or Monte Carlo technique. The method has been applied recently to the problem of multiple scattering by clouds and fog.

We present a review of the known properties of haze, fog, and clouds, an outline of Mie theory and the equations of radiative transfer, and finally a brief discussion of some computer results for haze and clouds. The entire discussion is restricted to unpolarized light — a scalar theory. Consideration of polarization complicates the discussion slightly, and for that reason it has been neglected here. There is no doubt, however, that polarization should be accounted for in any multiple-scattering calculation, since, in almost any real medium, an initially unpolarized beam of light will be partially polarized after a single scattering event.

### Quantum Detection Theory

C. W. Helstrom

Detection theory evaluates communication receivers on the basis of the average probability  $P_e$  of error and seeks that receiver structure for which  $P_e$  is minimum. In a binary communication system the receiver must decide periodically, on the basis of its

input, between two hypotheses  $H_0$  and  $H_1$ . When the input can be described by classical physics, the optimum receiver bases its decision on the likelihood ratio. When, as in optical communications, it is necessary to describe the input data quantum-mechanically, the receiver is required to choose which of two density operators  $\rho_0$  and  $\rho_1$  best represents the input optical field. It does so by measuring a projection operator,  $\Pi$ . If the outcome of the measurement is 0, hypothesis  $H_0$  is chosen; if it is 1,  $H_1$  is chosen. The best projection operator is the one for which the average probability of error is minimum; it can be expressed in terms of the eigenvectors of the operator  $(1-\xi)\rho_1 - \xi\rho_0$ , where  $\xi$  is the prior probability of hypothesis  $H_0$ . The detection of a coherent signal in the absence of any noise was discussed; it can be represented as a choice between two coherent states. The average error probability so attained is lower than that for either counting the number of photons or for using the "classical" detection operator, which corresponds to matched filtering. The detection of a coherent signal with random phase occupying many modes of the receiver can also be treated; the error probability is expressed in terms of the Laguerre distribution. In instances when the optimum detection operator is difficult to calculate, a threshold detector, which is its limiting form at low signal-to-noise ratio, can often be specified.

#### References

1. "Detection Theory and Quantum Mechanics," Inform. Contr. 10, 254-291 (1967).
2. "Detection Theory and Quantum Mechanics (II)" (to appear in Information and Control).
3. "Fundamental Limitations on the Detectability of Electromagnetic Signals," Int. J. Theoret. Phys., Vol. 1, No. 1, May 1968.
4. "Minimum Mean-Squared Error of Estimates in Quantum Statistics," Phys. Letters 25A, 101-102 (July 31, 1967).
5. "The Minimum Variance of Estimates in Quantum Signal Detection," IEEE Trans. on Information Theory, Vol. IT-14, pp. 234-242, March 1968.

#### Optical-Communication Space Experiments

J. L. Randall

The following subjects were discussed: (i) an Optical Guidance Sensor system for automatic (unmanned) rendez-vous and docking of two spacecraft, (ii) a precision pointing and tracking system for laser communication, (iii) a 30 megabit/sec optical communication system, (iv) a 10.6- $\mu$  coherent communication system, and (v) OTES (Optical Technology Experiment System) proposed optical-communication experiments.

The rendez-vous and docking guidance system utilizes a GaAs laser and incoherent diode as light sources and an image dissector for tracking to provide range, range rate, angle, and angle rate which are necessary to perform the rendez-vous and docking maneuver. The system is light-weight and requires low power. The system is being built by ITT for NASA.



The precision pointing and tracking system for optical communication is a prototype of an optical communication system that is being tested to study the problems of space communication with emphasis on the pointing output. Thus far, all problems appear solvable insofar as pointing is concerned. This system was built for NASA by Perkin-Elmer.

The 30-megabit PCM optical communication system uses a 6328 Å He-Ne laser with incoherent detection. It uses a binary bit code and operates either on-off or in a dual channel capacity where the ones and zeros are marked by two circular states of polarization. Error rates per bit have been measured under varying conditions. Results indicate that the system operates near that which can be theoretically predicted with the Poisson statistics of photoelectrons. The system was built for NASA by ITT.

The 10.6- $\mu$  communication system utilizes coherent detection by mixing the signal with a 10.6- $\mu$  local oscillator. The signal is impressed on the transmitter laser by frequency-modulating the laser cavity with a piezoelectric transducer. Two voice channels are transmitted in the way. The system operates within about a factor of two of that predicted. That is, the measured value and calculated value of  $S/N$  agree within a factor of two. The system was built by Honeywell for NASA.

A brief discussion of some of the space experiments that are being studied on the OTES was held. These experiments under consideration have as a goal the solving of all of the practical problems associated with operating an optical communication system in space. This would include communication through the atmosphere and from ground to space. The results of these experiments would be the proof of the feasibility of optical communication in space and to provide the data necessary to design communication systems for mission applications.

## Appendix B

### Designing "Real" Optical Communication Systems

T. Kinsel

A communication theorist can generally derive a block diagram for an optimum communication system designed to operate with a specific transmission channel. This block diagram will specify various operations to be performed on the signal waveform. The problem considered here is the embodiment of such a block diagram either with existing devices or physically possible devices. For an optical communication system a useful way for the experimentalist to attack this problem is to consider the location of the interfaces where information conveyed by electrons is converted to photons or vice versa. There are two such interfaces, one at the transmitter and one at the receiver. It may be asserted that the location of these interfaces is variable, within limits, and their optimum location in a real system will depend upon the characteristics of the various devices. It would seem that the performance of "real" systems would generally be improved by separating the interfaces as much as possible. In other words, it is desirable to do as much of the signal processing by optical techniques as possible.

An example of a system that represents an attempt to maximize the separation of these interfaces is the time-multiplexed optical PCM system as described by the author. At the transmitter terminal the interface is located at the array of optical modulators which is near the front end of the transmitter. Following this the processing is done optically using active devices. At the receiving terminal the interface is located at the array of photodetectors. Before this, the process of recovering the separate channels is done optically.

Let us also consider the diversity-channel receiver suggested by Kennedy and Hoversten to operate with a turbulent channel. Here the input to the receiver has a distribution of phases and propagation directions. A block diagram of the optimum receiver is shown in Fig. B-1.

One design of a real receiver that performs the operations is shown in Fig. B-2. The input beam is mixed with a local oscillator whose output is a combination of plane waves. Various pencils of light in the input beam will then be selected and imaged onto an array of photodetectors. The photon-electron interface is located at this photodetector array which is near the front end of the receiver. Subsequent processing would be done electrically using either integrated circuit approaches or possibly bulk effects in solids.

An alternative approach that places the interface at a later stage of the receiver is shown in Fig. B-3. This approach is highly speculative and not fully analyzed. It is used to suggest possible approaches to the problem and possible areas of device research. The mosaic array of photodetectors is replaced by an optically nonlinear crystal. The number of diversity channels would be determined by the number of plane waves in the local oscillator. The output of the nonlinear crystal at the sum frequency

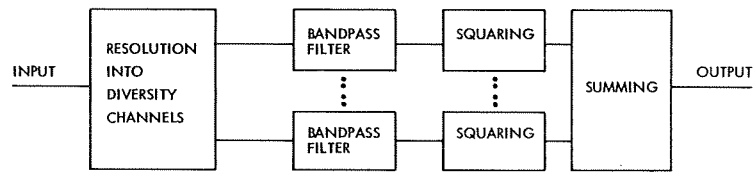


Fig. B-1. Idealized diversity receiver.

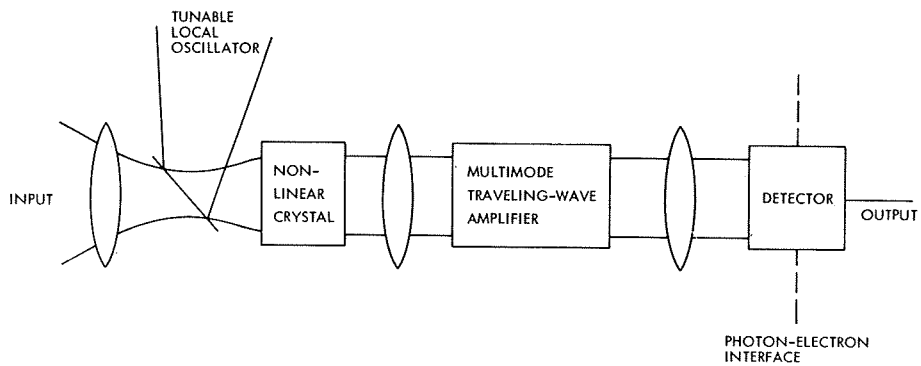


Fig. B-2. Discrete realization of the diversity receiver.

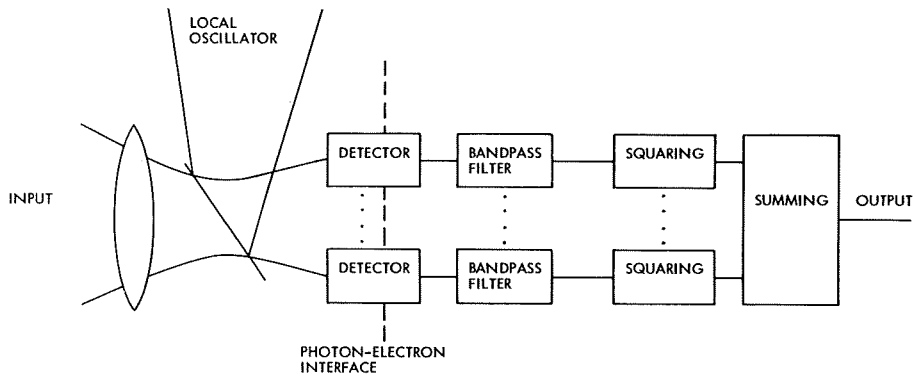
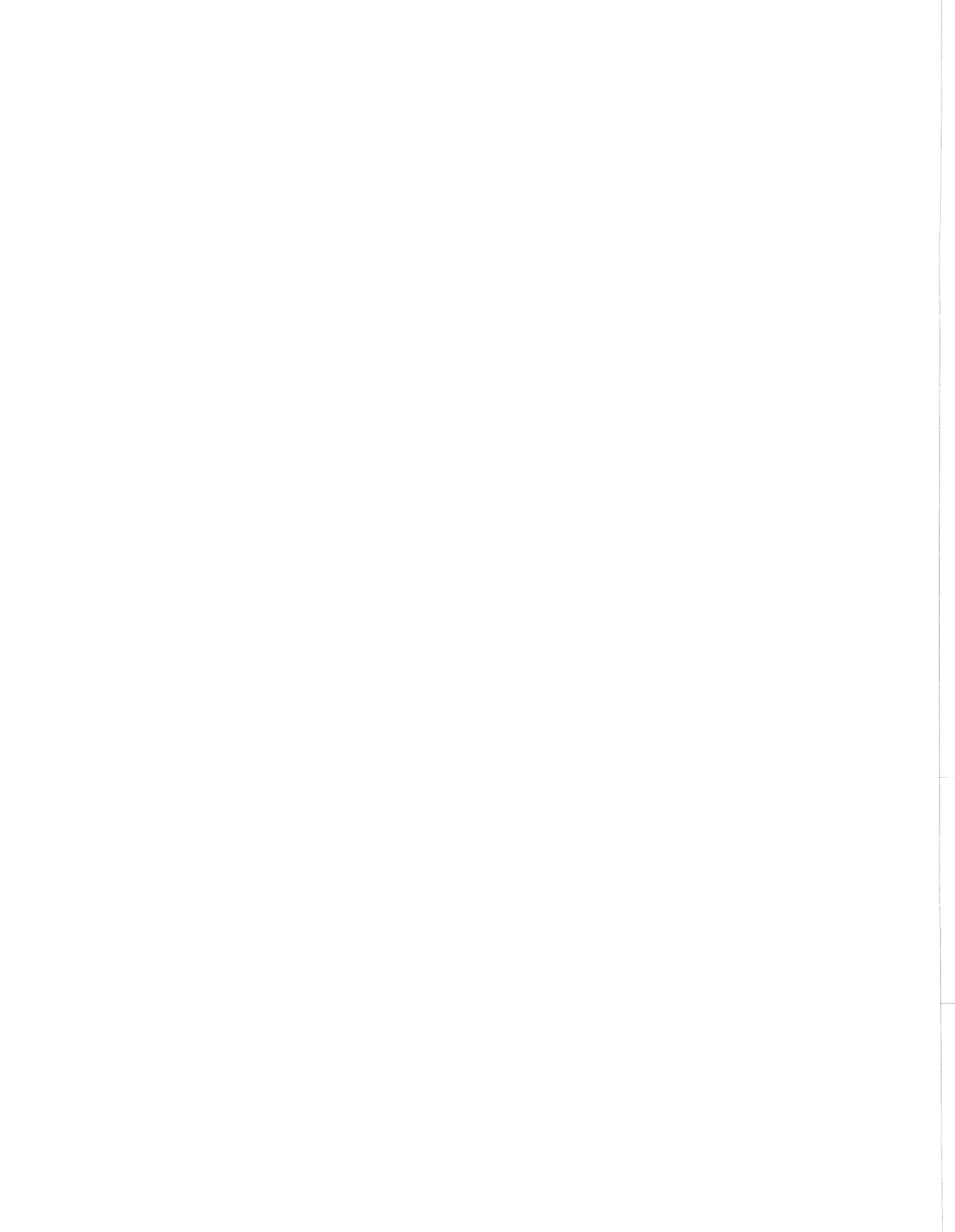


Fig. B-3. Distributed realization of the diversity receiver.

would be coupled into a multimode traveling-wave optical amplifier such as has been discussed by Geusic and Scovil. The bandpass filter would be supplied by the oscillation bandwidth of the gain medium. The output would be imaged onto a single photodetector which would provide the squaring process and sum over the different channels.

There are obviously many problems in actually constructing such a receiver and this discussion should serve only as an example of the types of signal processing in the optical domain that should be explored.



## Appendix C

### On the Reliability of Quantum-Mechanical Communication Systems

Jane W. S. Liu

A model of typical communication systems is shown in Fig. C-1. The input to the system in any given signaling interval of duration  $\tau$  is one of  $M$  input symbols denoted  $m_1, m_2, \dots, m_M$ . For simplicity, it is assumed that the  $M$  input symbols occur with equal probabilities. Thus the information rate of the system is  $R = \frac{\ln M}{\tau}$  nats/sec. Correspondingly, the output of the system in every signaling interval is also one of  $M$  symbols which is the estimation  $\hat{m}$  of the transmitted input symbol  $m$ . The receiver is designed so that the probability of error  $P(\epsilon)$  defined as

$$P(\epsilon) = \Pr [\hat{m} \neq m] \tag{1}$$

is minimum.

When the input symbol is  $m_j$ , the electromagnetic field generated by the signal source is in a state specified by the density operator  $\rho_j^t$ . Therefore, in any given signaling

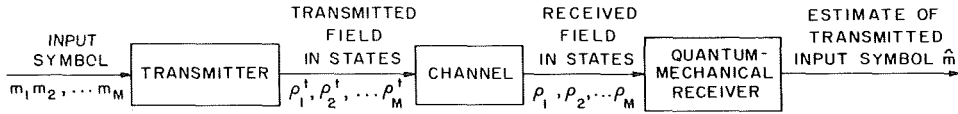


Fig. C-1. Quantum-mechanical model of a communication system.

interval, the signal field is in one of the  $M$  states specified by the density operators  $\rho_1^t, \rho_2^t, \dots, \rho_M^t$ . Because of the presence of the additive thermal noise field, the total received field at the receiver is in one of the  $M$  states specified by the density operators  $\rho_1, \rho_2, \dots, \rho_M$ .

Within the time interval during which the signal is expected to arrive, the receiver makes measurements on the received field and decides which one of the  $M$  input symbols is transmitted on the basis of the outcomes of the measurements. We shall assume that measurements made by the receiver are ideal.<sup>1</sup> An ideal measurement is one in which the state of the field immediately after the measurement depends only on the outcome of the measurement and not on the initial state of the field. It follows that the probability distribution of the outcome of any subsequent measurement on the field does not depend on the transmitted symbol. Therefore, any measurement made after an ideal measurement is useless in estimating the transmitted symbol.

An ideal measurement can be characterized by a complete set of measurement state vectors  $\{|x_{\underline{k}}\rangle\}$ , where  $x_{\underline{k}}$  denote  $K$ -tuples of parameters that are the possible outcomes of the measurement associated with the set  $\{|x_{\underline{k}}\rangle\}$ . [To simplify our discussions, we

assume that the parameters  $x_{\underline{k}}$ , as well as the eigenspectra of the density operators  $\rho_i$ , are discrete. It is evident that this assumption imposes no real restriction, since all discussions can be generalized by replacing the sums with integrals and probability distributions with their corresponding probability density functions.] When the field is in the state specified by the density operator  $\rho_i$ , the probability of the outcome being  $x_{\underline{k}}$  is given by

$$P(x_{\underline{k}} | m_i) = P(x_{\underline{k}} | \rho_i) = \langle x_{\underline{k}} | \rho_i | x_{\underline{k}} \rangle. \quad (2)$$

Clearly, measurements of dynamical variables that are represented by commutative Hermitian operators are ideal measurements. In this particular case, measurement state vectors associated with the measurement are just the simultaneous eigenstates of the commutative Hermitian operators; hence,  $\{|x_{\underline{k}}\rangle\}$  is not only a complete set but also an orthogonal one.

Hereafter, we shall specify the structure and discuss the performance of an optimum receiver when the density operators  $\rho_j$  are commutative. As an example, the performance of a system in which signal fields have known orthogonal classical waveforms but random phase is derived. The performance of the heterodyne receiver and photon counter are also included for comparison.

Another example of systems in which the density operators are commutative is one in which both the signal fields and the noise field are in completely incoherent states. A fading dispersive channel can thus be modeled quantum-mechanically. The optimum performance of the optimum equal-strength diversity system is derived in Eq. (3). We shall not discuss this system here.

### I. STRUCTURE AND PERFORMANCE OF AN OPTIMUM RECEIVER FOR EQUIPROBABLE SIGNALS REPRESENTED BY COMMUTATIVE DENSITY OPERATORS

Let us assume that the receiver makes an ideal measurement characterized by the set of measurement state vectors  $\{|x_{\underline{k}}\rangle\}$  on the received field. Then, the receiver uses a randomized strategy to estimate the transmitted symbol as  $m_j$  with probability  $p_{j\underline{k}}$  when the outcome of the measurement is  $x_{\underline{k}}$ . Clearly,

$$j = \sum_{j=1}^M p_{j\underline{k}} = 1; \quad \text{for all } \underline{k}.$$

It follows that the probability of error defined in Eq. (1) is

$$P(\epsilon) = 1 - \frac{1}{M} \sum_{j=1}^M \sum_{\underline{k}} p_{j\underline{k}} \langle x_{\underline{k}} | \rho_j | x_{\underline{k}} \rangle. \quad (3)$$

In general, it is difficult to find the set  $\{|x_{\underline{k}}\rangle\}$  and probabilities  $p_{j\underline{k}}$  that minimize the right-hand side of Eq. (3). Here, we shall consider only the case in which the  $\rho_j$  are commutative.

Since the density operators are commutative, they can be written

$$\rho_j = \sum_i \gamma_{ji} |\gamma_i\rangle \langle \gamma_i|, \quad (4)$$

where  $\{|\gamma_i\rangle\}$  is the set of simultaneous eigenstates of the density operators  $\rho_j$ ;  $j = 1, 2, \dots, M$ .  $\gamma_{ji}$  is the eigenvalue of  $\rho_j$  corresponding to the eigenstate  $|\gamma_i\rangle$ . Substituting this expression for  $\rho_j$  in Eq. (3), we obtain

$$P(\epsilon) = 1 - \frac{1}{M} \sum_{j=1}^M \sum_{\underline{k}} \sum_i p_{j\underline{k}} \gamma_{ji} |\langle x_{\underline{k}} | \gamma_i \rangle|^2. \quad (5)$$

Let  $\left(\max_{1 \leq j \leq M} \gamma_{ji}\right)$  denote the largest of the  $M$  eigenvalues  $\gamma_{1i}, \gamma_{2i}, \dots, \gamma_{Mi}$  for a given  $i$ . It is clear that the right-hand side of Eq. (5) is bounded from below by

$$\begin{aligned} P(\epsilon) &\geq 1 - \frac{1}{M} \sum_i \left(\max_{1 \leq j \leq M} \gamma_{ji}\right) \sum_{\underline{k}} \langle \gamma_i | x_{\underline{k}} \rangle \langle x_{\underline{k}} | \gamma_i \rangle \\ &= 1 - \frac{1}{M} \sum_i \left(\max_{1 \leq j \leq M} \gamma_{ji}\right). \end{aligned} \quad (6)$$

This lower bound of  $P(\epsilon)$  given by Eq. (6) can be achieved by using a receiver that measures dynamical variables of the received field which are represented by the set of operators  $\{X\}$  with simultaneous eigenstates  $|\gamma_i\rangle$ . Since the  $|\gamma_i\rangle$  form an orthogonal complete set, these operators are Hermitian and commutative. Moreover, it follows from Eqs. (2), (4), and (6) that the optimum decision procedure is the maximum-likelihood decision rule. That is,  $p_{j\underline{k}}$  is set to be one if

$$p(\{x_{\underline{k}}\}/m_j) \geq p(\{x_{\underline{k}}\}/m_i); \quad i, j = 1, 2, \dots, M,$$

where  $\{x_{\underline{k}}\}$  denotes the set of eigenvalues of the operators in  $\{X\}$  associated with the state  $|\gamma_k\rangle$ .  $p_{i\underline{k}}$  is equal to zero for all  $i$ ;  $i \neq j$ .

## II. PERFORMANCE OF THE OPTIMUM RECEIVER FOR FREQUENCY ORTHOGONAL SIGNALS WITH KNOWN AMPLITUDE AND RANDOM PHASE

We shall now discuss the optimum performance of a specific system in which the  $M$  density operators  $\rho_i$  are commutative. For the sake of simplicity, we shall assume



that the signal field is linearly polarized and is not spatially modulated. Therefore, in the absence of additive noise, the field over the receiving aperture is a plane wave. Let us assume that the signal field propagates along the z-axis. Hence, the relevant component of the electric field operator can be expanded as

$$E(\underline{r}, t) = i \sum_{\mathbf{k}} \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2}} \left\{ a_{\mathbf{k}} \exp\left[i\omega_{\mathbf{k}}\left(\frac{z}{c} - t\right)\right] - a_{\mathbf{k}}^{\dagger} \exp\left[-i\omega_{\mathbf{k}}\left(\frac{z}{c} - t\right)\right] \right\}.$$

When the transmitted input symbol is  $m_j$ , the field present at the receiver in the absence of the additive thermal noise field is in a state specified by the density operator

$$\rho_j^t = \frac{1}{2\pi} \int_0^{2\pi} d\phi \int_{\mathbf{k}} \delta^{(2)}(a_{\mathbf{k}} - \sigma_{j\mathbf{k}} e^{i\phi}) |a_{\mathbf{k}}\rangle \langle a_{\mathbf{k}}| d^2 a_{\mathbf{k}} \quad (7)$$

in the P representation.<sup>2</sup> In Eq. (7), the  $|a_{\mathbf{k}}\rangle$  are the right eigenstates of the annihilation operators  $a_{\mathbf{k}}$ . The  $\sigma_{j\mathbf{k}}$  are complex quantities known to the receiver, while  $\phi$ , which corresponds to the absolute phase of the received field, is assumed to be a random variable evenly distributed over the interval  $(0, 2\pi)$ . That is, the signal field at the receiver is in one of the coherent states  $\prod_{\mathbf{k}} |\sigma_{j\mathbf{k}} e^{i\phi}\rangle$ .

In the presence of the additive thermal-noise field, the total received field is in a state specified by the density operators

$$\rho_j = \int_0^{2\pi} \frac{d\phi}{2\pi} \int_{\mathbf{k}} \exp\left[-\frac{|a_{\mathbf{k}} - \sigma_{j\mathbf{k}} e^{i\phi}|^2}{\langle n \rangle}\right] |a_{\mathbf{k}}\rangle \langle a_{\mathbf{k}}| \frac{d^2 a_{\mathbf{k}}}{\pi \langle n \rangle} \quad (8)$$

when the message  $m_j$  is transmitted. That is, the additive thermal-noise field alone is in a state specified by the density operator

$$\rho^{(n)} = \int_{\mathbf{k}} \exp\left[-\frac{|a_{\mathbf{k}}|^2}{\langle n \rangle}\right] |a_{\mathbf{k}}\rangle \langle a_{\mathbf{k}}| \frac{d^2 a_{\mathbf{k}}}{\pi \langle n \rangle}, \quad (9)$$

where  $\langle n \rangle$  is the average number of thermal-noise photons in each mode of the noise field. For simplicity, we assume that over the frequency range of interest, it is a constant.

It can be shown that the density operators given by Eq. (8) are commutative when the complex numbers  $\sigma_{j\mathbf{k}}$  satisfy the condition

$$\sum_{\mathbf{k}} \sigma_{j\mathbf{k}} \sigma_{j'\mathbf{k}}^* = \delta_{jj'} |\sigma_{j\mathbf{k}}|^2. \quad (10a)$$

Hereafter, we shall restrict our attention to the set of  $M$  orthogonal signals

corresponding to signals in frequency position modulation systems. In this case, the  $\sigma_{jk}$  also satisfy the condition

$$\sigma_{jk} \sigma_{j'k}^* = |\sigma_{jk}|^2 \delta_{jj'}; \quad \text{for all } k. \quad (10b)$$

(In the classical limit, signals specified by Eqs. (7) and (10b) have known waveforms that constitute a frequency orthogonal set. The absolute phase of the signal field is, however, unknown to the receiver. We assume that it is a random variable evenly distributed over the interval  $(0, 2\pi)$ . The classical electric field amplitude of the additive noise field can be considered as a white Gaussian random process.)

As a simple example of the type of signals described above, let us consider the case in which the signal field is unmodulated. In this case, the density operator in Eq. (8) simplifies to

$$\rho_j = \int_0^{2\pi} \frac{d\phi}{2\pi} \left( \frac{1}{\pi \langle n \rangle} \right)^M \int \left\{ \prod_{\substack{k=1 \\ k \neq j}}^M \exp \left[ -\frac{|a_k|^2}{\langle n \rangle} \right] \right\} \exp \left[ -\frac{|a_j - \sigma_j e^{i\phi}|^2}{\langle n \rangle} \right] \prod_{k=1}^M |a_n\rangle \langle a_k| d^2 a_k. \quad (11)$$

That is, when the transmitted input symbol is  $m_j$ , only the  $j^{\text{th}}$  mode of the signal field is excited while the other modes remain in their vacuum states. Here, we denote the relevant modes of the field as the 1st, 2nd, ... Mth modes. It is clear that these density operators can be written

$$\rho_j = \sum_{n_1=0}^{\infty} \dots \sum_{n_M=0}^{\infty} \left( \frac{1}{1+\langle n \rangle} \right)^M \left( \frac{\langle n \rangle}{1+\langle n \rangle} \right)^{n_1+n_2+\dots+n_M} \left\{ \sum_{\gamma=0}^{n_j} \binom{n_j}{\gamma} \frac{1}{\gamma!} \left[ \frac{|\sigma_j|^2}{(1+\langle n \rangle)\langle n \rangle} \right]^\gamma \right\} \exp \left[ -\frac{|\sigma_j|^2}{1+\langle n \rangle} \right] |n_1, n_2, \dots, n_M\rangle \langle n_1, \dots, n_M|$$

where  $|n_1, n_2, \dots, n_M\rangle$  denotes the simultaneous eigenstate of the number operators  $a_1^\dagger a_1, a_2^\dagger a_2, \dots, a_M^\dagger a_M$  of the individual modes corresponding to eigenvalues  $n_1, n_2, \dots, n_M$ , respectively. Therefore an optimum receiver for this set of M signals measures simultaneously the variables represented by operators  $a_1^\dagger a_1, a_2^\dagger a_2, \dots, a_M^\dagger a_M$ .

Given that  $m_j$  is transmitted, the joint conditional probability distribution of the observed numbers of photons in modes 1, 2, ... M, respectively, being  $n_1, n_2, \dots, n_M$  is

$$P(n_1, n_2, \dots, n_M/m_j) = \left(\frac{1}{1+\langle n \rangle}\right)^M \left(\frac{\langle n \rangle}{1+\langle n \rangle}\right)^{n_1+n_2+\dots+n_M} \left\{ \sum_{\gamma=0}^{n_j} \binom{n_j}{\gamma} \frac{1}{\gamma!} \left[ \frac{|\sigma_j|^2}{\langle n \rangle (1+\langle n \rangle)} \right]^\gamma \right\} \exp\left[-\frac{|\sigma_j|^2}{1+\langle n \rangle}\right]$$

Therefore, the receiver estimates that the transmitted symbol is  $m_j$  if the quantity

$$f_j = \sum_{\gamma=0}^{n_j} \binom{n_j}{\gamma} \frac{1}{\gamma!} \left[ \frac{|\sigma_j|^2}{\langle n \rangle (1+\langle n \rangle)} \right]^\gamma \exp\left[-\frac{|\sigma_j|^2}{1+\langle n \rangle}\right]$$

is maximum over all  $j$ . For simplicity, we shall further assume that the magnitudes of the complex quantities  $\sigma_j$  are such that

$$|\sigma_j|^2 = |\sigma|^2 = p\tau. \quad (12)$$

That is, the average number of photons in the signal field is independent of transmitted input symbols. Then,  $f_j$  is the largest among all  $j$ 's if and only if  $n_j$  is the largest.

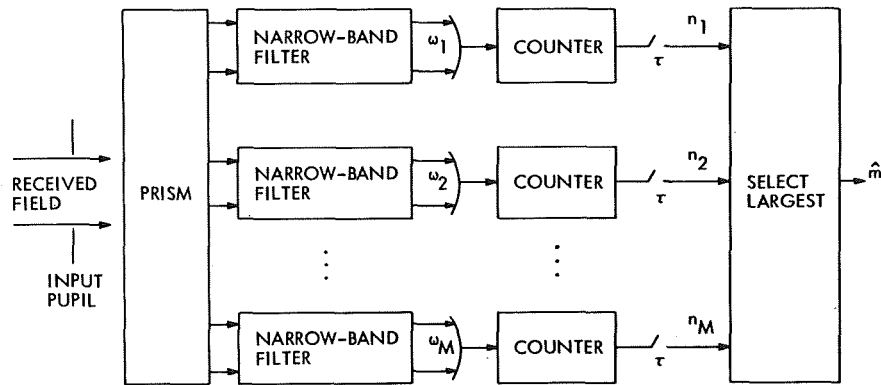


Fig. C-2. Optimum receiver for equal-strength unmodulated signals with random phase.

Therefore, an optimum receiver is the system of photon counters shown in Fig. C-2.

We shall express the bounds to the error probability  $P(\epsilon)$  in the form

$$K_1 \exp[-\tau CE(R)] \leq p(\epsilon) \leq K_2 \exp[-\tau CE(R)],$$

where  $E(R)$  is the system reliability function, and  $C$  is the channel capacity. We shall

be concerned only with the exponential term in the bounds of  $P(\epsilon)$ , since the coefficients of  $K_1$  and  $K_2$  do not depend exponentially on  $\tau$ .

It has been found<sup>3</sup> that the channel capacity  $C$  is given by the expression

$$C = p \ln \frac{1 + \langle n \rangle}{\langle n \rangle}. \quad (14a)$$

For rates  $R$  in the range  $0 \leq R \leq R_C$ , where

$$\frac{R_C}{C} = \frac{\langle n \rangle (1 + \langle n \rangle)}{(1 + 2\langle n \rangle)^2}, \quad (14b)$$

the system reliability function  $E(R)$  is bounded from below by

$$E(R) \geq \frac{1}{(1 + 2\langle n \rangle) \ln \frac{1 + \langle n \rangle}{\langle n \rangle}} - \frac{R}{C}. \quad (14c)$$

For rates  $R$  within the range  $R_C \leq R \leq C$ , the function  $E(R)$  is given by

$$E(R) = \frac{1}{\langle n \rangle \ln \frac{1 + \langle n \rangle}{\langle n \rangle}} \left\{ \frac{1 + 2\langle n \rangle \frac{R}{C} - \sqrt{1 + 4 \frac{R}{C} \langle n \rangle (1 + \langle n \rangle)}}{1 - \sqrt{1 + 4 \frac{R}{C} \langle n \rangle (1 + \langle n \rangle)}} \right\} \\ - \frac{R}{C} \frac{1}{\ln \frac{1 + \langle n \rangle}{\langle n \rangle}} \left\{ \frac{1 + 2\langle n \rangle (1 + \langle n \rangle) \frac{R}{C} + \sqrt{1 + 4 \frac{R}{C} \langle n \rangle (1 + \langle n \rangle)}}{2 \frac{R}{C} (1 + \langle n \rangle)^2} \right\}. \quad (14d)$$

The general behavior of the exponential factor  $CE(R)$  is shown in Fig. C-3 for several values of  $\langle n \rangle$  and  $p/\langle n \rangle = 1.0$ . Also shown in Fig. C-3 is the exponential factor in the expression of  $P(\epsilon)$  when a heterodyne receiver is used. It can be shown that when a heterodyne receiver is used for the reception of the signal specified by Eq. (12), the system capacity is given by

$$C_h = \frac{p}{\langle n \rangle + \frac{1}{2}}. \quad (15a)$$

For rates  $R$  in the range  $0 \leq \frac{R}{C_h} \leq \frac{1}{4}$ , the reliability function is lower-bounded by

$$E_h(R) \geq \frac{1}{2} - \frac{R}{C_h}. \quad (15b)$$

For rates  $R$  in the range  $\frac{1}{4} \leq \frac{R}{C_h} \leq 1$ , the system reliability function is given by

$$E_h(R) = \left(1 - \sqrt{\frac{R}{C_h}}\right)^2. \quad (15c)$$

It has been shown<sup>3</sup> that for arbitrary narrow-band frequency orthogonal signals given by the set of density operators Eq. (8), the optimum system reliability function and channel capacity are also given by Eq. (14) when

$$\sum_k |\sigma_{jk}|^2 = |\sigma_j|^2 = p\tau.$$

The receiver shown in Fig. C-2, however, is no longer an optimum receiver. According to the result in Section I, an optimum receiver should measure dynamical variables

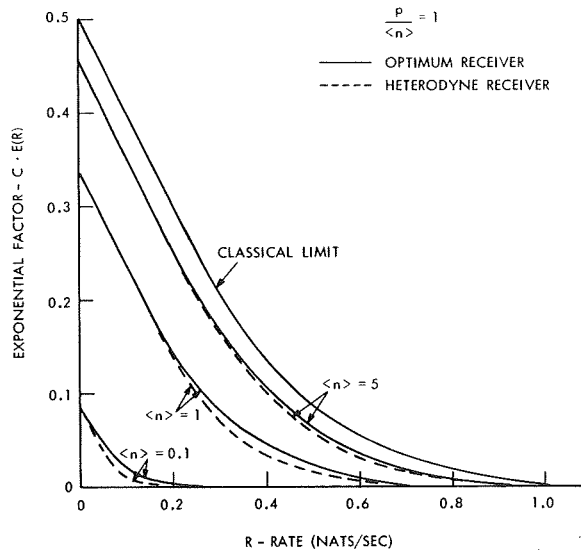


Fig. C-3. System performance of the direct-detection scheme illustrated in Fig. C-2.

represented by operators  $b_1^+ b_1, b_2^+ b_2, \dots, b_M^+ b_M$  simultaneously, where the operators  $b_j$  are given by

$$b_j = \sum_k \sigma_{jk}^* a_k; \quad j = 1, 2, \dots, M.$$

Let us again denote the outcome of the measurement of  $b_1^+ b_1, \dots, b_M^+ b_M$  by  $n_1, n_2, \dots, n_M$ . Then the receiver estimates that  $m_j$  is transmitted if  $n_j$  is the largest among

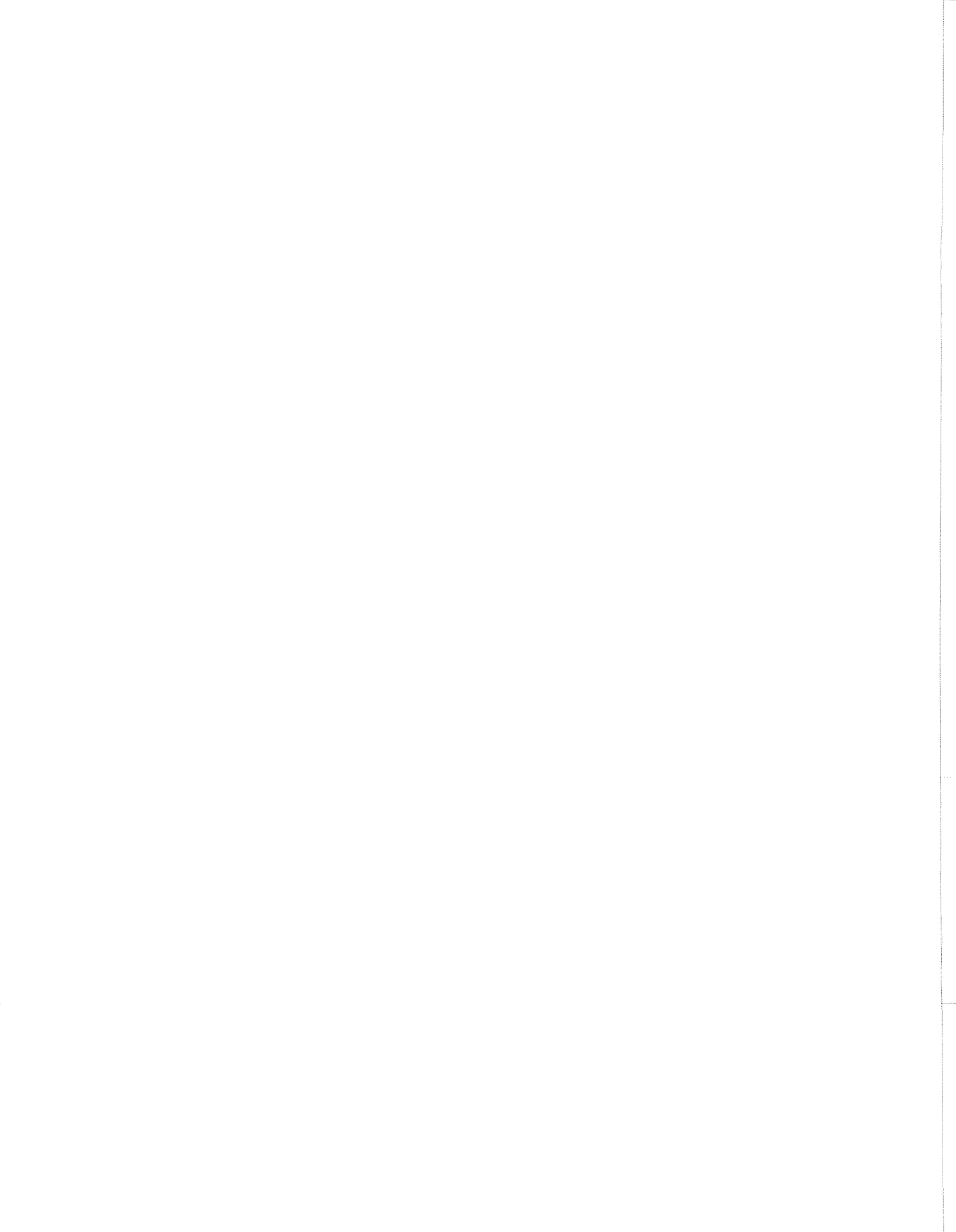
all  $n_j$ . (The  $b_j$  can be considered as annihilation operators of a set of new normal modes with mode functions.

$$V_j(\underline{r}, t) = \frac{1}{|\sigma|} \sum_k \sigma_{jk} \exp\left[-i(\omega_j - \omega_k)\left(\frac{z}{c} - t\right)\right],$$

where the  $\omega_j$  are chosen so that the  $V_j(\underline{r}, t)$  satisfy the orthonormality condition.) For this set of signals, the performance of the heterodyne receiver is again given by Eq. (15) and Fig. C-3. But the system performance of the direct-detection scheme shown in Fig. C-2 is not known.

#### References

1. J. P. Gordon and W. H. Louisell, "Simultaneous Measurement of Non Commuting Observables," in Physics of Quantum Electronics, edited by P. L. Kelley et al. (McGraw-Hill, New York, 1966), pp. 833-840.
2. R. J. Glauber, "Coherent and Incoherent States of the Radiation Field," Phys. Rev. 131, 2766-2788 (1963).
3. Jane W. S. Liu, "Reliability of Quantum-Mechanical Communication Systems," Technical Report 468, Research Laboratory of Electronics, M. I. T., Cambridge, Massachusetts, December 31, 1968.



# Appendix D

## Some Quantum-Mechanical Aspects of Optical Communications

---

Carl W. Helstrom

### I. QUANTUM DETECTION THEORY

#### A. Ideal Receiver

A typical optical communication system will transmit periodically one of  $M$  signals, which, after passing through the intervening medium, appears at the receiver mixed with background radiation. The receiver must decide which of the  $M$  signals is present in its input during each reception interval of duration  $T$ ; that is, it must choose every  $T$  seconds one of  $M$  hypotheses  $H_1, H_2, \dots, H_M$ . Hypothesis  $H_k$  states that the input to the receiver consists of the  $k^{\text{th}}$  signal and background radiation.

In determining the fundamental limitation on the performance of a receiver in such a communication system, it is expedient to imagine the signals as picked up by an ideal receiver consisting of a lossless box or cavity with an aperture facing the transmitter. Initially empty, the cavity is exposed to the input field by opening the aperture at the beginning of the reception interval. After the signal field can be expected to have entirely reached the aperture, the aperture is closed. An observer then proceeds to make the most effective observations of the field within the cavity, and on the basis of their outcomes he decides which of the signals arrived. The same procedure is carried out for each interval during which a signal might appear. By virtue of its independence of particular choices of detector and amplifier structures, this ideal model brings out the limitations on signal detectability imposed by the stochastic and quantum natures of the background radiation and of the signal itself.

The field in the cavity will be described quantum-mechanically by the density operator  $\rho_k$  when hypothesis  $H_k$ , that the  $k^{\text{th}}$  signal was sent, is true. The observer must perform whatever measurements he can on the field in order most efficiently to decide among the  $M$  density operators  $\rho_1, \rho_2, \dots, \rho_M$ . For greatest generality one admits a randomized decision procedure whereby the measurements yield  $M$  numbers  $\pi_1, \pi_2, \dots, \pi_M$ ,

$$0 \leq \pi_k \leq 1, \quad k = 1, 2, \dots, M$$

$$\sum_{k=1}^M \pi_k = 1, \tag{1}$$

of which  $\pi_k$  is the probability with which the observer should choose hypothesis  $H_k$ . He presumably then uses some chance device to make his decision in accordance with the  $M$  probabilities  $\pi_1, \pi_2, \dots, \pi_M$ .



The number  $\pi_k$  can be regarded as an eigenvalue of an operator  $\Pi_k$ . The observer measures the  $M$  operators  $\Pi_1, \Pi_2, \dots, \Pi_M$ , and the outcomes of these measurements are the probabilities  $\pi_1, \pi_2, \dots, \pi_M$ . Because of Eq. (1), the  $M$  operators must sum to the identity operator  $\underline{1}$ ,

$$\Pi_1 + \Pi_2 + \dots + \Pi_M = \underline{1}, \quad (2)$$

and it is actually necessary to measure only  $M - 1$  of them.

From the standpoint of the axioms of quantum mechanics as established, for example, by von Neumann, the  $M$  operators  $\Pi_k$ , necessarily to be measured on one and the same system, must commute. Hence they can be measured simultaneously or one after another without mutual interference.

### B. Broader Formulation

A less restrictive specification of the  $M$  operators is based on the following reasoning. Suppose that an overcomplete set of states  $|\beta\rangle$  characterizes the system,  $\beta$  being a continuous  $n$ -dimensional parameter. If these are to constitute a resolution of the identity, the completeness relation

$$\int |\beta\rangle\langle\beta| d^n\beta = \underline{1} \quad (3)$$

must hold, where  $d^n\beta$  is the volume element of the  $n$ -dimensional parameter space and the integral is taken over all values of  $\beta$ . An example of such a set is furnished by the coherent states of the harmonic oscillator, for which  $n = 2$ , and  $\beta$  is conveniently a complex number.

The decision procedure is then based on determination of the parameter  $\beta$  by some means, followed by the choice of hypothesis  $H_k$  with probability  $\pi_k(\beta)$ ,  $k = 1, 2, \dots, M$ , where the  $\pi_k(\beta)$  are  $M$  functions of  $\beta$  summing to 1,

$$\sum_{k=1}^M \pi_k(\beta) = 1. \quad (4)$$

If hypothesis  $H_k$  is true and the system has the density operator  $\rho_k$ , the probability that the determination\* of  $\beta$  yields an outcome in the element of volume  $d^n\beta$  about the point  $\beta$  is

$$p_k(\beta) d^n\beta = \langle\beta|\rho_k|\beta\rangle d^n\beta. \quad (5)$$

Because of Eq. (3), the  $n$ -dimensional probability density function  $p_k(\beta)$  is properly normalized,

---

\*We speak of "determining"  $\beta$ , using the word "measure" only in the conventional quantum-mechanical sense.

$$\int \rho_k(\beta) d^n\beta = \text{Tr } \rho_k = 1, \quad (6)$$

where 'Tr' stands for the trace of an operator, and the integral is carried out over the entire range of  $\beta$ .

The probability that hypothesis  $H_j$  will be chosen when hypothesis  $H_k$  is really true now is

$$\begin{aligned} \text{Pr } \{H_j | H_k\} &= \int \pi_j(\beta) \rho_k(\beta) d^n\beta \\ &= \int \pi_j(\beta) \langle \beta | \rho_k | \beta \rangle d^n\beta = \text{Tr } (\rho_k \Pi_j) \end{aligned} \quad (7)$$

if the operator  $\Pi_j$  is defined by

$$\Pi_j = \int |\beta\rangle \pi_j(\beta) \langle \beta| d^n\beta. \quad (8)$$

Because of Eqs. (4), and (3), the  $M$  operators  $\Pi_1, \Pi_2, \dots, \Pi_M$  so defined sum to the identity operator as in Eq. (2). If the states  $|\beta\rangle$  are orthonormal,

$$\langle \beta | \beta' \rangle = \delta_n(\beta - \beta'), \quad (9)$$

where  $\delta_n$  is the  $n$ -dimensional delta function, and the operators  $\Pi_k$  commute. The present formulation, therefore, includes the possibility that the operators  $\Pi_k$  commute.

### C. Optimization Problem

The average probability of correctly identifying the transmitted signal is

$$\bar{Q} = \sum_{k=1}^M \zeta_k \text{Pr } \{H_k | H_k\} = \sum_{k=1}^M \zeta_k \text{Tr } (\rho_k \Pi_k) \quad (10)$$

where  $\zeta_k$  is the prior probability of hypothesis  $H_k$ ,

$$\sum_{k=1}^M \zeta_k = 1. \quad (11)$$

The best receiver can be defined as the one for which  $\bar{Q}$  is maximum. If the  $M$  density operators  $\rho_k$  commute, the most effective measurements to make on the system are of a complete set of commuting observables in terms of whose eigenstates the operators  $\rho_k$  are simultaneously diagonal. The receiver can then process the outcomes of its measurements as prescribed by conventional detection theory.

When the  $M$  density operators  $\rho_k$  do not commute, the determination of the  $M$  operators  $\Pi_k$  that maximize  $\bar{Q}$  is, for  $M > 2$ , an unsolved problem. For  $M = 2$ , measurement

of only a single operator  $\Pi$  is necessary, and it has been shown<sup>1</sup> that  $\Pi$  is a projection operator on the eigenstates of the operator  $\xi_2\rho_2 - \xi_1\rho_1$  having positive eigenvalues. Even for  $M = 2$ , however, there are few pairs of noncommuting density operators  $\rho_1, \rho_2$  for which the eigenvalue problem can be solved and the optimum detection operator  $\Pi$  determined.

One problem that can be solved for  $M = 2$  has to do with the optimum choice between two pure quantum-mechanical states  $|1\rangle$  and  $|2\rangle$ , which might, for example, be two coherent states of the harmonic oscillator.<sup>2</sup> The density operators are

$$\rho_1 = |1\rangle\langle 1|, \quad \rho_2 = |2\rangle\langle 2|,$$

and the optimum operator  $\Pi$  is

$$\Pi = |\eta_1\rangle\langle \eta_1|,$$

where  $|\eta_1\rangle$  is the eigenstate of  $\xi_2\rho_2 - \xi_1\rho_1$  having a positive eigenvalue. The two eigenstates of this operator have the form  $|\eta\rangle = x_1|1\rangle + x_2|2\rangle$ , with  $x_1$  and  $x_2$  coefficients that can be determined algebraically.

In the choice between a coherent state  $|a\rangle$  and the ground state  $|0\rangle$  of the harmonic oscillator, the probability of error attained by measuring  $\Pi$  is smaller than that incurred by measuring the number of photons in the system or by measuring the oscillator coordinate parallel to the complex number  $a$ , these last two methods corresponding to incoherent and coherent detection, respectively. The analysis and the resulting error probabilities have been presented elsewhere.<sup>2</sup>

#### D. Choice among $M$ Pure States

If the receiver can be in one of  $M$  pure states, represented by the vectors  $|1\rangle, |2\rangle, \dots, |M\rangle$ , and the observer is to decide which one, it is necessary to consider only operators that can be expressed in terms of eigenvectors in the  $M$ -dimensional space spanned by the  $M$  state vectors  $|k\rangle, k = 1, 2, \dots, M$ . The states  $|\beta\rangle$  in Eqs. (3)-(8) can now be limited to  $M$  discrete vectors  $|\beta_1\rangle, |\beta_2\rangle, \dots, |\beta_M\rangle$  spanning the same  $M$ -dimensional space. The integrals in Eqs. (3)-(8) are replaced by sums over this discrete set.

The "resolution of the identity"

$$\sum_{k=1}^M |\beta_k\rangle\langle \beta_k| = \mathbb{1} \quad (12)$$

now requires the vectors  $|\beta_k\rangle$  to be orthonormal  $k = 1, 2, \dots, M$ . Indeed, Eq. (12) implies that the scalar product of any two of these vectors is

$$\langle \beta_m | \beta_n \rangle = \sum_{k=1}^M \langle \beta_m | \beta_k \rangle \langle \beta_k | \beta_n \rangle. \quad (13)$$

Let  $\underline{B}$  be the  $M \times M$  matrix of the scalar products  $\langle \beta_m | \beta_n \rangle$ ; Eq. (13) is then

$$\underline{B} = \underline{E}^2.$$

Hence  $\underline{B}$  must be the identity matrix, the solution  $\underline{B} = 0$  being untenable, and the vectors  $|\beta_k\rangle$  must be orthonormal.

As the  $|\beta_k\rangle$  are now orthonormal, the operators

$$\Pi_j = \sum_{k=1}^M |\beta_k\rangle \pi_j(\beta_k) \langle \beta_k|, \quad (14)$$

defined as in Eq. (8), commute, for they are simultaneously diagonal in the  $|\beta_k\rangle$ -representation. The average probability of correct decision is

$$\bar{Q} = \sum_{j=1}^M \zeta_j \sum_{k=1}^M \pi_j(\beta_k) |\langle \beta_k | j \rangle|^2, \quad (15)$$

where  $\zeta_j$  is the prior probability of hypothesis  $H_j$ ,  $j = 1, 2, \dots, M$ . It is apparent that the orthonormal set  $|\beta_k\rangle$  can be so oriented in the  $M$ -dimensional space that one vector, say  $|\beta_j\rangle$ , is closest to each state vector  $|j\rangle$ , and

$$|\langle \beta_j | j \rangle|^2 > |\langle \beta_k | j \rangle|^2, \quad k \neq j.$$

The average probability  $\bar{Q}$  is then maximum when  $\pi_j(\beta_j) = 1$ ,  $\pi_k(\beta_j) = 0$ ,  $k \neq j$ . Randomization is unnecessary, and

$$\bar{Q} = \sum_{j=1}^M \zeta_j |\langle \beta_j | j \rangle|^2. \quad (16)$$

How to orient the orthonormal set  $|\beta_j\rangle$  to maximize the average probability  $\bar{Q}$  of correct decision is an unsolved problem for  $M > 2$ . An analysis showing the difficulties involved is presented in Section II.

#### E. Afterthoughts

The procedure outlined in part (B) involved determining—somehow—the  $n$ -dimensional parameter  $\beta$  and then choosing hypothesis  $H_k$  with probability  $\pi_k(\beta)$ ,  $k = 1, 2, \dots, M$ , where the  $\pi_k(\beta)$  are such functions that the average probability  $\bar{Q}$  of correct decision is maximum. This probability was expressed in Eq. (10) in terms of  $M$  operators

$$\Pi_k = \int |\beta\rangle \pi_k(\beta) \langle \beta| d^n \beta. \quad (17)$$

The integral is taken over the entire range of  $\beta$ . The decision procedure is not necessarily equivalent to measuring these  $M$  operators  $\Pi_1, \Pi_2, \dots, \Pi_M$ .

To see this, consider an operator of similar structure, namely the density operator  $\rho$  of a harmonic oscillator in thermal equilibrium at temperature  $\tau$ . This operator can be expressed in terms of the coherent states  $|\beta\rangle$  of the harmonic oscillator by the expression

$$\rho = \int |\beta\rangle P(\beta) \langle\beta| d^2\beta,$$

$$P(\beta) = (N)^{-1} \exp(-|\beta|^2/N), \quad (18)$$

where

$$N = 1 - e^{-w}, \quad w = h\nu/K\tau,$$

$h\nu$  being the energy per quantum and  $K$  Boltzmann's constant.<sup>3</sup> This way of writing  $\rho$  resembles (17), but another way of writing it is in terms of the energy eigenstates  $|n\rangle$  of the oscillator,

$$\rho = \sum_{n=0}^{\infty} |n\rangle q_n \langle n|$$

$$q_n = (N+1)^{-1} \left( \frac{N}{1+N} \right)^n. \quad (19)$$

A measurement of the operator  $\rho$  can yield only one of the eigenvalues  $q_n$ . What corresponds to the procedure in part (B), however, is determination first – somehow – of the complex number  $\beta$ , and then evaluation of the function  $P(\beta)$  in Eq. (18). This can yield any of a continuum of positive values between 0 and  $(N)^{-1}$ . It is possible, therefore, that once the function  $\pi_k(\beta)$  has been specified, the operator  $\Pi_k$  may acquire a spectrum of eigenvalues that does not coincide with the domain of the function  $\pi_k(\beta)$ . Measurement of  $\Pi_k$  may not be equivalent to first determining  $\beta$  and then calculating  $\pi_k(\beta)$ . Furthermore, if the  $\Pi_k$  do not commute and one tries to measure them simultaneously, one cannot be sure that the average probability of correct decision is still given by Eq. (10).

Two possibilities then arise: either the set of states  $|\beta\rangle$  to be used in the procedure is given in advance, or it is not. If it is given, the probability densities  $p_k(\beta)$  are also known, and the average probability  $\bar{Q}$  of correct decision can be written in the usual way as

$$\bar{Q} = \sum_{k=1}^M \zeta_k \int \pi_k(\beta) p_k(\beta) d^n\beta. \quad (20)$$

This is maximized in the same manner as in conventional detection theory, and the best receiver determines the  $M$  likelihoods  $\zeta_{k^*p_k}(\beta)$ ,  $k = 1, 2, \dots, M$ , and picks up the hypothesis corresponding to the largest.

If the set of states  $|\beta\rangle$  is not given in advance, there must be some way of determining it from the operators  $\Pi_k$  that maximize  $\bar{Q}$  of Eq.(10), provided the  $\Pi_k$  can be found, for we have seen that it may not be possible to attain the same maximum  $\bar{Q}$  by measuring the  $\Pi_k$  themselves if these do not commute. How one can specify a set  $|\beta\rangle$  permitting the decision scheme to be realized in terms of the probabilities  $\pi_k(\beta)$  is unclear, and the formulation of part (B) may be incomplete.

## II. OPTIMUM CHOICE AMONG $M$ PURE STATES

Continuing with the problem of choosing with maximum probability of correctness which of  $M$  state vectors  $|1\rangle, |2\rangle, \dots, |M\rangle$  describes a quantum-mechanical system, we write these states as linear combinations of the  $M$  orthonormal vectors  $|\beta_j\rangle$ , using (12),

$$|k\rangle = \sum_{j=1}^M |\beta_j\rangle \langle \beta_j | k \rangle = \sum_{j=1}^M \zeta_{kj} |\beta_j\rangle,$$

$$\zeta_{kj} = \langle \beta_j | k \rangle. \quad (21)$$

We wish to find how the orthonormal set  $|\beta_j\rangle$  should be oriented to maximize the probability  $\bar{Q}$  of Eq. (16). For simplicity, we take the states as equally likely,  $\zeta_j \equiv 1/M$ .

The probability of choosing hypothesis  $H_j$ , that the system is in state  $|j\rangle$ , when it really is in state  $|j\rangle$ , is

$$\Pr \{ \Pi_j \rightarrow 1 | H_j \} = |\langle j | \beta_j \rangle|^2 = |\zeta_{jj}|^2. \quad (22)$$

Since the vector  $|k\rangle$  can be multiplied by an arbitrary phase factor without changing any probabilities, we are free to assume that  $\zeta_{jj} = \langle \beta_j | j \rangle$  is real, and we do so henceforth.

The average probability of correct decision when the states  $|k\rangle$  are equally probable is

$$\bar{Q} = \frac{1}{M} \sum_{j=1}^M \zeta_{jj}^2. \quad (23)$$

We are to maximize this by varying the  $\zeta_{kj}$  under the constraint that the cosines  $\langle j | k \rangle$  of the angles between the state vectors  $|j\rangle$  and  $|k\rangle$  be fixed for all  $j$  and  $k$ :

$$\begin{aligned}
\gamma_{jk} &= \langle j|k \rangle = \sum_{i,m} \zeta_{ji}^* \zeta_{km} \langle \beta_i | \beta_m \rangle \\
&= \sum_i \zeta_{ji}^* \zeta_{ki}
\end{aligned} \tag{24}$$

since  $\langle \beta_i | \beta_m \rangle = \delta_{im}$ .

Denote by  $\underline{Y}_k^+$  the row vector  $(\zeta_{1k}^* \zeta_{2k}^* \dots \zeta_{Mk}^*)$ , and by  $\underline{Y}_k$  the complex conjugate column vector of the  $\zeta_{ik}$ ,  $j = 1, 2, \dots, M$ . Let  $\underline{J}_k$  be a matrix that has a 1 in the  $k^{\text{th}}$  diagonal element and zeros elsewhere,

$$(\underline{J}_k)_{ij} = \delta_{ij} \delta_{jk} \tag{25}$$

Then  $\underline{J}_k \underline{Y}_k \underline{Y}_k^+ \underline{J}_k$  is an  $M \times M$  matrix with  $\zeta_{kk}^2$  as its  $k^{\text{th}}$  diagonal element and zeros elsewhere, and

$$\underline{D} = \sum_{k=1}^M \underline{J}_k \underline{Y}_k \underline{Y}_k^+ \underline{J}_k \tag{26}$$

is a diagonal matrix with diagonal elements  $\zeta_{kk}^2$ . Furthermore,

$$\begin{aligned}
M\bar{Q} &= \text{Tr } \underline{D} = \text{Tr} \sum_{k=1}^M \underline{J}_k \underline{Y}_k \underline{Y}_k^+ \underline{J}_k \\
&= \text{Tr} \sum_{k=1}^M \underline{J}_k \underline{Y}_k \underline{Y}_k^+
\end{aligned} \tag{27}$$

The matrix  $\underline{\Gamma}$  whose elements are  $\Gamma_{kj} = \gamma_{jk}$  is given by

$$\underline{\Gamma} = \sum_{k=1}^M \underline{Y}_k \underline{Y}_k^+ \tag{28}$$

by virtue of (24). We want to maximize  $\text{Tr } \underline{D}$  for fixed  $\underline{\Gamma}$ , and we can do so by introducing the matrix  $\underline{\Lambda}$  of Lagrange multipliers  $\lambda_{kj}$  and maximizing the function

$$\begin{aligned}
F(\{\underline{Y}_j\}) &= \text{Tr } \underline{D} - \sum_{kj} \lambda_{kj} \Gamma_{jk} \\
&= \text{Tr} (\underline{D} - \underline{\Lambda} \underline{\Gamma}) \\
&= \sum_k \text{Tr} (\underline{J}_k - \underline{\Lambda}) \underline{Y}_k \underline{Y}_k^+
\end{aligned} \tag{29}$$

of the elements  $\zeta_{kj}$  of the vectors  $\underline{Y}_j$ .

If we make arbitrary increments  $\xi_k$  in the vectors  $\underline{Y}_k$ , we find, to first order in the  $\epsilon_{jk}$ , where  $\xi_k^+ = (\epsilon_{1k}^+ \ \epsilon_{2k}^* \ \dots \ \epsilon_{Mk}^*)$ ,

$$F(\{\underline{Y}_j + \xi_j\}) - F(\{\underline{Y}_j\}) = \sum_k \text{Tr} (\underline{J}_k - \underline{\Lambda}) (\underline{Y}_k \xi_k^+ + \xi_k \underline{Y}_k^+).$$

At the maximum this must be zero for arbitrary  $\xi_k$  and  $\xi_k^+$ , since the components  $\epsilon_{jk}$  may be complex. Hence we obtain

$$(\underline{J}_k - \underline{\Lambda}) \underline{Y}_k = 0,$$

$$\underline{Y}_k^+ (\underline{J}_k - \underline{\Lambda}) = 0,$$

which leads to the equations

$$\begin{aligned} \underline{\Lambda} \underline{Y}_k &= \underline{J}_k \underline{Y}_k = \zeta_{kk} \underline{G}_k \\ \underline{Y}_k^+ \underline{\Lambda} &= \underline{Y}_k^+ \underline{J}_k = \zeta_{kk} \underline{G}_k^+, \end{aligned} \quad (30)$$

where  $\underline{G}_k$  is a column vector with 1 in the  $k^{\text{th}}$  place, and zeros elsewhere. Taking the conjugate transpose of the second of these, we see that  $\underline{\Lambda}$  must be Hermitian,  $\underline{\Lambda} = \underline{\Lambda}^+$ .

Writing out the first equation in (30), we find

$$\sum_{j=1}^M \Lambda_{ij} (\zeta_{jk} / \zeta_{kk}) = \delta_{ik}.$$

If we define the matrix  $\underline{U} = \|\underline{U}_{ij}\|$  as the  $M \times M$  array of elements

$$\underline{U}_{ij} = \zeta_{ij} / \zeta_{jj}, \quad (31)$$

this equation can be written  $\underline{\Lambda} \underline{U} = \underline{I}$ , where  $\underline{I}$  is the  $M \times M$  identity matrix. Hence  $\underline{U} = \underline{\Lambda}^{-1}$  is Hermitian, and its diagonal elements are 1.

The "Hermiticity" of  $\underline{U}$  imposes a condition on the coefficients  $\zeta_{ij}$ , namely

$$\frac{\zeta_{ji}^*}{\zeta_{ii}} = \frac{\zeta_{ij}}{\zeta_{jj}}. \quad (32)$$

The matrix



$$\underline{Y} = \|\zeta_{ij}\| = (\underline{Y}_1 \underline{Y}_2 \dots \underline{Y}_M) \quad (33)$$

is related to the known matrix  $\underline{\Gamma}$  by (28), or by

$$\underline{\Gamma} = \underline{Y} \underline{Y}^+ \quad (34)$$

Because of Eq. (32),  $\underline{Y}$  has only  $\frac{1}{2} M(M+1)$  independent elements, and by writing out Eq. (34) we obtain  $\frac{1}{2} M(M+1)$  distinct equations for them. Unfortunately, the equations are nonlinear in a most complicated way.

From Eqs. (26), (28), and (30) we find

$$\begin{aligned} \underline{\Lambda} \underline{\Gamma} \underline{\Lambda} &= \sum_{k=1}^M \underline{\Lambda} \underline{Y}_k \underline{Y}_k^+ \underline{\Lambda} \\ &= \sum_{k=1}^M \underline{J}_k \underline{Y}_k \underline{Y}_k^+ \underline{J}_k = \underline{D}. \end{aligned} \quad (35)$$

Hence the unknown matrix  $\underline{\Lambda}$  is a Hermitian matrix that diagonalizes  $\underline{\Gamma}$ , if such can be found. Equivalently,

$$\underline{\Gamma} = \underline{U} \underline{D} \underline{U} \quad (36)$$

which also yields a set of  $\frac{1}{2} M(M+1)$  equations for the  $M$  diagonal elements  $\zeta_{jj}^2$  of  $\underline{D}$  and the  $\frac{1}{2} M(M-1)$  elements above the diagonal of the Hermitian matrix  $\underline{U}$ . ( $U_{kk} = 1$ ,  $k = 1, \dots, M$ .) These equations also are nonlinear.

#### References

1. C. W. Helstrom, "Detection Theory and Quantum Mechanics," Inform. Contr. **10**, 254-291 (March 1967).
2. C. W. Helstrom, "Detection Theory and Quantum Mechanics II," Inform. Contr. **13**, 156-171 (August 1968).
3. R. J. Glauber, "Coherent and Incoherent States of the Radiation Field," Phys. Rev. **131**, 2766-2788 (1963).

## Appendix E

### Estimation of Modulation Frequency of a Light Beam

Carl W. Helstrom

In a system to measure the velocity of a target, the intensity of a beam of light is modulated sinusoidally and directed toward it. The reflected light has all its frequencies shifted by the Doppler effect, and the frequency of the intensity modulation is also changed, by an amount proportional to the modulation frequency and to the radial velocity of the target. The reflected light falls on a photoelectric surface, and the emitted photoelectrons excite a tuned circuit. The average rate of emission is sinusoidally modulated, and by finding the resonant frequency at which the response of the tuned circuit is maximum, the modulation frequency of the light and hence the radial component of the target velocity can be determined. A device for accomplishing this was proposed by Nelson McEvoy.

The question arises how accurate an estimate of the modulation frequency can be attained by such a method, and it can be answered at least roughly by an analysis involving the Cramér-Rao inequality of statistical estimation theory.

The photoelectrons are observed over an interval of duration  $T$ , which can be taken as the integration time of the tuned circuit. Divide this interval into subintervals of duration  $\Delta t$  that are much shorter than the period of the modulation frequency. Let  $t_i$  be a time within the  $i^{\text{th}}$  of these subintervals, and let  $n_i$  be the number of photoelectrons emitted in that subinterval. The  $n_i$  can be taken as independent random variables with a Poisson distribution; the mean value of  $n_i$  is

$$\mathbb{E}(n_i) = \lambda(t_i) \Delta t, \quad (1)$$

where  $\lambda(t)$  is the rate of photoelectron emission and is sinusoidally modulated,

$$\lambda(t) = \lambda_0(1 + \cos \Omega t). \quad (2)$$

It is the angular frequency  $\Omega$  that is to be estimated.

The joint probability of observing  $n_i$  photoelectrons in the  $i^{\text{th}}$  interval,  $i = 1, 2, \dots$ , is

$$p(\{n_i\}) = \prod_i (n_i!)^{-1} [\lambda(t_i)\Delta t]^{n_i} \exp[-\lambda(t_i)\Delta t]. \quad (3)$$

The Cramér-Rao inequality tells us that the variance of an unbiased estimate  $\hat{\Omega}$  of the parameter  $\Omega$  of this distribution is bounded below by

$$\text{Var } \hat{\Omega} \geq \left\{ \mathbb{E} \left[ \frac{\partial}{\partial \Omega} \ln p(\{n_i\}) \right]^2 \right\}^{-1}. \quad (4)$$

We proceed to evaluate the right-hand side of this expression.

The logarithm of the probability in (3) is

$$\ln p(\{n_i\}) = \sum_i \{n_i \ln[\lambda(t_i)\Delta t] - \lambda(t_i)\Delta t - \ln(n_i!)\}$$

and its derivative is

$$\frac{\partial}{\partial \Omega} \ln p(\{n_i\}) = \sum_i \left[ n_i \lambda^{-1} \frac{\partial \lambda}{\partial \Omega} - \frac{\partial \lambda}{\partial \Omega} \Delta t \right]. \quad (5)$$

By virtue of (1), the expected value of this expression is zero, and its mean-square value is, therefore,

$$\mathbb{E} \left[ \frac{\partial}{\partial \Omega} \ln p(\{n_i\}) \right]^2 = \sum_i (\text{Var } n_i) \lambda^{-2} \left( \frac{\partial \lambda}{\partial \Omega} \right)^2. \quad (6)$$

Since

$$\text{Var } n_i = \mathbb{E}(n_i) = \lambda(t_i) \Delta t,$$

we find

$$\begin{aligned} \mathbb{E} \left[ \frac{\partial}{\partial \Omega} \ln p(\{n_i\}) \right]^2 &= \sum_i [\lambda(t_i)]^{-1} \left( \frac{\partial \lambda}{\partial \Omega} \right)^2 \Delta t \\ &\doteq \int_0^T \frac{dt}{\lambda(t)} \left( \frac{\partial \lambda}{\partial \Omega} \right)^2 \end{aligned} \quad (7)$$

when we pass to the limit  $\Delta t \ll 1/\Omega$ .

From (2),

$$\partial \lambda / \partial \Omega = -\lambda_0 t \sin \Omega t,$$

and

$$\begin{aligned} \int_0^T \frac{dt}{\lambda(t)} \left( \frac{\partial \lambda}{\partial \Omega} \right)^2 &= \lambda_0 \int_0^T \frac{t^2 \sin^2 \Omega t}{1 + \cos \Omega t} dt \\ &= \lambda_0 \int_0^T t^2 (1 - \cos \Omega t) dt \\ &= \frac{1}{3} \lambda_0 T^3 [1 + O(1/\Omega T)]. \end{aligned} \quad (8)$$

Under the assumption that  $\Omega T \gg 1$  (many cycles of the modulation are encompassed), we obtain from (4), (7), and (8) the lower bound

$$\text{Var } \hat{\Omega} \gtrsim 3/\lambda_0 T^3 \quad (9)$$

The Cramér-Rao inequality provides only a lower bound, and there is no assurance that it can be approached. At high signal-to-noise ratio a maximum-likelihood estimate of  $\Omega$  will have an error variance given approximately by the right-hand side of Eq. (9). Hence for such an estimate

$$\text{Var } \hat{\Omega} \sim 3/\lambda_0 T^3, \quad \lambda_0 T \gg 1, \quad (10)$$

since with "shot noise" such as we are measuring here the mean number  $\lambda_0 T$  of photoelectrons serves as a signal-to-noise ratio. Thus

$$\text{Var } \hat{\Omega} \sim \left(\frac{3}{T^2}\right) \frac{1}{\lambda_0 T}, \quad \lambda_0 T \gg 1 \quad (11)$$

has the usual form of the error variance of a frequency estimate, being inversely proportional to the square of the observation time  $T$  and to the signal-to-noise ratio.

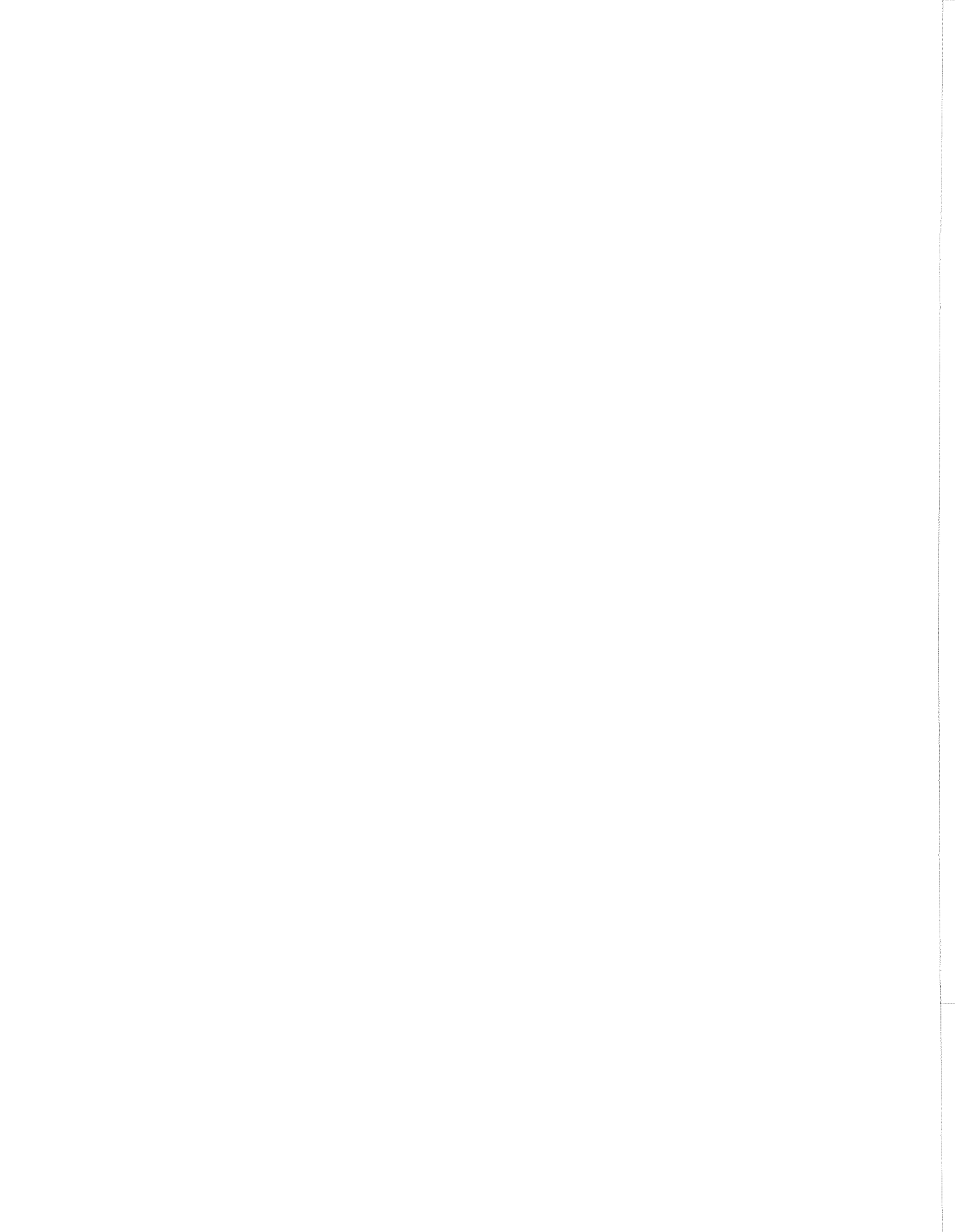
If the frequency to be measured is changing at a rate  $\dot{\Omega}$ , the observation time  $T$  should be short enough that

$$\dot{\Omega} T \ll \sqrt{\text{Var } T} = \sqrt{3/\lambda_0 T^3}$$

or

$$T \ll \left(3/\lambda_0 \dot{\Omega}^2\right)^{1/5}.$$

Otherwise an analysis taking both  $\Omega$  and  $\dot{\Omega}$  as unknown parameters to be estimated must be carried out. For this the two-dimensional form of the Cramér-Rao inequality can be employed.



# Appendix F

## Comments on the Quantal Theory of Communication

James P. Gordon

### I. INTRODUCTION

Quantum-mechanical effects, or simply quantal effects, are important to communications using optical-frequency radiation. In communication theory, one must recognize that information is carried by fields or particles that obey the quantal laws. The rewriting of information theory to take this fact into account poses a number of interesting problems, only a few of which have been solved thus far. On the other hand, most systems envisaged now involve some linear processing of an incident field (amplification, addition of locally generated fields, etc.), followed by a detector characterized by photoproduction of electrons at a mean rate proportional to the short-time average of the optical power incident on it. By short time is meant a time short compared with the reciprocal of the signal bandwidth or intermediate frequency, if applicable, and yet long compared with the reciprocal of the optical carrier frequency. In these cases it appears that the quantal nature of the field can be rigorously ignored if the basic photoproduction process is assumed to have Poisson statistics.

In Section II we shall briefly discuss the application of quantal laws to communication theory, and in Section III we shall give a simple example of how classical fields and Poisson photoproduction statistics give correct quantal results.

### II. QUANTAL DESCRIPTION OF COMMUNICATION

The communication system that we envisage comprises a modulator-transmitter that implants information onto a field (or other quantal objects), a transmission path through which the field is sent, and finally a receiver that measures some property or properties of the field in order to extract the implanted information. We consider that the field strength is much attenuated by the transmission path, and perhaps some noise (such as thermal background) is added to it. In order to avoid undue complications, we shall consider that only one transverse field mode and polarization is important, so the path is analogous to a single-mode waveguide. In some reasonably long time interval  $T$ , each of the independent Fourier components satisfying periodic boundary conditions over time  $T$ , and hence separated by  $\delta\nu = 1/T$  from its neighbors, behaves like a harmonic oscillator. The number of such independent harmonic oscillators is equal to the time-bandwidth product, and we may liken the model system to an equivalent one in which the information is implanted by suitable preparation of harmonic oscillators, which are then sent to the receiver at a rate equal to the optical-frequency bandwidth. Now consider an imaginary plane cutting the system at the point where the transmitted power

reaches its lowest ebb, namely just at the receiver aperture. Everything to the left (transmitter side) of this plane may be associated with the preparation of each harmonic oscillator as it arrives at the plane, while everything to the right may be associated with its measurement.

A harmonic oscillator is described classically by a coordinate  $q$  and a canonically conjugate momentum  $p$ . The proportionality between  $q, p$  and the electromagnetic fields  $E, H$  is well documented. The energy  $W$  of the oscillator is

$$W = \frac{1}{2} (p^2 + \omega^2 q^2). \quad (1)$$

In quantal theory  $q$  and  $p$  become noncommuting operators satisfying Heisenberg's uncertainty relation, and  $W$  becomes the Hamiltonian operator.

A pure state of preparation of an oscillator is described no longer by assigning values of  $q$  and  $p$ , but rather by assigning a complex wave function  $\psi_t(q)$  after Schrödinger, or a ket vector  $|t\rangle$  in Hilbert space after Dirac. The index  $t$  is a (possibly multidimensional) index ranging over all possible states that the transmitter may prepare. The index may be thought of as identifying the various possible knob settings of the modulator. As a pertinent example we have the coherent states championed by Glauber, for which the normalized wave functions are given by

$$\psi_{\bar{q}, \bar{p}}(q) = \left[ \frac{\omega}{\pi \hbar} \right]^{1/4} \exp \left\{ -\frac{\omega(q-\bar{q})^2}{2\hbar} + \frac{i\bar{p}q}{\hbar} \right\}. \quad (2a)$$

Following Klauder, we can represent these states by kets labeled

$$|\bar{q}, \bar{p}\rangle. \quad (2b)$$

These are the states which would be prepared by a highly coherent laser followed by a modulator that could adjust amplitude and phase, followed by a noise-free ( $0^\circ\text{K}$ ) path. Another example might be the energy states, whose wave functions are the Hermite-Gaussian functions, and which we can represent by the kets

$$|n\rangle. \quad (3)$$

Here the index  $n$  represents the number of quanta of energy  $h\nu$  in the state. Because of the attenuating path, these pure states cannot be prepared in our system; and this is unfortunate because they would give us a communication channel with the maximum possible information rate. An exception is the ground state  $|0\rangle$ , which is identical with the coherent state  $\bar{q} = \bar{p} = 0$ .

A rather general description of the measurement process involves two stages: first, a reversible coupling of the original quantal system to an "auxiliary" quantal system whose initial state is set by the receiver, followed after a suitable interval by the irreversible exact "observation" of a complete set of commuting observables for the expanded

(original plus auxiliary) system. This description allows for minimum uncertainty combined measurements of noncommuting observables of the original system. For example, if we wish to measure both position and momentum of a harmonic oscillator, we might couple half its energy out into a second identical harmonic oscillator which was initially in its ground state, whereupon exact "observation" of the momenta of the two separate oscillators at suitable later times would give us measures of both the position and momentum of the original oscillator. It turns out that the results of all such measurements if done as well as quantal theory allows can be expressed in the form

$$P(r/t) = \left| \int \psi_r^*(q) \psi_t(q) dq \right|^2 = |\langle r | t \rangle|^2. \quad (4)$$

Here  $r$  is an index that covers all possible measurement results (meter readings), and  $P(r/t)$  is the conditional probability that the measurement will give a reading  $r$  (again possibly multidimensional), given that the transmitter prepared the systems in the state  $|t\rangle$ . The state  $|r\rangle$ , or  $\psi_r(q)$ , represents a state of the original system (harmonic oscillator) which is associated through the measurement process with the result  $r$ . If the set of prepared states  $|t\rangle$  are unit-normalized so that

$$\langle t | t \rangle = \int \psi_t^*(q) \psi_t(q) dq = 1$$

and we write

$$P(r/t) = \langle t | r \rangle \langle r | t \rangle,$$

then we see that  $p(r/t)$  is a probability function over  $r$  if for any  $t$  we have

$$\sum_r P(r/t) = \langle t | \left[ \sum_r |r\rangle \langle r| \right] |t\rangle = 1.$$

Thus we require only that the set of measurement states satisfy a completeness relation

$$\sum_r |r\rangle \langle r| = I, \quad (5)$$

where  $I$  is the unit operator. The notation  $|r\rangle \langle r|$  represents the outer (tensor) product of the ket  $|r\rangle$  with its corresponding bra  $\langle r|$  in the conjugate space. An important point here is that the set of measurement states  $|r\rangle$  need not be orthogonal. For pertinent examples, the measurement states corresponding to energy measurement are the energy states  $|n\rangle$ , which satisfy the completeness relation

$$\sum_n |n\rangle \langle n| = I.$$

Correspondingly, for energy measurements, we find

$$P(n/t) = |\langle n | t \rangle|^2.$$



Likewise the position eigenkets  $|q\rangle$  satisfy the completeness relation

$$\int |q\rangle \langle q| dq = I,$$

and correspondingly

$$P(q/t) = |\langle q|t\rangle|^2 = |\psi_t(q)|^2.$$

Here, of course, the probability function  $p(q/t)$  is continuous rather than discrete as was the case for the energy measurement, and we achieve the standard result that the absolute square of the wave function is the probability function for position measurement. But now our formulation covers minimum uncertainty measurements of both position and momentum. The coherent states also satisfy a completeness relation

$$\frac{1}{2\pi\hbar} \int |q_m, p_m\rangle \langle q_m, p_m| dq_m dp_m = I, \quad (6)$$

where the individual vectors are unit-normalized, and  $q_m, p_m$  replace  $\bar{q}, \bar{p}$  in (2) and represent the measurement results. Correspondingly, a minimum uncertainty measurement of both  $p$  and  $q$  can result in a conditional probability function

$$\begin{aligned} P(q_m, p_m/\bar{p}, \bar{q}) &= \frac{1}{2\pi\hbar} |\langle p_m, q_m | \bar{p}, \bar{q} \rangle|^2 \\ &= \frac{1}{2\pi\hbar} \exp \left\{ -\frac{\omega(q_m - \bar{q})^2}{2\hbar} - \frac{(p_m - \bar{p})^2}{2\hbar\omega} \right\}, \end{aligned} \quad (7)$$

where in the last line we have worked out the answer, using the wave functions of (2a) and the indicated integration of (4). Note that the variance of the  $q$  part of the measurement, namely  $\hbar/\omega$ , is twice as large as we could have achieved by an exact measurement of  $q$ , for which the result could be

$$P(q_m/\bar{q}, \bar{p}) = |\psi_{\bar{q}, \bar{p}}(q_m)|^2 = \left[ \frac{\omega}{\pi\hbar} \right]^{1/2} \exp \left\{ -\frac{\omega(q_m - \bar{q})^2}{\hbar} \right\}, \quad (8)$$

where we have used (2a). Thus we cannot obtain information about  $\bar{p}$  without giving up some ground in the measurement of  $\bar{q}$ , which is also in accord with the uncertainty principle. We may note that the uncertainty principle has two faces, one saying that we cannot prepare a system with exactly defined  $p$  and  $q$ , the other saying that in addition we cannot measure both  $p$  and  $q$  with arbitrary precision. Note the similarity between (7) and (8) and the quantal noise in heterodyne and homodyne receivers. In the former case, both in phase and quadrature components of the field can be used to carry information, and we find a minimum effective input noise energy per mode of  $h\nu$ , just as in an ideal laser amplifier.

In the latter, while we give up all notion of using the quadrature field component, we find the minimum effective quantal input noise reduced by exactly a factor of two, to  $h\nu/2$ .

With one further step, we shall have completed the formalities. Equation (4) gives the conditional probability result for the case in which the transmitter can prepare pure states and the receiver can make what we may call ideal measurements, namely the best measurements of a particular type that quantal theory will allow. But in real systems neither of these is the case. For example, the transmitter may attempt to prepare a coherent state  $|\bar{p}, \bar{q}\rangle$ , but what appears is a state  $|\bar{p}+p_n, \bar{q}+q_n\rangle$ , where  $p_n$  and  $q_n$  represent the random-noise field added in the path. Given the statistical probability distribution  $P(p_n, q_n)$ , Eq. (4) must now be modified to read

$$P(r|\bar{q}, \bar{p}) = \sum_{p_n, q_n} P(p_n, q_n) |\langle r | \bar{p}+p_n, \bar{q}+q_n \rangle|^2 \quad (9)$$

which may be rearranged to read

$$P(r|\bar{q}, \bar{p}) = \langle r | \rho(\bar{p}, \bar{q}) | r \rangle, \quad (10)$$

where

$$\rho(\bar{p}, \bar{q}) \equiv \sum_{p_n, q_n} P(p_n, q_n) |\bar{p}+p_n, \bar{q}+q_n\rangle \langle \bar{p}+p_n, \bar{q}+q_n|$$

is the statistical density operator representing the noisily prepared system. For thermal noise the distribution  $P(p_n, q_n)$  is Gaussian, just as in the classical theory, and, of course, then the double sum indicated in (9) and (10) becomes a double integral. Finally, the receiver may not be ideal. For example, its meters may be perturbed by unessential noise, or perhaps the observer simply does not care to distinguish each possible value of  $r$ . He might then obtain a more coarse-grained conditional probability distribution such as

$$P(\Delta|\bar{q}, \bar{p}) = \sum_{r \in \Delta} \langle r | \rho(\bar{p}, \bar{q}) | r \rangle \quad (11)$$

where the set of  $\Delta$ 's represent a complete set of mutually exclusive ranges of  $r$ . Equation (11) can be re-expressed in the form

$$P(\Delta|\bar{q}, \bar{p}) = \text{Trace} [\rho(\bar{p}, \bar{q}) \sigma(\Delta)], \quad (12)$$

where

$$\sigma(\Delta) = \sum_{r \in \Delta} |r\rangle \langle r|.$$

Equation (12) is the generalization of (4) to include noisy preparation and measurement of the quantal systems. Without ambiguity, we can rewrite (12) in the general notation of (4), namely

$$P(r/t) = \text{Trace} [\rho(t)\sigma(r)]. \quad (13)$$

To recapitulate,  $\rho(t)$  is a density operator for the information-carrying quantal system normalized to unit trace, which is associated with the particular transmitter setting  $t$ , and includes the effects of the lossy and noisy transmission path. And  $\sigma(r)$  is a density operator for the information-carrying quantal system, but is associated with the particular receiver reading  $r$ . The set  $\sigma(r)$  must be complete, in the sense that

$$\sum_r \sigma(r) = I, \quad (14)$$

so that the conditional probability sums to unity over  $r$ . The possible set of measurement operators, however, need not be restricted to orthogonal sets.

If (13) encompasses all possible measurements of quantal observables, a backwards question may be asked; namely, Does every complete set of density operators  $\sigma(r)$  correspond to some measurement? The answer is not obvious, and deserves some thought. A simple example may serve to illustrate the problem. Helstrom has considered a binary system in which the transmitter prepares each oscillator either in the ground state  $|0, 0\rangle$  or in a coherent state  $|\bar{p}, \bar{q}\rangle$  with some a priori probability  $\lambda_0$  for the ground state, and  $\lambda_1 = 1 - \lambda_0$  for the excited state. For this problem the two states by themselves form a complete set. Helstrom has shown that a certain pair of orthogonal measurement states, which are the eigenstates of the Hermitian operator

$$\lambda_0 |0, 0\rangle \langle 0, 0| - \lambda_1 |\bar{p}, \bar{q}\rangle \langle \bar{p}, \bar{q}|, \quad (15)$$

minimize the average error probability for the system. These measurement states have the form

$$|r\rangle = a_r |0, 0\rangle + b_r |\bar{p}, \bar{q}\rangle, \quad r = 1, 2. \quad (16)$$

The results above can be shown solely on the basis of Eq. (4), without any reference to how one might make such a measurement, even in a gedanken world. In this case, however, the measurement turns out to be possible in principle if one admits the possibility only of energy measurements. Without loss of generality we can set  $\bar{p} = 0$ . We must wait only a suitable fraction of a cycle for this to be true, since  $\bar{p}$  and  $\bar{q}$  obey the classical harmonic equations of motion, while the oscillator is left undisturbed. When  $\bar{p} = 0$ , then  $a_r$  and  $b_r$  turn out to be real, and hence

$$\psi_r(q) = \langle q|r\rangle = a_r \psi_{0,0}(q) + b_r \psi_{0,\bar{q}}(q) \quad (17)$$

is a real wave function in which there is a weighted sum of two coherent-state wave functions. Thinking of our oscillator as a mechanical oscillator, we can alter suddenly the potential energy function so that one of the two measurement states (16) becomes an energy eigenstate of the new Hamiltonian. Energy eigenstates are found from Schrödinger's equation (for unit mass)

$$\frac{\hbar^2}{2} \frac{\partial^2 \psi}{\partial q^2} + (W - V(q))\psi = 0. \quad (18)$$

Hence, given  $\psi_r(q)$ , we have for the new potential

$$V(q) = W + \frac{\hbar^2}{2\psi_r(q)} \frac{\partial^2 \psi_r(q)}{\partial q^2}.$$

It does not matter whether the other measurement state is an energy state for the new potential or not. We calculate all of the energy states for this new potential, and then having made the sudden change, make an energy measurement at leisure. Either the oscillator was in the singled-out  $\psi_r(q)$  state or not, and we find the answer we want. This idea of making measurements via a sudden alteration of the potential function in the Hamiltonian of the system involved is due to W. E. Lamb.

Based on (13) as the fundamental premise of quantal communication, a number of general things can be said. First, an upper bound to the information rate of the channel per quantal system (per harmonic oscillator here) exists which is based solely on the set of prepared density operators  $\rho(t)$  and the a priori probability  $\lambda(t)$  that each is used. In quantal theory, the entropy function  $S(\rho)$  is given by

$$\begin{aligned} S(\rho) &= -\text{Trace} [\rho \log \rho] \\ &= -\sum_k \rho_k \log \rho_k, \end{aligned}$$

where the  $\rho_k$  are the eigenvalues of  $\rho$ . The upper bound is given by

$$B \equiv S(\bar{\rho}) - \sum_t \lambda(t) S(\rho(t)),$$

where

$$\bar{\rho} \equiv \sum_t \lambda(t) \rho(t).$$

The properties of the transmitter determine the set  $\rho(t)$ . Thus we can optimize the transmitter in a rough way by choosing a set of  $\rho(t)$  and corresponding  $\lambda(t)$  so as to maximize  $B$ . Given any particular variety of receiver, the set of measurement operators  $\sigma(r)$  can be determined, and then the information capacity of that channel can be determined from the set of conditional probabilities  $P(r/t)$  and the  $\lambda(t)$ , and then we can see how good the channel is by comparison with  $B$ .

Helstrom has considered  $M$ -nary channels in which the set  $\rho(t)$  is a set of  $M$  pure quantal states

$$\rho(t) = |t\rangle \langle t|, \quad t = 1, 2 \dots M.$$

In this case any set of  $M$  measurement states  $|r\rangle$  always forms an orthogonal set, and in the case of the binary channel he has been able to find the pair that minimizes the error probability. We have discussed how the binary measurement might in principle be accomplished, but the question of what the optimum set of measurement states is for the  $M$ -nary channel, and whether the Hermitian operator that generates them is in any sense an "observable," remain unanswered. Also, the question of how additive noise affects Helstrom's results should be looked into.

### III. BASIS FOR CLASSICAL FIELDS AND POISSON STATISTICS

To return to a more practical level, we may consider the class of prepared density operators for a harmonic oscillator which can be described in the so-called  $P$  representation of Glauber, namely

$$\rho = \iint P(p, q) |p, q\rangle \langle p, q| dpdq,$$

with a real, non-negative probability function  $P(p, q)$ . The probability distribution for the number of quanta in such an oscillator is given by

$$\langle n | \rho | n \rangle = \iint P(p, q) \frac{m^n}{n!} e^{-m} dpdq,$$

where

$$m \equiv \frac{1}{2\hbar\omega} (p^2 + \omega^2 q^2)$$

is the average energy of the state  $|p, q\rangle$  expressed in photon units. Thus the photon number distribution is simply a Poisson convolution of the "classical" distribution  $P(p, q)$ . An oscillator weakly coupled to a heat bath and in equilibrium with it is statistically described by a Gaussian  $P$  function,

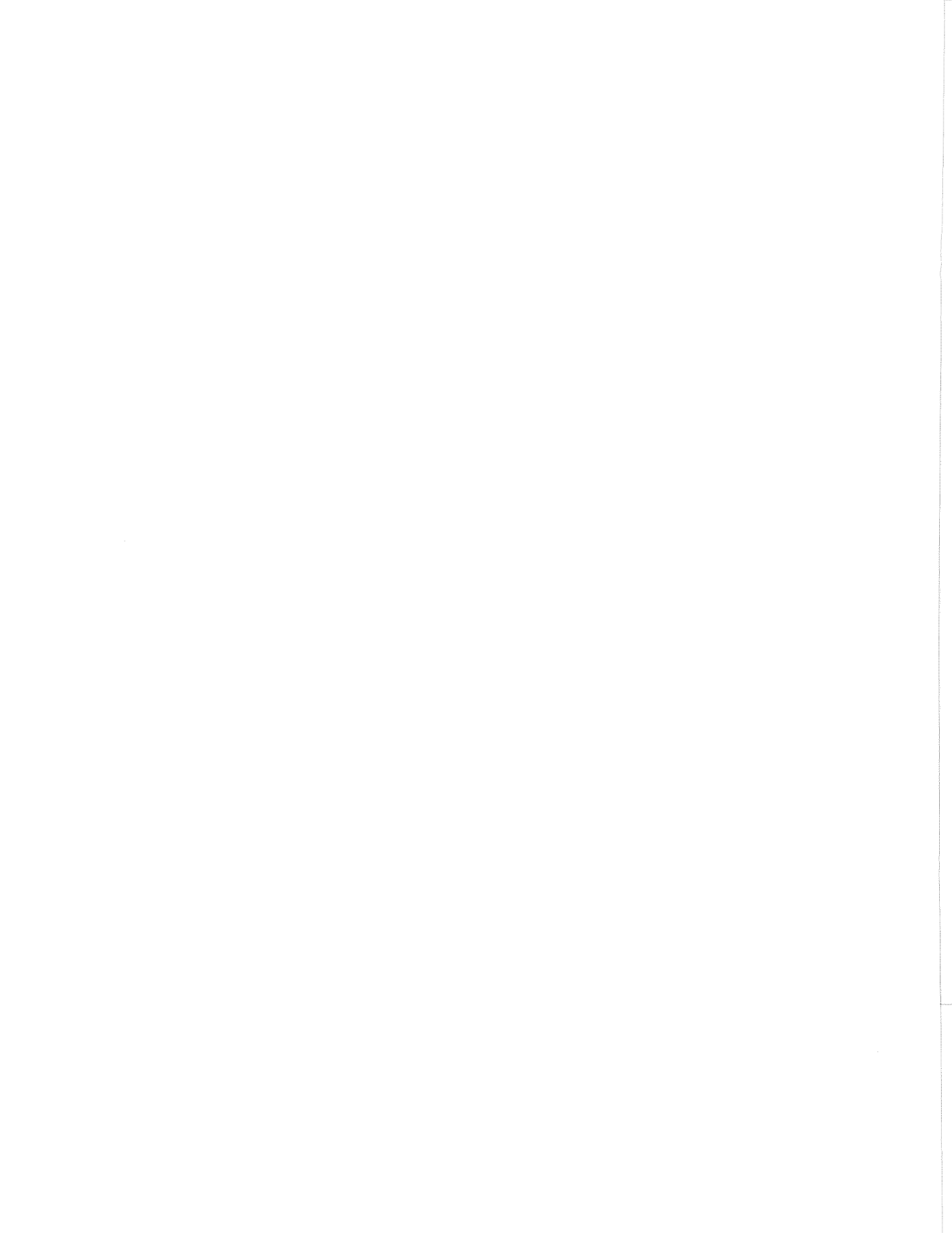
$$P(p, q) \propto \exp \left\{ - \frac{p^2 + \omega^2 q^2}{2\bar{n}\hbar\omega} \right\},$$

where

$$\bar{n} = \left[ \exp\left(\frac{\hbar\omega}{kT}\right) - 1 \right]^{-1}$$

The field emitted by an ideal oscillator is very nearly described by a  $P$  function that is a  $\delta$  function. It is just this  $P$  distribution that we can consider as describing "the" classical field distribution. It obeys the classical laws of motion of the harmonic

oscillator throughout all of the linear operations that are performed at the receiver. The Poisson convolution makes its appearance only at the detecting element, where in essence we ask the question, How many photons are there? While there are many conceivable fields that are not described by such non-negative P functions, it is not likely that they will ever be encountered in any practical problem in communication.



# Appendix G

## Equivalence of Transmitter and Receiver Antenna Gain in a Random Medium

David L. Fried

We are concerned here with the relationship between the performance of a ground-based optical heterodyne receiver and a ground-based laser transmitter operating through the atmosphere in a ground-space link. In particular, we are interested in the effect of the random refractive index on the system performance. We shall show that the atmosphere has an identical influence on both systems.

For the purpose of analysis, rather than consider the actual configuration of an optical heterodyne receiver we shall consider a much simpler system composed of a diffraction-limited lens and a pinhole, at a distance one focal length away from the lens. The pinhole diameter is small enough compared with the Airy disk, which the lens might form, that the pinhole diameter may be considered to be effectively of zero extent. In back of the pinhole a photodetector is placed to measure the optical power coming through the pinhole. The significance of this very simple configuration is that it corresponds exactly to the heterodyne receiver. It is easy to show, either formally or heuristically, that the wavefront distortion effect of atmospheric refractive index inhomogeneities is identical in reducing heterodyne signal power and in reducing the optical power that can pass through the pinhole. The simplest proof of this follows from recognition that (i) the optical power through the pinhole measures the Strehl definition of the composite lens-atmosphere optical system, (ii) the Strehl definition is equal to the integrated MTF of the composite system,<sup>1</sup> and (iii) the expression for the integrated MTF as given elsewhere<sup>2</sup> is identical in its form to the expression for the heterodyne receiver relative to signal-to-noise ratio as previously given.<sup>3</sup> Although these expressions, as given previously,<sup>2,3</sup> apply to values averaged over an ensemble of atmospheric distortions, equivalent results can be obtained for the instantaneous, or unaveraged performance.

In place of the laser transmitter, we consider the same lens/pinhole arrangement but with a light source placed in back of the pinhole. The equivalence to a diffraction-limited laser transmitter is obvious. In this case we seek to determine how atmospheric turbulence affects the on-axis far-field intensity.

For the purpose of analysis we define the three coordinate systems  $\vec{x}_1$ ,  $\vec{x}_2$ , and  $\vec{\theta}$ , which are (for  $\vec{x}_1$ ) linear position in the plane of the pinhole, denoted  $X_1$ , (for  $\vec{x}_2$ ) linear position in the plane  $X_2$ , a plane just above the atmosphere, and (for  $\vec{\theta}$ ) angular position in the far field away from the atmosphere. ( $\vec{\theta}$  is a vector in the direction space just as  $\vec{x}$  is a vector in the ordinary space  $X$ .) The pinhole and lens define a center line for the system.  $\vec{x}_1 = 0$  (or  $\vec{x}_2$ ) occurs where the center line passes through  $X_1$  (or  $X_2$ ), and  $\vec{\theta} = 0$  is the direction vector parallel to the center line.

Our analysis is based on the well-known fact that if  $G(\vec{x}_1, \vec{x}_2)$  is the Green's function



that propagates a wave function  $u_1(\vec{x}_1)$  or  $X_1$  into the wave function  $u_2(\vec{x}_2)$  or  $X_2$ ; that is

$$u_2(\vec{x}_2) = \int_{X_1} d\vec{x}_1 G(\vec{x}_1, \vec{x}_2) u_1(\vec{x}_1), \quad (1)$$

then  $G^*(x_1, x_2)$  propagates a wave function from  $X_2$  to  $X_1$ ; that is,

$$u_1(\vec{x}_1) = \int_{X_2} d\vec{x}_2 G^*(\vec{x}_1, \vec{x}_2) u_2(\vec{x}_2). \quad (2)$$

The Green's function  $G$  reflects all of the boundary conditions associated with the region between  $X_1$  and  $X_2$ , including both the presence of the lens and the irregular refractive-index pattern of the atmosphere.

The wave function  $u_3(\vec{\theta})$  associated with an infinite plane wave beyond  $X_2$  traveling in the direction  $\vec{\theta}$  may be computed from Eq. (3),

$$u_3(\vec{\theta}) = \int_{X_2} d\vec{x}_2 \exp(ik\vec{\theta} \cdot \vec{x}_2) u_2(\vec{x}_2), \quad (3)$$

where we have suppressed a constant of proportionality as being irrelevant to the details of our calculation. The inverse relationship permitting us to calculate  $u_2(\vec{x}_2)$  for a plane wave traveling from the direction  $\vec{\theta}_2$ , incident on  $X_2$  is

$$u_2(\vec{x}_2) = \int_{\textcircled{H}} d\vec{\theta} \exp(-ik\vec{\theta} \cdot \vec{x}_2) u_3(\vec{\theta}). \quad (4)$$

To analyze the problem of transmitter performance, we assume that  $u_1(\vec{x}_1)$  is a delta function centered at  $\vec{x}_1 = 0$ ; that is,

$$u_1(\vec{x}_1) = \delta(x_1) \quad (5a)$$

$$u_3(\vec{\theta}) = (\vec{\theta} - \vec{\theta}_0), \quad (5b)$$

and solve for the power density in the far-field direction  $\vec{\theta}_0$ , that is, for  $u_3^*(\vec{\theta}_0) u_3(\vec{\theta}_0)$ . This is a measure of the efficiency of a ground-based transmitter in directing optical power to a receiver in space, an angular distance  $\vec{\theta}_0$  away from the optical center line. This is directly proportional to the transmitter antenna gain,  $G_T$ . Combining Eqs. (5a), (1), and (3), we get

$$u_3(\vec{\theta}_0) = \int_{X_2} d\vec{x}_2 \exp(ik\vec{\theta}_0 \cdot \vec{x}_2) G(0, \vec{x}_2), \quad (6a)$$

so that  $G_T$  is proportional to

$$u_3^*(\vec{\theta}_0) u_3(\vec{\theta}_0) = \int_{X_2} \int_{X'_2} d\vec{x}_2 d\vec{x}'_2 \exp[ik\vec{\theta}_0 \cdot (\vec{x}_2 - \vec{x}'_2)] G(0, \vec{x}_2) G^*(0, \vec{x}'_2). \quad (7a)$$

To analyze the problem of receiver antenna gain, we assume that  $u_3(\vec{\theta})$  is a delta function centered at  $\vec{\theta}_0$ , corresponding to a transmitter in space an angular distance  $\vec{\theta}_0$  away from the ground-based receiver center line, and ask about the power density at the pinhole; that is, we solve for  $u_1^*(0) u_1(0)$ . This quantity determines the received signal power and is proportional to the heterodyne receiver antenna gain,  $G_R$ . By combining Eqs. (5b), (2), and (4), we get

$$u_1(0) = \int_{X_2} d\vec{x}_2 \exp(-ik\vec{\theta}_0 \cdot \vec{x}_2) G^*(0, \vec{x}_2). \quad (6b)$$

Hence the receiver antenna gain,  $G_R$ , is proportional to

$$u_1^*(0) u_1(0) = \int_{X_2} \int_{X'_2} d\vec{x}_2 d\vec{x}'_2 \exp[ik\vec{\theta}_0 \cdot (\vec{x}_2 - \vec{x}'_2)] G(0, \vec{x}_2) G^*(0, \vec{x}'_2). \quad (7b)$$

Considering the fact that all of the atmospheric refractive-index irregularities are contained in the Green's function  $G$ , and noting the exact equivalence of the right-hand sides of Eqs. (7a) and (7b), we can conclude that a given optical system, whether used as a diffraction-limited transmitter or a heterodyne receiver will be equally affected by the atmosphere in terms of instantaneous and average optical signal power transfer. This means, for example, that if a large transmitter and heterodyne receiver share common optics, then if, because of a particular atmospheric refractive index configuration, the heterodyne signal increases in value (or fades) then the fraction of transmitter optical power reaching the space platform will increase in value (or fade). Such a situation offers, for instance, the possibility of burst communication. It also means that if the optical axis is rapidly adjusted to maximize the heterodyne signal strength, thereby tracking out some of the atmospherically induced wavefront distortion, this pointing will automatically maximize transmitter performance.

A particularly useful consequence of the result that we have obtained is that we can use the results of the previous paper,<sup>3</sup> which describes optical heterodyne receiver performance with atmospheric effects considered, or even of the other paper,<sup>2</sup> which analyzes the performance of an imaging system to predict the average effect of the atmosphere on a laser transmitter. We could, for example, conclude from this that a laser transmitter would not be limited by atmospheric effects if its diameter were  $\sim r_0/2$ , where  $r_0$  is a coherence length previously defined.<sup>3</sup> For a transmitter of diameter  $r_0$

the average antenna gain will be down approximately a factor of 2 from the ideal achievable if there were not atmospheric refractive-index inhomogeneities.

The reciprocity between transmitter and receiver that we have demonstrated also permits us to use other results<sup>4</sup> to calculate the variance of transmitter antenna gain as affected by the changing pattern of atmospheric inhomogeneities.

It should be noted that although this analysis has been couched in terms of a ground-space link, it should be possible to modify the proof so that it will apply equally well for the two ends of the link both in the atmosphere and on the ground.

#### References

1. E. L. O'Neill, Introduction to Statistical Optics (Addison-Wesley Publishing Co., Reading, Mass., 1963).
2. D. L. Fried, J. Opt. Soc. Am. 56, 1372 (1966).
3. D. L. Fried, Proc. IEEE 55, 57 (1967).
4. D. L. Fried, J. Quant. Elec. 3, 213 (1967).

## Appendix H

### Atmospheric Turbulence Considerations for a Ground-Based Laser Transmitter in a Ground-Space Link

D. L. Fried

The kinds of questions that we wish to be able to answer about atmospheric turbulence effects on a ground-based laser transmitter concern (i) average antenna gain, (ii) the variance and other statistical aspects of antenna gain, and (iii) to what extent can antenna gain be adjusted by matching the transmitted wavefront to the distorted wavefront of a beacon arriving from the receiver position, and what is the speed with which these effects fluctuate.

There are two ways in which we can approach the problem. The first method involves direct calculation of laser beam statistics starting from Schmeltzer's<sup>1</sup> results or performing an equivalent more extensive treatment. It is then necessary to carry out numerical evaluations from this. Published results are available for laser beam intensity variance<sup>2,3</sup> for both ground-to-ground and ground-to-space links. The published results are available in the form of log-amplitude variance data,  $C_\ell(0)$ . To obtain intensity variance data  $\sigma_I^2$ , we use the assumption<sup>4</sup> that the fluctuations are log-normally distributed so that

$$\sigma_I^2 = \exp[4 C_\ell(0)] - 1.$$

Predicted values for transmitter antenna gain,  $G_{\text{atmos}}$ , relative to the ideal gain,  $G_{\text{ideal}}$ , obtainable in the absence of atmospheric turbulence can be determined from the log-amplitude mean,  $\bar{\ell}$ , and variance,  $C_\ell(0)$  (again under the assumption of a log-normal distribution), from the formula

$$G = \frac{G_{\text{atmos}}}{G_{\text{ideal}}} = \exp\{2[C_\ell(0) + \bar{\ell}]\},$$

where gains are directly proportional to power (rather than in the dB form, proportional to the logarithm of power). The value of  $\bar{\ell}$  has recently been calculated, from Schmeltzer,<sup>1</sup> by F. Gebhart (O. S. U. Electro Sciences Laboratory) and the results should soon be published. The calculations, however, only apply to a horizontal path and so are not directly applicable to our ground-space link problem.

Because a Gaussian tapered aperture is assumed in all of the published work on laser beam statistics, and it appears that a uniformly illuminated finite-diameter source may have significantly different scintillation statistics, we do not believe the results of Fried<sup>3</sup> are directly applicable to our problem.<sup>5</sup> As a consequence of this and the lack of calculated data for  $\bar{\ell}$  on the ground-to-space path<sup>6</sup> there is little that we can determine

with certainty from a direct study of laser beam propagation statistics. Consequently, most of what we can say about transmitter performance must be deduced from our second approach to the problem. This approach is based on the exact reciprocity, as demonstrated in the Appendix, between optics acting as a transmitter or as a receiver over a given randomly inhomogeneous path. This very powerful result permits us to consider the receiver in space as replaced by a laser source, and the transmitter on the ground, whose performance we wish to study, as replaced by a heterodyne receiver, and equate instantaneous, as well as average atmospheric effects, on the heterodyne receiver antenna gain with what the ground-based optics could achieve if it were functioning as a transmitter. This immediately permits us to apply the results of Fried<sup>7,8</sup> to the evaluation of the transmitter problem. Moreover, since it can be shown that imagery resolution as defined by Fried<sup>9</sup> is equivalent to antenna gain as applied to heterodyne reception, we can also use these results<sup>9</sup> to study transmitter performance. Fried's results<sup>7-9</sup> together permit us to talk about transmitter antenna gain, gain variance, and the effects of rapid angle tracking on gain, for a uniformly illuminated transmitter aperture operating on a ground-to-space link.

Using this approach, we draw the following conclusions regarding the performance of a continuous (as distinct from burst operation) static (not angle-tracking) transmitter. We restrict attention here and throughout the rest of this appendix to wavelengths of approximately  $0.5 \mu$  and  $10.6 \mu$ . The maximum average achievable antenna gain is that associated with a diameter  $r_o$ . For  $\lambda = 0.5 \mu$ ,  $r_o$  is of the order of a few inches. Between one and four inches may be taken as a good nominal range of values, although the exact value will depend on local atmospheric conditions, the zenith angle of the path, and possibly nearby geographic features. For  $\lambda = 10.6 \mu$ , we expect  $r_o$  to be larger by a factor of  $(10.6/0.5)^{6/5}$ , with a nominal range from just under 1 m to possibly as much as 3 m. (It is interesting to note that since ideal antenna gain is proportional to  $(\lambda/D)^2$ , the limiting antenna gain at  $10.6 \mu$  will be better than that at  $0.5 \mu$  by a factor of  $(10.6/.5)^{2/5} \approx 3$ .) The variance of antenna gain will show, however, a strong dependence on the ratio of diameter to  $r_o$ . For  $D = r_o/2$ , the rms fluctuation will be approximately 30%, but for  $D = r_o$  it will be approximately 70%. In view of the fact that we expect the fluctuations to be log-normal, the 70% rms fluctuation is probably intolerable, and even the 30% represents a problem, although probably a tolerable one. In forming these estimates for scintillation, we have assumed that there was no gross intensity fluctuation of the kind that in the reciprocal problem would make the total power entering the aperture vary. A consideration of stellar scintillation data and our estimate of its wavelength dependence (that is,  $C_\theta(0) \propto \lambda^{-7/6}$ ) indicates that for  $\lambda = 0.5 \mu$  we may have an additional 30% rms fluctuation to convolve with the effect discussed above (giving a total of approximately 40% for  $D = r_o/2$ ), but for  $\lambda = 10.6 \mu$  there should be little additional effect. We conclude that for a static laser transmitter to minimize fluctuations of gain, the optics diameter should be  $\sim r_o/2$ , in which case we shall have 30-40% rms fluctuation. It may be desirable to build a transmitter with a variable diameter.

If a transmitter is designed to operate on a burst communications basis, then it would be desirable to have a diameter much larger than  $r_0$  and maximize the antenna gain fluctuations. Transmission would then be initiated by monitoring reception from a beacon on the space receiver. Whenever the beacon signal strength rose above a critical value a burst of transmission would be initiated. It is difficult to estimate the mean time between the occurrence of burst conditions, but it is reasonable to estimate that it is of the order of  $D/v$  or  $r_0/v$  when  $v$  is the velocity with which the distorting medium moves through the optical beam. We may estimate  $v$  as being comparable to the wind speed. It is questionable whether, at the low rate at which bursts occur, adequate message memory is available for this kind of operation.

It is possible that if the transmitter pointing is slaved to a receiver that tracks a beacon on the space station, the transmitter pointing will track out that part of the atmospheric disturbance which appears as wavefront tilt. Using the results of Fried<sup>9</sup> and noting that we expect very much less than 1 Np rms log-amplitude fluctuation in all cases (so that we use what are referred to as the near-field results<sup>9</sup>), we may expect average performance with a diameter  $D \approx 1.75 r_0$  as close to the diffraction limit as the  $r_0/2$  diameter provides in the static case, that is, a gain down  $\sim 30\%$  from the diffraction limit for that diameter. We believe, but this is only conjecture, that this means a tracking system with  $D = 1.75 r_0$  will have the same rms fluctuation of gain that a static system with  $D = r_0/2$  has, namely 30%. Experimental data for the spectrum of image motion indicates that most of the motion is below 10 Hz, and all of it below 100 Hz in the visible range. We believe the frequency will be lower by a factor of approximately 20 for 10- $\mu$  observation.

Perhaps the most serious questions to be raised concerning beacon angle-tracking transmitters involves the question of "point ahead." To reach the beacon, the transmitted signal must be pointed ahead of the beacon an angle  $\theta = 2V/c$ , where  $V$  is the relative velocity perpendicular to the line of sight between the transmitter and the receiver, and  $c$  is the speed of light. For a nominal velocity  $V = 10^4$  miles/hour,  $\theta = 3 \times 10^{-5}$  rad. We may take the scale height of the atmosphere as  $H = 6300$  m, the distance in which the density of the atmosphere falls to  $e^{-1}$  of its sea-level value. We argue that most of the wavefront distortion is produced where the atmosphere is most dense, and if the received and transmitted beams are nearly congruent over that range, that they will see the same atmospheric effects. Hence  $H\theta \approx 20$  cm. This is larger than  $1.75 r_0$  from most cases when  $\lambda = 0.5 \mu$ , and is much smaller than  $1.75 r_0$  for  $\lambda = 10.6 \mu$ . Thus it appears likely that point-ahead problems will prevent any significant improvement in transmitter antenna gain by beacon angle-tracking in the visible range, but will definitely be usable at 10  $\mu$ . We may comfortably expect to be able to use a 1.5-m or 2.0-m diameter, beacon-tracking, 10- $\mu$  laser transmitter at 70% of diffraction-limited antenna gain, with an estimated 30% rms fluctuation of gain.

## References

1. R. A. Schmeltzer, *Quart. Appl. Math.* 24, 339 (1967).
2. D. L. Fried and J. B. Seidman, *J. Opt. Soc. Am.* 57, 181 (1966).
3. D. L. Fried, *J. Opt. Soc. Am.* 59, 980 (1967).
4. In a paper by R. Mitchell (Aerospace Corporation) soon to be published in *J. Opt. Soc. Am.* it is shown that the sum of several log-normal variables is very well described by considering it to be log-normal. We extrapolate this to apply to all phase-normal random variables, note that the product of phase-normal variables is definitely phase-normal, and infer that in all cases of interest the effect of the atmosphere is to produce phase-normal effects and a log-normal distribution of signal strength in any instrument. This conclusion is in reasonable agreement with available experimental results.
5. To extend these results so that it would be applicable to our interests the work of Schmeltzer<sup>1</sup> would have to be expanded to include as a source, not just the lowest order Gauss-Hermite laser excitation pattern but all of the higher order patterns.
6. At present, there is sufficient question about the vertical distribution of turbulence in the atmosphere that the value of calculating  $\bar{\ell}$  for the vertical path is doubtful until this question is resolved.
7. D. L. Fried, *Proc. IEEE* 55, 57 (1967).
8. D. L. Fried, *J. Quant. Elec.* 3, 213 (1967).
9. D. L. Fried, *J. Opt. Soc. Am.* 56, 1372 (1966).

# Appendix I

## Spatially Distributed Coherent Heterodyne Reception through a Fluctuating Atmosphere

P. E. Green

In order to develop maximum SNR at the receiver output, it is most desirable to sample more of the received signal wavefront than simply that over which phase and amplitude remain more or less constant. The means to be described are available in principle to extend the usual notion of coherent heterodyne optical reception so as to embody the optimum use of energy over a large area of wavefront.

We shall use as our starting point receivers (discussed elsewhere<sup>1</sup>) that receive a signal at each of D spatially separate detectors, pass each detector output through appropriate matched filter-envelope detector elements (or the equivalent correlators), and then, after a nonlinear weighting of each signal, combine outputs, as shown in Fig. I-1b. (For convenience, we shall assume throughout that binary frequency-shift-keyed signals are transmitted, as in Fig. I-1a.) Each input element is a detector whose aperture is a portion of the received wavefront roughly equal to the area over which phase is constant. The nonlinear operations serve the function of emphasizing in the output the contribution of those elements whose signal strength is greatest while de-emphasizing the weaker ones.

We are motivated here by the notion that a considerable improvement in performance against noise may be expected from a receiver which, instead of using nonlinear elements to do the weighting as in Fig. I-1, multiplies each channel before combining by the long-time average strength of that channel. To use all of the available data in the received wavefront, it is desirable not only to do such a weighting of individual spatial regions by the most noiseless estimates possible of their current amplitudes but also to estimate the (slowly varying) received phase at each element and compensate for this.\*

We discuss two procedures for accomplishing the phase compensation and amplitude-weighting functions. First, we describe a receiver structure that is discrete in space; that is, it is a mosaic of elements, as is the structure of Fig. I-1. The second structure is identical in function to the first but treats the spatial coordinate as a continuum. It appears on cursory examination that the spatially discrete version could be built now with state-of-the-art components, whereas the spatially continuous receiver, although visualizable in detail, requires components that do not exist at present.

---

\*Devices that combine signals in such a way may be viewed as "optimum multipath receivers," "adaptive arrays," "optimum diversity combiners," "phase-front correcting antennas," etc., depending on one's point of view.



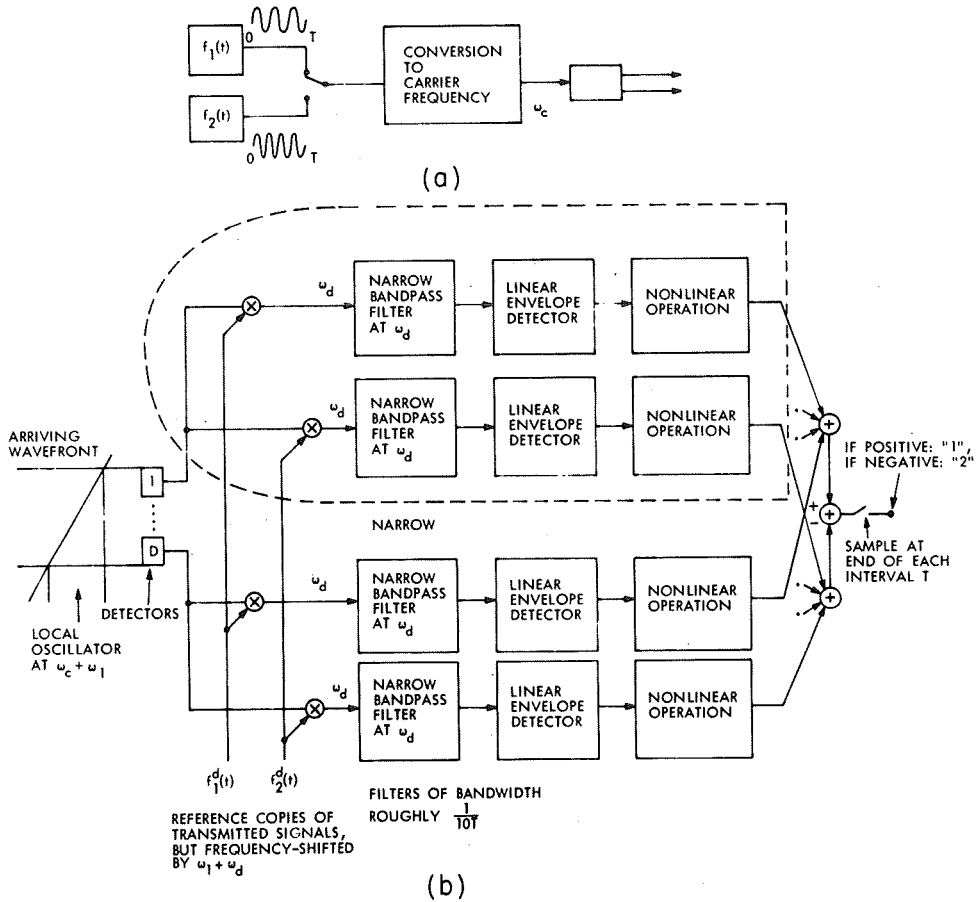


Fig. I-1. (a) Transmitter. (b) Optimum receiver when averaging of slow channel fluctuations is ignored ("min-max phase" receiver – Kennedy and Hoversten<sup>1</sup>). Bandpass correlators are shown instead of matched filters.

A. Spatially Discrete Version

A receiver<sup>2</sup> that weights each of the  $D$  diversity element outputs by a smoothed estimate of its current amplitude, and at the same time corrects phase differences from element to element, can be realized by replacing the operations within the dashed line of Fig. I-1 with the operations shown in Fig. I-2. Figure I-2 is labeled to show the amplitude, phase, and frequency at various points; the bar indicates a time averaging that is assumed to take place by virtue of the narrowness of the "very narrow bandpass filter" or channel-measuring filter shown in the center. The filter is shown driven all the time, no matter which signal "1" or "2" is transmitted. For optimum performance the squared-amplitude passband characteristic of this filter should match the fluctuation power spectrum of the path. In each of the two final bandpass filters all channel

contributions add in phase with the right amplitude weighting, as follows. Note that the input to each bandpass filter has the desired amplitude (product of current value  $A$  and average  $A$ ) and that the phase is constant at  $\phi_{\Delta}$  (that of the oscillator at  $\omega_{\Delta}$ ) plus the small fluctuation ( $\phi - \bar{\phi}$ ). By driving all  $D$  elements that are like that in Fig. I-2 from the same oscillator at  $\omega_{\Delta}$ , all channels can be added together before the bandpass filtering, thereby achieving the desired phase realignment.

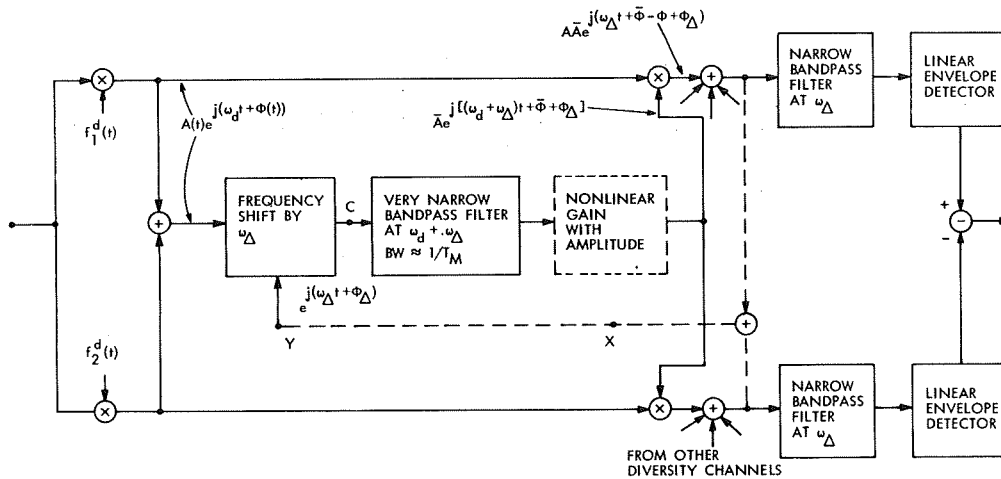


Fig. I-2. A modification of one element of the structure of Fig. I-1b to include channel amplitude weighting and phase correction. (The final narrow BPF's and linear envelope detectors appear only once. All else is iterated  $D$  times to form the complete receiver.)

At least two questions have to be investigated further to determine whether this receiver is optimum for the optical channel. Note that in going from Fig. I-1 to Fig. I-2 some sort of "nonlinear operation" has been retained, and the first question concerns whether this is a requirement in the optical transmission case. If the optical channel had been nonfluctuating, the optimum receiver (under the assumption of only additive white Gaussian noise) would have been the single elements within the dotted line of Fig. I-1, with the nonlinearity replaced by a linearity. If the channel fading statistics were Gaussian, the nonlinearity in Fig. I-2 becomes a linearity and so all diversity channels may be combined before bandpass filtering (as just described) with the nonlinear gain element taken out. One may conjecture that when the fading statistics are log-normal, as is presumed to be the case for the optical atmospheric channel, some form of nonlinear gain operation may be needed. This is tentatively shown by the dashed-line "nonlinear gain" elements in Fig. I-2. A second point to be mentioned is that, as shown, an equal smoothing time for amplitude and phase is implied. Equality of fluctuation rates of amplitude and phase, a property of Gaussian fading, is not a property of the optical channel. These points and presumably others, concerning the required modification of

the Fig. I-2 receiver, optimum for Gaussian fading, to convert it to a receiver that is optimal for log-normal fading ought to be investigated further.

The biggest problem in actually building such a system would appear to be the relative frequency stability of the transmitter and receiver. Assuming that the transmitter frequency is constant, we note that the various conversions ahead of the "very narrow-band filter" must translate the signal spectrum to an accuracy of, say, one-tenth of the bandwidth  $1/T_M$ . Under the assumption of atmospheric fluctuation rates of roughly 1 kHz, this implies frequency stability to 100 Hz. This order of raw laser frequency stability might be achievable, but it appears that the use of a locked loop (completed via path A-B in Fig. I-2) would be more convenient. In space applications this approach will be required because of the large Doppler shifts.

If frequency stability problems become too great, one can always fall back on the transmitted-reference technique. Here, we receive on each detector element not only the modulated signal but also a reference sinusoid transmitted at the desired nearby frequency (close enough so that fluctuations are highly correlated). The receiver reference tone is then injected at point C of each diversity element (Fig. I-2) in place of the signal shown. This wastes part of the transmitter power, but might be interesting in certain cases. In particular, this artifice appears attractive for the spatially continuous system to be discussed next.

#### B. Spatially Continuous Version

In Fig. I-2 the key element was the feed-forward loop shown isolated in Fig. I-3a. This operation had to be iterated as many times as there were diversity elements, in the structure proposed in the preceding section. A more elegant approach and one that might become quite necessary if the number of diversity elements becomes truly large\* is to set up such a feed-forward loop of the entire spatially distributed field. To do this in a manner exactly analogous to Fig. I-3a requires that the received field (amplitude and phase as a function of time and two spatial coordinates) be both frequency-shifted and time-averaged before being injected as a "local oscillator" in a final conversion. Of these two operations, only the latter appears possible. Therefore, we must fall back on the transmitted-reference technique just mentioned. For this situation, Fig. I-3a becomes Fig. I-3b.

The corresponding two-dimensional receiver is shown in Fig. I-3c. The pilot signal excites a laser amplifier whose bandwidth (at a frequency shown as  $\omega_1 + \omega_2$ ) is the required  $1/T_M$  and whose apertures at the two ends are adjusted so that it can amplify many modes simultaneously (each mode corresponding to a different spatial Fourier component of the received wavefront). We require an amplifier, rather than simply a resonant cavity, so that the local-oscillator signal amplitude will be very large

---

\*Systems of the Rake<sup>2</sup> type, based on Fig. I-2, have been built at megacycle frequencies embodying 50 or more diversity elements.

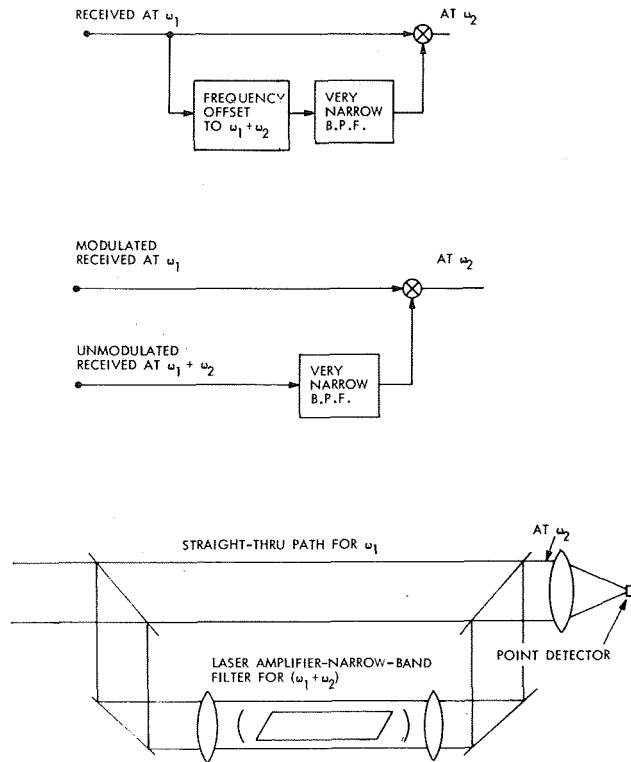


Fig. I-3. (a) Feed-forward loop.  
 (b) Alternative approach using transmitted pilot sinusoid (unmodulated at  $\omega_1 + \omega_2$ ).  
 (c) Spatially continuous version of Fig. I-2, using the method of deriving the conversion signal given in Fig. I-3b.

compared with that of the arriving signal, thereby minimizing the effect of noise in the final detecting element.

If these operations can be carried out, then, at all points across the aperture ahead of the final lens that focuses energy onto the difference-frequency detector, the difference-frequency component has almost constant phase, and an amplitude distribution in which those points receiving most strongly are accentuated, and those receiving most weakly are suppressed.

#### References

1. R. S. Kennedy and E. V. Hoversten, "On the Atmosphere as an Optical Communication Channel," IEEE Trans. on Information Theory, Vol. IT-14, No. 5, pp. 716-725, September 1968.
2. R. Price and P. E. Green, "A Communication Technique for Multipath Channel," Proc. IRE 46, 550-570 (1958).

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION  
WASHINGTON, D. C. 20546  
OFFICIAL BUSINESS

FIRST CLASS MAIL



POSTAGE AND FEES PAID  
NATIONAL AERONAUTICS AND  
SPACE ADMINISTRATION

02U 001 32 51 3DS 70043 00903  
AIR FORCE WEAPONS LABORATORY /WL0L/  
KIRTLAND AFB, NEW MEXICO 87117

ATT E. LOU BOWMAN, CHIEF, TECH. LIBRARY

POSTMASTER: If Undeliverable (Section 158  
Postal Manual) Do Not Return

*"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."*

— NATIONAL AERONAUTICS AND SPACE ACT OF 1958

## NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

**TECHNICAL REPORTS:** Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

**TECHNICAL NOTES:** Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

**TECHNICAL MEMORANDUMS:** Information receiving limited distribution because of preliminary data, security classification, or other reasons.

**CONTRACTOR REPORTS:** Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

**TECHNICAL TRANSLATIONS:** Information published in a foreign language considered to merit NASA distribution in English.

**SPECIAL PUBLICATIONS:** Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

**TECHNOLOGY UTILIZATION PUBLICATIONS:** Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Notes, and Technology Surveys.

*Details on the availability of these publications may be obtained from:*

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION  
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION  
Washington, D.C. 20546