Quantum-Mechanical Communication Theory

C. W. Helstrom,* Jane W. S. Liu,† J. P. Gordon‡

Abstract – We are concerned with the problem of finding the structure and performance of the receiver that yields the best performance in the reception of signals that are described quantum-mechanically. The principles of statistical detection and estimation theory are discussed, with the laws of quantum mechanics taken into account. Several specific communication systems of practical interest are studied as examples of applying these principles. Basic concepts in quantum mechanics that are needed in these discussions are briefly reviewed.

*Department of Applied Electrophysics, University of California, San Diego, La Jolla, California. Research supported by NASA Grant NGL 05-009-079.

†Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts. Research supported by NASA Grant NGL 22-009-013.

‡Bell Telephone Laboratories, Inc., Holmdel, New Jersey.

I. INTRODUCTION

The performance and structure of a receiver in a communication system depend on the form of the signals used to transmit messages and the nature of the random noise that accompanies the signals. For a given performance measure, the best structure of the receiver can be determined by the principles of detection and estimation theory, which views the receiver as an instrument for testing certain hypotheses about its input and applies the methods of statistical decision theory [1-4].

Although real receivers are perturbed by a variety of noises, whose characteristics differ from one application to another, thermal noise is always present. By thermal noise, we mean the type of noise that arises from the chaotic thermal agitation of the atoms and molecules composing the receiver and its surroundings [5-6]. The upper limit to the receiver performance in the reception of a given set of signals can be ascertained by calling upon detection or estimation theory to determine the best receiver for receiving these signals when they are accompanied only by thermal noise.

In most treatments of ideal reception of signals at microwave frequencies, it is assumed that the electromagnetic fields of signal and noise, the receiver, and the interaction between them behave in accordance with the classical laws of electromagnetism. The noise generated by an ideal receiver itself can be accounted for in the thermal noise accompanying the signals at its input. The ideal receiver can examine its input in every detail in order to extract all information relevant to the optimum reception of the given signals without introducing more uncertainty about them. Usually, the input to the receiver is described by the waveforms of the electric field over the receiving aperture.

When the signals to be detected are composed of optical rather than microwave frequencies, the input fields and their interaction with matter can be described accurately only by the laws of quantum mechanics. The postulate that the ideal receiver can make use of every detail of its input without introducing further uncertainty must be scrutinized. By virtue of the uncertainty principle of quantum mechanics, for instance, the amplitude and phase of the input field cannot be determined simultaneously with arbitrary accuracy. Indeed, an electric field having both precise amplitude and phase cannot in principle be generated. Whereas the limitations imposed by the laws of quantum mechanics negligibly affect the reliability of a communication system at microwave frequencies, they are often more influential than the thermal noise at optical frequencies.

In early studies of the quantum-mechanical aspects of communication systems, emphasis was placed on finding upper bounds to channel capacity and on evaluating performance of systems incorporating specific receivers. Channel capacities were derived for communication systems using known receivers, such as linear amplifiers, heterodyne and homodyne receivers, and photon counters [7-13]. Quantum-mechanical limitations on the accuracy of measurements made by a phase-sensitive receiver are taken into account by introducing at its input a frequency-dependent noise statistically similar to thermal noise. Quantum limitations on the detection of known and random signals in thermal noise were found when the receiver measures the strength of the electric field as in a heterodyne receiver [14].

We are concerned with the problem of finding the structure and performance of the receiver that yields the best performance in the reception of signals that are described quantum-mechanically. In Sections IV and V, the principles of statistical detection and estimation theory, taking into account the laws of quantum mechanics [15-18], are discussed. Several specific communication systems of practical interest are studied as examples of applying these principles. Basic concepts in quantum mechanics needed in these discussions are briefly reviewed in Section II.

II. QUANTUM MECHANICS

A thorough introduction to quantum mechanics can hardly be fitted into the compass of this paper; at most we can present the basic rules and concepts. The reasons behind them and the techniques of applying them to physical problems have been discussed in textbooks [19,20]. We shall content ourselves with asserting that from these principles a broad and accurate understanding of the physical world in general, and of the properties of matter and radiation in particular, have been achieved.

To great accuracy, the behavior of all physical communication systems is governed by the laws of quantum mechanics. Whereas the classical laws predict the behavior of communication systems operating at microwave frequencies or below with reasonable accuracy, the radiation fields and the receivers in communication systems at optical frequencies can only be adequately described quantum-mechanically. We shall call a system classical or quantum-mechanical depending on whether a sufficiently accurate description of its behavior requires classical or quantum-mechanical laws.

State Vectors and Operators

As in the case of a classical system, the condition of a quantum-mechanical system at any instant of time is completely specified by its state. Mathematically, the state of a quantum-mechanical system is described by a state vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$ over the field of complex numbers [21]. Any state vector $|\psi\rangle$ can be expressed in terms of a linear combination of vectors $|\phi_n\rangle$ in a basis – coordinate system – $\{|\phi_n\rangle\}$ in the Hilbert space $\mathcal{H}$ [22]

$$|\psi\rangle = \sum_{n=1}^{\infty} a_n |\phi_n\rangle.$$

Hence without loss of generality, $|\psi\rangle$ can be thought of as a column vector having an infinite number of components that are the complex numbers a_n. State vectors can be combined linearly to form a new vector that also represents a possible state of the system. Associated with any $|\psi\rangle$ is a Hermitian conjugate $\langle\psi|$, which can be regarded as a row vector whose components are complex conjugates of those in $|\psi\rangle$:

The scalar product of two state vectors $|\phi\rangle$ and $|\psi\rangle$ is a complex number, written $\langle\phi|\psi\rangle$; in terms of the components $\{a_n\}$ and $\{b_n\}$ of these vectors,

$$\langle\phi|\psi\rangle = \sum_n a_n^* b_n. \tag{1}$$

Following Dirac, we call $|\psi\rangle$ a ket, and $\langle\phi|$ a bra because together they form a bracket $\langle\phi|\psi\rangle$. Although the components a_n and b_n of the kets $|\phi\rangle$ and $|\psi\rangle$ depend on the basis to which they are referred, the value of the scalar product $\langle\phi|\psi\rangle$ is independent thereof. The squared length of the ket $|\psi\rangle$ is $\langle\psi|\psi\rangle$; when it is finite, the ket is normalized to have unit length: $\langle\psi|\psi\rangle = 1$. Not all state vectors in quantum mechanics can be assigned a finite length. Two kets differing only by a phase factor $e^{i\theta}$ that is common to all components describe physically identical states.

The kets $|\psi\rangle$ are transformed by linear operators. A linear operator Ξ can be expressed in term of the kets $|\phi_n\rangle$ in a basis $\{|\phi_n\rangle\}$

$$\Xi = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\phi_m\rangle\langle\phi_m|\Xi|\phi_n\rangle\langle\phi_n|.$$

When represented in terms of the basis $\{|\phi_n\rangle\}$, the operator is associated with a square matrix, albeit usually infinite in extent, whose elements are

$$\Xi_{mn} = \langle\phi_m|\Xi|\phi_n\rangle.$$

We can, therefore, regard a linear operator as a square matrix. An operator Ξ is said to be Hermitian when its associated matrix equals the transpose conjugate matrix Ξ^+. Ξ^+ is called the Hermitian adjoint of Ξ.

Suppose that a Hermitian operator Ξ has a discrete set $(\xi_1, \xi_2, \ldots \xi_n, \ldots)$ of eigenvalues. The associated eigenvectors $|\xi_1\rangle$, $|\xi_2\rangle, \ldots |\xi_n\rangle, \ldots$ are defined by the eigenvalue equation

$$\Xi|\xi_n\rangle = \xi_n|\xi_n\rangle. \tag{2}$$

Because Ξ is Hermitian, the eigenvalues ξ_n are real. The eigenvectors are orthogonal. Being normalized to unit length, they have the scalar products

$$\langle\xi_n|\xi_m\rangle = \delta_{nm}, \tag{3}$$

where δ_{nm} is the Kronecker delta symbol. Our interest will be restricted to those Hermitian operators whose eigenvectors form a complete set,

$$\sum_{n=1}^{\infty} |\xi_n\rangle\langle\xi_n| = \underset{\sim}{1}, \tag{4}$$

where $\underset{\sim}{1}$ is the identity operator: $\underset{\sim}{1}|\psi\rangle = |\psi\rangle$, for all $|\psi\rangle$. Having properties (3) and (4), the eigenvectors $|\xi_n\rangle$ form a basis in term of which any ket $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \sum_{n=1}^{\infty} c_n|\xi_n\rangle, \qquad c_n = \langle\xi_n|\psi\rangle. \tag{5}$$

If, on the other hand, a linear operator Z has a continuum of eigenvalues ζ,

$$Z|\zeta\rangle = \zeta|\zeta\rangle \tag{6}$$

the eigenvectors $|\zeta\rangle$ have infinite length and are so normalized that

$$\langle\zeta'|\zeta''\rangle = \delta(\zeta'-\zeta''), \tag{7}$$

where $\delta(\zeta'-\zeta'')$ is the Dirac delta function. Their completeness relation now is

$$\int_{-\infty}^{\infty} |\zeta\rangle\langle\zeta| \, d\zeta = \underset{\sim}{1}, \tag{8}$$

by virtue of which any ket $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \int_{-\infty}^{\infty} \gamma(\zeta)|\zeta\rangle \, d\zeta, \qquad \gamma(\zeta) = \langle\zeta|\psi\rangle. \tag{9}$$

Observables and Quantum-Mechanical Measurement

Each measurable physical quantity, or observable, of the system, such as position, momentum, or angular momentum of a particle, is associated with a Hermitian operator that has a complete set of eigenvectors. The eigenvalues of such a Hermitian operator may form a discrete set, a continuous set, or a combination of both [23]. Without loss of generality, we shall discuss observables and their operators as though their spectra were discrete. Modifications to cover those with a continuous spectrum of eigenvalues, or a combination of discrete and continuous spectra, will involve changing Kronecker deltas to Dirac deltas, and sums to integrations.

Quantum mechanics postulates that an exact measurement of an observable whose operator is Ξ always yields as an outcome one or another of the eigenvalues ξ_n of Ξ. Immediately after a measurement of Ξ that yields the eigenvalue ξ_k, the measured system is in the corresponding eigenstate $|\xi_k\rangle$, and repeating the measurement immediately afterwards on the same system would yield the same value ξ_k [24].

Furthermore, if before measurement the system is in state $|\psi\rangle$, the probability that the measurement will yield the value ξ_k is given by

$$\Pr(\xi_k) = |\langle \xi_k | \psi \rangle|^2 = |c_k|^2 \tag{10}$$

where c_k is the coefficient in the expansion of $|\psi\rangle$ as in (5). By the closure relation (4), these probabilities sum to 1,

$$\sum_k \Pr(\xi_k) = \sum_k \langle \psi | \xi_k \rangle \langle \xi_k | \psi \rangle = \langle \psi | \psi \rangle = 1.$$

If the observable, as Z, has a continuous spectrum, $|\langle \zeta | \psi \rangle|^2 \, d\zeta$ is the probability that the outcome of a measurement lies between ζ and $\zeta + d\zeta$. Hence $|\langle \zeta | \psi \rangle|^2$ is the probability density function of the outcome of the measurement. From (2) and (4), the expected value of the outcome of a measurement of Ξ is

$$\underset{\sim}{E}[\Xi] = \sum_k \xi_k \Pr(\xi_k) = \langle \psi | \Xi | \psi \rangle, \tag{11}$$

when the system is in state $|\psi\rangle$. This expression for the expected value is independent of the coordinate system used to describe the ket $|\psi\rangle$.

The outcome of a measurement of the observable Ξ is the answer to the question, "What is the value of Ξ?". Instead, suppose that the weaker question is asked, "Does the value of Ξ lie between a and b?", and the actual value within the range (a, b) is of no concern. While this question can also be answered by measuring Ξ itself, it is sufficient to measure the observable represented by the operator

$$\Pi_{ab} = \sum_{R(a,b)} |\xi_n\rangle \langle \xi_n|, \tag{12}$$

where R(a, b) is the set of eigenvalues ξ_n of Ξ lying within (a, b). A complete set of eigenvectors of Π_{ab} is the eigenvectors $|\xi_n\rangle$ of Ξ. Π_{ab} has only the two eigenvalues, 1 and 0, however. The outcome of the measurement of Π_{ab} will be 1 if the value of Ξ lies in R(a, b) and 0 otherwise. If the state of the system before measurement is $|\psi\rangle$, the expected value of the outcome is

$$\underset{\sim}{E}[\Pi_{ab}] = \langle\psi|\Pi_{ab}|\psi\rangle = \sum_{R(a,b)} |\langle\xi_n|\psi\rangle|^2 = \Pr\{\xi \in R(a,b)\}, \tag{13}$$

which is the probability that the value ξ of the observable Ξ lies between a and b.

The operator Π_{ab} in (12) is a projection operator. It obeys the defining equation

$$\Pi_{ab}^2 = \Pi_{ab} \tag{14}$$

for projection operators, as can be seen by using (12) and the orthonormality of the kets $|\xi_n\rangle$, (3). Since $\Pi_{ab}(\Pi_{ab} - \underset{\sim}{1}) = 0$, the only eigenvalues of Π_{ab} are 0 or 1, as we have already observed. The operator Π_{ab} projects the ket $|\psi\rangle$ onto the linear subspace spanned by the kets $|\xi_n\rangle$ for which $\xi_n \in R(a, b)$. The statement, "The value of the observable Ξ lies between a and b," is a proposition that is either true or false, and such propositions correspond in the logic of quantum mechanics to projection operators like Π_{ab}. The decisions among hypotheses treated in detection theory will be expressed in this form.

If two observables, say Ξ and Υ, are to be measured exactly and simultaneously on the same system, it must be left after the measurement in a state $|\xi_n, \upsilon_m\rangle$ that is an eigenstate of both operators Ξ and Υ with eigenvalues ξ_n, υ_m, respectively [25],

$$\begin{aligned} \Xi|\xi_n, \upsilon_m\rangle &= \xi_n|\xi_n, \upsilon_m\rangle \\ \Upsilon|\xi_n, \upsilon_m\rangle &= \upsilon_m|\xi_n, \upsilon_m\rangle, \end{aligned} \tag{15}$$

and this must hold true for all possible outcomes ξ_n, υ_m of the measurement. It follows from (15) and the completeness of the states $|\xi_n, \upsilon_m\rangle$ that a necessary and sufficient condition is that the operators Ξ and Υ commute,

$$\Xi\Upsilon = \Upsilon\Xi. \tag{16}$$

Two observables are said to be compatible when their corresponding operators commute. As an example, the three Cartesian coordinates x, y, z of a particle are compatible, for the three associated operators X, Y, Z commute.

Uncertainty Principle

When the two operators Ξ and Υ do not commute, the corresponding observables cannot be measured simultaneously on the same system with complete precision. This is a crude statement of the Heisenberg uncertainty principle. To express this principle more precisely, we suppose that the operators Ξ and Υ satisfy the commutation relation,

$$[\Xi, \Upsilon] = \Xi\Upsilon - \Upsilon\Xi = iZ,$$

where Z is either a constant times the identity operator $\underset{\sim}{1}$ or another operator. Let

$$\sigma_{\Xi}^2 = \underset{\sim}{E}[\Xi^2] - \underset{\sim}{E}[\Xi]^2 \tag{17}$$

denote the mean-square deviation of the outcomes of a measurement of Ξ. Similarly, we define

$$\sigma_{\Upsilon}^2 = \underset{\sim}{E}[\Upsilon^2] - \underset{\sim}{E}[\Upsilon]^2. \tag{18}$$

When the system is in the state $|\psi\rangle$ before the measurement,

$$\sigma_{\Xi}^2 = \langle\psi|\Xi^2|\psi\rangle - (\langle\psi|\Xi|\psi\rangle)^2$$

and

$$\sigma_{\Upsilon}^2 = \langle\psi|\Upsilon^2|\psi\rangle - (\langle\psi|\Upsilon|\psi\rangle)^2$$

By using the Schwarz inequality, it can be shown [26] that the product $\sigma_{\Xi}\sigma_{\Upsilon}$ of the standard deviations satisfies the inequality

$$\sigma_{\Xi}\sigma_{\Upsilon} \geq |\langle\psi|Z|\psi\rangle|/2 \tag{19}$$

for all state vectors $|\psi\rangle$. Equation (19) is called the Heisenberg uncertainty relation.

This principle can be illustrated by an example. The two operators P and Q satisfy the commutation relation

$$[Q, P] = QP - PQ = i\hbar,$$

where $\hbar = h/2\pi$ is Planck's constant. The corresponding observables are the coordinate q and momentum p, respectively, of a particle with one degree of freedom. The same commutation relation holds for the operators corresponding to the charge and current in a lossless LC circuit. From (19),

$$\sigma_Q\sigma_P \geq \hbar/2. \tag{20}$$

In order to interpret this relation, we think of a large ensemble of N independent systems, all of them in the same state $|\psi\rangle$. On some systems Q is measured, on others P. The outcomes will be random variables, differing from one system to another, but when $N \gg 1$, the average values will be near their mean values and the mean-square deviations near σ_Q^2 and σ_P^2, respectively. Just what average values and mean-square deviations σ_Q^2, σ_P^2 are obtained depends on the state $|\psi\rangle$ in which the systems were originally prepared. The uncertainty principle (20) asserts that for no state $|\psi\rangle$ can the product of the standard deviations $\sigma_Q\sigma_P$ be less than $\frac{1}{2}\hbar$.

Density Operators

Given an ensemble of independent systems, all in the same state $|\psi\rangle$, measurement of an observable such as Ξ on each will in general produce a random collection of results, the probability of obtaining the value ξ_n being $|\langle \xi_n|\psi\rangle|^2$, n = 1, 2, ... This randomness is strictly a quantum phenomenon. The kind of randomness met in classical physics must also somehow be incorporated into the framework of quantum mechanics, so that we can treat problems involving noise or random signals. The means of doing so is provided by the density operator ρ [27,28].

Let a large number of systems of the same kind be prepared, each in one of a set of orthonormal states $|\phi_n\rangle$, and let the fraction of systems in state $|\phi_n\rangle$ be P_n, n = 1, 2, ..., with

$$\langle \phi_m|\phi_n\rangle = \delta_{mn} \tag{21}$$

$$\sum_{n=1}^{\infty} P_n = 1. \tag{22}$$

P_n is the prior probability of the state $|\phi_n\rangle$. If now we measure the observable Ξ, the probability of obtaining the value ξ_k will be

$$\Pr\{\xi_k\} = \sum_{n=1}^{\infty} P_n |\langle \xi_k|\phi_n\rangle|^2 = \langle \xi_k|\rho|\xi_k\rangle. \tag{23}$$

In this expression, the operator ρ defined by

$$\rho = \sum_{n=1}^{\infty} P_n |\phi_n\rangle\langle\phi_n| \tag{24}$$

is called the <u>density operator</u>. Since ρ is a linear operator, it can also be thought of as a square matrix. The expected value of the outcome of our measurement of Ξ is

$$\underset{\sim}{E}(\Xi) = \sum_{k=1}^{\infty} \xi_k \Pr\{\xi_k\} = \sum_{k=1}^{\infty} \xi_k \langle \xi_k|\rho|\xi_k\rangle = \sum_{k=1}^{\infty} \langle \xi_k|\rho\Xi|\xi_k\rangle$$

$$= \mathrm{Tr}(\rho\Xi), \tag{25}$$

where Tr stands for the trace of a matrix, the sum of its diagonal elements.

Clearly, the density operator ρ is Hermitian. It has a complete set of orthonormal eigenvectors $|\phi_n\rangle$ corresponding to non-negative eigenvalues P_n and $\mathrm{Tr}\, \rho = 1$. Moreover, any Hermitian operator with non-negative eigenvalues and trace 1 may be considered as a density operator that describes an ensemble of quantum-mechanical systems.

A density operator ρ also has the property $\mathrm{Tr}\, \rho^2 \leq 1$, with equality if and only if one of the prior probabilities P_n equals 1 and all the rest 0. When this is so, the density operator is a projection operator, $\rho = |\phi_n\rangle\langle\phi_n|$, and the ensemble is a collection of systems all in the same state $|\phi_n\rangle$. We say when ρ is a projection operator that it represents a system in a pure state; otherwise, with $\mathrm{Tr}\, \rho^2 < 1$, it represents a mixed state.

Time Dependence

Thus far, we have not shown how a quantum-mechanical system behaves dynamically. To discuss the manner in which the state of a system changes with time, let us denote the state vector at time by $|\psi(t)\rangle$. It is clear from the discussions on quantum-mechanical measurements that the state vector changes irreversibly in an unpredictable way when the system interacts with a measuring device. But when a closed quantum-mechanical system is not perturbed by any measurement, its state vector $|\psi(t)\rangle$ at time t obeys a linear differential equation of the first order in time [29]

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H|\psi(t)\rangle, \tag{26}$$

where H is an operator called the Hamiltonian of the system. We shall be concerned solely with conservative systems. The Hamiltonian of a conservative system does not contain time explicitly; (26) can be solved to obtain

$$|\psi(t)\rangle = \exp\left[-i\frac{H}{\hbar}(t-t_0)\right]|\psi(t_0)\rangle, \tag{27}$$

which relates the state vector $|\psi(t)\rangle$ at time t to that at an earlier time t_0.

The Hamiltonian H is the operator corresponding to the energy of the system. For many systems H can be obtained from the classical expression for the energy in terms of coordinates and momenta, these being simply replaced by their quantum-mechanical operators and properly symmetrized. Thus, for a simple harmonic oscillator of mass m and frequency ω, the Hamiltonian operator is

$$H = (2m)^{-1}(P^2+\omega^2Q^2) \tag{28}$$

in terms of the operators P and Q for its momentum and its coordinate.

The operator $\exp\left[-i\frac{H}{\hbar}(t-t_0)\right]$ executes a unitary transformation on the state vectors $|\psi\rangle$, as in (27). Under this transformation the lengths of the kets $|\psi\rangle$ and the angles between them in the Hilbert space do not change. This representation of the time dependence of a system as a rigid rotation of its state vectors is called the "Schrödinger picture." In the Schrödinger picture, operators not depending explicitly on the time are taken as constants.

We shall use a different, but physically equivalent, representation of the time dependence of quantum-mechanical systems, which is called the "Heisenberg picture" [30]. The state vector in the Heisenberg picture is independent of time. At time t, an observable $\Xi(t)$ is related to the operator Ξ in the Schrödinger picture by

$$\Xi(t) = \exp\left[i\frac{H}{\hbar}(t-t_0)\right]\Xi\exp\left[-i\frac{H}{\hbar}(t-t_0)\right] \tag{29}$$

From this transformation and (26), the equation of motion for the observable Ξ in the Heisenberg picture is

$$i\hbar \frac{d\Xi(t)}{dt} = [\Xi(t), H] + i\hbar \frac{\partial \Xi(t)}{\partial t} \tag{30}$$

The Hamiltonian H equals the Hamiltonian in the Schrödinger picture, since the system is conservative.

If a system is in a statistical mixture of states represented by a density operator ρ, ρ is independent of time in the Heisenberg picture. The expectation value of a measurement of an observable $\Xi(t)$ at time t will be

$$E_t[\Xi] = \mathrm{Tr}\,[\rho \Xi(t)].$$

It is easy to see that this expectation value equals $\mathrm{Tr}\,[\Xi\, \rho(t)]$, where $\rho(t) = \exp\left[-\frac{i}{\hbar} H(t-t_o)\right] \rho \exp\left[\frac{i}{\hbar} H(t-t_o)\right]$ and Ξ are the operators at time t in the Schrödinger picture.

Ideal Measurement of Incompatible Observables

The class of exact measurements of compatible observables discussed above does not include those measurements yielding approximate values of several incompatible observables. As an example of devices that make such approximate measurements on the received field, we mention a high-gain laser amplifier followed by a classical receiver. The field at the output of such an amplifier can be treated as a classical field with precisely measurable amplitude and phase. Therefore, we may consider that the amplifier performs simultaneous approximate measurements of the amplitude and phase of its input field [31]. The additive Gaussian noise injected by the amplifier accounts for the inevitable error in the measurement imposed by the uncertainty principle.

We shall briefly consider a definition [32] for ideal measurements to include such approximate measurements of incompatible observables. Let x_n denote the set of numbers that are the observed values of the measured observables. By an ideal measurement, we mean one in which each possible outcome x_n is associated with a state vector $|\underset{\sim}{x}_n\rangle$ such that the probability for obtaining $\underset{\sim}{x}_n$ is

$$P(\underset{\sim}{x}_n) = w|\langle x_n|\psi\rangle|^2 \tag{31}$$

when the system is in the state $|\psi\rangle$ prior to the measurement, and w is a normalization constant. Moreover, the state of the system after the measurement depends only on the measurement result, and not at all on its initial state before the measurement. Thus subsequent measurements cannot yield additional information about that initial state.

The normalized state vector $|\underset{\sim}{x}_n\rangle$ is called a measurement state vector. Since some result must be obtained from any measurement, the set of measurement state vectors must satisfy a completeness relation of the form

$$\sum_n w|\underset{\sim}{x}_n\rangle\langle\underset{\sim}{x}_n| = \underset{\sim}{1}. \tag{32}$$

Any ideal measurement, therefore, is characterized by a complete set of measurement state vectors.

Exact measurements of compatible observables are ideal by the above-mentioned criterion. They are characterized by complete sets of orthonormal measurement state vectors that are the simultaneous eigenvectors of the measured observables. In such cases the relation (31) is satisfied with the normalization constant w equal to unity, and (32) is simply the completeness relation in (4). The measurement of field amplitude and phase as made by an ideal high-gain amplifier is also ideal. In this case, the measurement state vectors are the coherent state vectors defined in (50), and the appropriate completeness relation is given by (52). These measurement state vectors are not orthogonal.

Sets of nonorthogonal vectors that satisfy a completeness relation such as (32) are called overcomplete sets. Ideal measurements yielding approximate values of several incompatible observables can be characterized by overcomplete sets of measurement state vectors. It is probably true that for every such measurement, in principle, there is an equivalent exact measurement characterized by an orthonormal set of measurement state vectors. Many conveniently realizable measuring processes correspond, however, to overcomplete sets of measurement state vectors, while realization of the equivalent exact measurements might prove difficult.

A fairly general way of implementing ideal measurement of incompatible observables is to combine the system to be measured with an auxiliary system whose initial state is known. The two systems may be allowed to interact for a length of time. An exact measurement of a complete set of compatible observables for the expanded system is then made. As an example of this prescription, consider the observables Q and P for a simple harmonic oscillator. We may combine this system with an auxiliary system comprising a similar but independent oscillator in its ground state, whose corresponding observables are Q' and P'. The observables Q-Q' and P+P' of the expanded system are compatible, and their simultaneous exact measurement yields approximate values of Q and P. Further analysis [33] shows that this prescription indeed yields an ideal measurement of Q and P.

III. QUANTUM-MECHANICAL DESCRIPTION OF COMMUNICATION SYSTEMS

A typical communication system is shown in Fig. 1. In every signaling interval of duration T, the input m of the system is either one of M messages generated by a digital data source or a set of parameters carrying analog data. A signal field, whose characteristics depend on the input message, is generated by the transmitter and is sent through the channel to the receiver. During each signaling interval, the receiver makes an estimate $\hat{m}$ of the transmitted message. Our objective is to design the receiver so that $\hat{m}$ minimizes a given cost function used to measure the fidelity of this estimate. Examples of commonly used cost functions are the probability of error, $\Pr[\hat{m}\neq m]$, for digital messages, and the mean-square error, $\underset{\sim}{E}[|\hat{m}-m|]^2$, for analog data.

For simplicity, input messages in different signaling intervals are assumed to be statistically independent. Furthermore, our attention will be restricted to systems in which the channel is memoryless and no coding schemes are employed. In these systems, the receiver makes independent estimates of input messages in successive signaling intervals on the basis of the electromagnetic field observed during each of these intervals. Therefore, we need to be concerned only with the problem of making an optimum estimate of a single message. The signal field representing such a message is a time-limited one with nonzero instantaneous power only in a time interval of duration T. Without loss of generality, we let this interval be $(0, T)$.

The receiver admits the incident field through an area normal to the direction of the transmitter. In an ordinary receiver, this area corresponds to the effective area of the antenna, which in practice must be limited to a finite size. This area will be called the receiving aperture.

Since at any time the instantaneous power associated with a time-limited signal is nonzero only in a finite region in space, and the noise fields at different points in space are statistically independent, the receiver can be idealized as a large lossless box or cavity with perfectly conducting walls. During the time interval (0, T), the incident field is admitted into the cavity, initially empty, through the receiving aperture. At the end of this interval, the aperture is closed, and measurements are made by the receiver on the field inside the cavity, which is called the received field.

The received field can be represented as a superposition of normal modes of the cavity. Each mode behaves like a harmonic oscillator with frequency ω_k; the frequencies ω_k depend on the shape of the cavity. To be more specific, the classical waveform, $\underset{\sim}{\mathscr{E}}(\underset{\sim}{r}, t)$, of the received electric field can be expanded in terms of standing-wave, normal-mode functions $\underset{\sim}{u}_k(\underset{\sim}{r})$ of an appropriately chosen cavity of volume V

$$\underset{\sim}{\mathscr{E}}(\underset{\sim}{r}, t) = -\epsilon_o^{-1/2} \sum_k p_k(t)\, \underset{\sim}{u}_k(\underset{\sim}{r}), \tag{33}$$

where ϵ_o, the dielectric constant, is used here for normalization. As a result of boundary conditions at the walls of the cavity, the functions $\underset{\sim}{u}_k(\underset{\sim}{r})$ are orthonormal,

$$\int_{\text{cavity}} \underset{\sim}{u}_k(\underset{\sim}{r})\, \underset{\sim}{u}_n(\underset{\sim}{r})\, d^3\underset{\sim}{r} = \delta_{kn}. \tag{34}$$

Inside the cavity, $\underset{\sim}{u}_k(\underset{\sim}{r})$ is a solution of the Helmholtz equation

$$\nabla^2 \underset{\sim}{u}_k(\underset{\sim}{r}) + \left(\omega_k^2/c^2\right) \underset{\sim}{u}_k(\underset{\sim}{r}) = 0 \tag{35}$$

for all k, where c is the velocity of light in vacuum. The oscillation frequencies ω_k of the normal modes are determined by (34) and (35).

As a consequence of Maxwell's field equations, the functions $q_k(t)$ defined by

$$p_k(t) = dq_k(t)/dt, \tag{36}$$

in association with the mode amplitudes $p_k(t)$ in (33), satisfy the equation of motion

$$\frac{d^2}{dt^2} q_k(t) + \omega_k^2 q_k(t) = 0. \tag{37}$$

Therefore, we may associate each mode of the field with a harmonic oscillator of frequency ω_k. Furthermore, it can be shown from Maxwell's field equations and from (33)-(35) that the total energy H contained in the received field is the sum of the energies of the uncoupled harmonic oscillators [34]

$$H = \sum_k \left(p_k^2 + \omega_k^2 q_k^2\right)/2. \tag{38}$$

It is often preferable to represent the received field in terms of plane traveling waves rather than standing waves. A particular set of mode functions suitable for our purpose is the set of plane traveling-wave mode functions of a cubical cavity of volume V. That is,

$$\underset{\sim}{u}_k(\underset{\sim}{r}) = V^{-1/2} \underset{\sim}{e}_k \exp(i\underset{\sim}{k} \cdot \underset{\sim}{r}), \tag{39}$$

where $\underset{\sim}{e}_k$ is a unit polarization vector perpendicular to the propagation vector $\underset{\sim}{k}$, and $|\underset{\sim}{k}|^2 = \omega_k^2/c^2$, for all k. The complex amplitude $\alpha_k(t)$ of each plane traveling-wave mode is related to the real variables $p_k(t)$ and $q_k(t)$ by

$$\alpha_k(t) = (2\hbar\omega_k)^{1/2} [\omega_k q_k(t) + ip_k(t)]. \tag{40}$$

Hence the equations of motion for $\alpha_k(t)$ are

$$\frac{d}{dt} \alpha_k(t) = -i\omega_k \alpha_k(t),$$

which have solutions

$$\alpha_k(t) = \alpha_k \exp[-i\omega_k t]. \tag{41}$$

In terms of these complex amplitudes, the energy of each normal mode is

$$H_k = \hbar\omega_k |\alpha_k|^2. \tag{42}$$

Quantization of the Radiation Field

The quantum theory of radiation [35] also treats each mode of the field as a harmonic oscillator. The "coordinates" $q_k(t)$ and "momenta" $p_k(t)$ are replaced by their corresponding quantum-mechanical operators $Q_k(t)$ and $P_k(t)$, which obey the commutation rule

$$[Q_k(t), P_n(t)] = Q_k(t)P_n(t) - P_n(t)Q_k(t) = i\hbar\,\delta_{kn} \tag{43}$$

for all k and n. The complex amplitudes α_k in (41) are replaced by operators a_k that are related to the operators $Q_k(t)$ and $P_k(t)$ by

$$a_k \exp[-i\omega_k t] = (2\hbar\omega_k)^{1/2}\,[\omega_k Q_k(t) + iP_k(t)]. \tag{44}$$

It follows from the commutation rules (43) that the commutation relations between the operators a_k and their Hermitian adjoints a_k^+ are

$$[a_k, a_n^+] = \delta_{kn}\underset{\sim}{1} \qquad \text{for all n and k.}$$

In terms of these operators, the electric field operator is

$$\underset{\sim}{E}(\underset{\sim}{r}, t) = i \sum_k (\hbar\omega_k/2\epsilon_o V)^{1/2}\,\underset{\sim}{e}_k\{a_k \exp[-i(\omega_k t - \underset{\sim}{k}\cdot r)] - a_k^+ \exp[+i(\omega_k t - \underset{\sim}{k}\cdot \underset{\sim}{r})]\}. \tag{45}$$

The Hamiltonian of the field becomes

$$H = \sum_k H_k,$$

where

$$H_k = \hbar\omega_k\left(a_k a_k^+ + a_k^+ a_k\right) = \hbar\omega_k\left(a_k^+ a_k + \frac{1}{2}\right). \tag{46}$$

For reasons that will appear immediately, the operator $a_k^+ a_k$ is called the number operator of the k^{th} mode. In texts on quantum mechanics, it is shown that the eigenvalues of the operators $a_k^+ a_k$ are the positive integers and zero [36]. We denote the eigenvectors by the corresponding eigenvalues

$$a_k^+ a_k |n_k\rangle = n_k |n_k\rangle. \tag{47}$$

Hence by (46), the eigenvalues E_k of the energy in the k^{th} mode are

$$E_k = \hbar\omega_k\left(n_k + \frac{1}{2}\right).$$

When the mode is in the state $|n_k\rangle$, it is customary to say that it contains n_k photons, each of which carries an energy of $\hbar\omega_k$. The state $|0\rangle$, the ground state, possesses a zero-point fluctuation energy $\frac{1}{2}\hbar\omega_k$.

When the operator a_k acts to the right on an eigenvector $|n_k\rangle$ of the number operator $a_k^+ a_k$, it converts the eigenvector $|n_k\rangle$ to $|n_k - 1\rangle$,

$$a_k |n_k\rangle = n_k^{1/2} |n_k - 1\rangle, \tag{48}$$

thereby reducing the number of photons of the mode by 1. For this reason, a_k is called the annihilation operator of the k^{th} mode. Its Hermitian adjoint a_k^+ raises the number of photons by 1 and is called the creation operator,

$$a_k^+ |n_k\rangle = (n_k + 1)^{1/2} |n_k + 1\rangle. \tag{49}$$

When the k^{th} mode of the field is in a state described by a state vector $|\alpha_k\rangle$ that is a right eigenvector of the annihilation operator a_k,

$$a_k |\alpha_k\rangle = \alpha_k |\alpha_k\rangle, \tag{50}$$

where $\alpha_k = \alpha_{kx} + i\alpha_{ky}$ is a complex eigenvalue, the mode is said to be in a coherent state [37]. Alternatively, we say that the mode contains a coherent signal. The coherent state vector $|\alpha_k\rangle$ can be expressed in terms of the eigenvectors $|n_k\rangle$ of the number operator $a_k^+ a_k$ [38]

$$|\alpha_k\rangle = \exp[-|\alpha_k|^2/2] \sum_{n_k=1}^{\infty} (n_k!)^{-1/2} \alpha_k^{n_k} |n_k\rangle, \tag{51}$$

and are normalized so that $\langle \alpha_k | \alpha_k \rangle = 1$. Moreover, they are complete, in the sense that

$$\int |\alpha_k\rangle \langle \alpha_k| d^2\alpha_k/\pi = \underline{1}, \tag{52}$$

where $d^2\alpha_k = d\alpha_{kx}\, d\alpha_{ky}$ is the element of integration in the complex plane, over the entirety of which the integration is performed. The coherent state vectors $|\alpha_k\rangle$ and $|\beta_k\rangle$ are not orthogonal, however. Their inner product is

$$\langle \alpha_k | \beta_k \rangle = \exp[\alpha_k^* \beta_k - |\alpha_k|^2/2 - |\beta_k|^2/2]. \tag{53}$$

The entire field is in a coherent state $|\{\alpha_k\}\rangle$ when all of its normal modes are in coherent states. The state vector $|\{\alpha_k\}\rangle$ is simultaneously a right eigenvector of all of the annihilation operators a_k.

$$a_k |\{\alpha_k\}\rangle = \alpha_k |\{\alpha_k\}\rangle. \tag{54}$$

It can be taken to be the direct product of the state vectors for the individual modes

$$|\{\alpha_k\}\rangle = |\alpha_1, \alpha_2, \ldots \alpha_k, \ldots\rangle = \prod_k |\alpha_k\rangle.$$

The vector space spanned by the vectors $|\{\alpha_k\}\rangle$ is the direct product space of those spanned by the vectors $|\alpha_k\rangle$.

It has been shown [39] that an antenna having a known current distribution and suffering no unpredictable reaction from the surrounding field will produce an electromagnetic field that is in a coherent state. When the field is in the coherent state $|\{\alpha_k\}\rangle$, the classical waveform of the electric field can be obtained from (45) and (54) as the expected value of the operator $\underset{\sim}{E}(\underset{\sim}{r}, t)$,

$$\underset{\sim}{\mathscr{E}}(\underset{\sim}{r}, t) = \langle\{\alpha_k\}|\underset{\sim}{E}(\underset{\sim}{r}, t)|\{\alpha_k\}\rangle$$

$$2 \,\mathrm{Im}\left[\sum_{k} (\hbar\omega_k/2\epsilon_o V)^{1/2}\, \alpha_k \exp[-i(\omega_k t - \underset{\sim}{k}\cdot\underset{\sim}{r})]\right]. \tag{55}$$

An extensive calculus involving coherent-state vectors has been developed by Glauber [40, 41]. In particular, it has been shown that a large class of density operators, including those met in communication theory, can be expanded in terms of them,

$$\rho = \int P(\{\alpha_k\}) \prod_{k=1}^{\infty} |\alpha_k\rangle\langle\alpha_k|\, d^2\alpha_k, \tag{56}$$

where the function $P(\{\alpha_k\})$ is called the weight function. This expansion is called the P-representation of the density operator ρ. The weight function $P(\{\alpha_k\})$ has many of the properties of a classical probability density function, but it is not always positive. In particular,

$$\int P(\{\alpha_k\}) \prod_{k=1}^{\infty} d^2\alpha_k = 1$$

follows from $\mathrm{Tr}\,\rho = 1$. The expected value of an operator Ξ is given by

$$\underset{\sim}{E}[\Xi] = \mathrm{Tr}[\rho\Xi] = \int P(\{\alpha_k\})\langle\{\alpha_k\}|\Xi|\{\alpha_k\}\rangle \prod_{k=1}^{\infty} d^2\alpha_k$$

when the state of the field is specified by the density operator ρ in (56).

Representation of Noise

For the moment, let us suppose that the field inside the cavity consists in thermal radiation alone. When this random field is in thermal equilibrium at an absolute temperature $\mathscr{T}$, the density operator ρ_k describing the state of the k^{th} normal mode in the P-representation is

$$\rho_k = \int \exp[-|\alpha|^2/\mathscr{N}_k]|\alpha\rangle\langle\alpha|\, d^2\alpha/\pi\mathscr{N}_k, \tag{57}$$

where

$$\mathscr{N}_k = \text{Tr}\,[\rho_k a_k^+ a_k] = \{\exp(\hbar\omega_k/K\mathscr{T})-1\}^{-1} \tag{58}$$

is the average number of photons in the k^{th} mode with frequency ω_k [42]. $K = 1.38 \times 10^{-23}$ J/deg is Boltzmann's constant. From (51), it follows that ρ_k can be expanded in terms of the eigenvectors $|n_k\rangle$ of the number operator $a_k^+ a_k$,

$$\begin{aligned} \rho_k &= \sum_{n_k=0}^{\infty} (1-v_k)\, v_k^{n_k} |n_k\rangle\langle n_k|, \\ v_k &= \mathscr{N}_k/(\mathscr{N}_k+1) = \exp(-\hbar\omega_k/K\mathscr{T}). \end{aligned} \tag{59}$$

In the classical limit, $K\mathscr{T} \gg \hbar\omega_k$, the weight function

$$\begin{aligned} P(\alpha_k) &= (\pi\mathscr{N}_k)^{-1} \exp[-|\alpha_k|^2/\mathscr{N}_k] \\ &= (\pi\mathscr{N}_k)^{-1} \exp\left[-\left(\alpha_{kx}^2+\alpha_{ky}^2\right)/\mathscr{N}_k\right] \end{aligned}$$

yields the joint probability density function of the real part α_{kx} and imaginary part α_{ky} of the complex amplitude α_k of the mode [43]. Since $\mathscr{N}_k$ in (58) becomes approximately equal to $K\mathscr{T}/\hbar\omega_k$, it follows from (42) that the average energy of this mode equals $K\mathscr{T}$ independently of its frequency. In classical communication theory, this type of noise is called the additive white Gaussian noise with spectral density $\mathscr{N}_o = K\mathscr{T}$

When a normal mode of the received field contains both thermal noise and a coherent signal that alone is represented by the coherent state vector $|\mu_k\rangle$, the center of the Gaussian weight function in the P-representation is simply shifted from the origin by a phasor μ_k. The density operator ρ_k becomes

$$\rho_k = (\pi \mathscr{N}_k)^{-1} \int \exp[-|\alpha-\mu_k|^2/\mathscr{N}_k]\, |\alpha\rangle\langle\alpha|\, d^2\alpha. \tag{60}$$

In the representation of the operator ρ_k as a matrix in terms of the basis specified by the eigenvectors $|n_k\rangle$ of the number operator $a_k^+a_k$, the matrix elements are [44]

$$\langle n|\rho_k|m\rangle = (1-v_k)(n!/m!)^{1/2}\, v_k^m \left(\mu_k^*/\mathscr{N}_k\right)^{m-n} \exp[-(1-v_k)|\mu_k|^2]$$

$$\times L_n^{m-n}[-(1-v_k)^2\,|\mu_k|^2/v_k], \qquad m \geq n \tag{61}$$

$$\langle n|\rho_k|m\rangle = \langle m|\rho_k|n\rangle^*, \qquad m < n$$

$$v_k = \mathscr{N}_k/(\mathscr{N}_k+1),$$

where $L_n^{m-n}(x)$ is the associated Laguerre polynomial.

The density operator ρ for the entire received field when it contains only thermal radiation in equilibrium is given by the direct product

$$\rho = \pi^{-\nu} \prod_k \int \exp\left[-|\alpha_k|^2/\mathscr{N}_k\right] |\alpha_k\rangle\langle\alpha_k|\, d^2\alpha_k/\pi\mathscr{N}_k, \tag{62}$$

where ν is the number of modes. The operator ρ is defined in the linear vector space which is the direct product of the linear vector spaces spanned by the coherent state vectors of the individual modes. When a coherent signal is present in the cavity with Gaussian thermal noise, the density operator, in the P-representation, is

$$\rho = \int \cdots \int P(\{a_k\}) \prod_k |a_k\rangle\langle a_k| \, d^2 a_k,$$

$$P(\{a_k\}) = \pi^{-\nu} |\det \underset{\sim}{\phi}|^{-1} \exp\left[- \sum_m \sum_n \left(a_m^* - \mu_m^*\right)(\underset{\sim}{\phi}^{-1})_{mn}(a_n - \mu_n) \right], \tag{63}$$

where μ_m is the complex amplitude of the coherent signal in mode m. Here $\underset{\sim}{\phi}$ is the mode correlation matrix, whose elements are

$$\phi_{nm} = \text{Tr}\left[\rho a_m^+ a_n\right] - \text{Tr}\left[\rho a_m^+\right] \text{Tr}\left[\rho a_n\right]. \tag{64}$$

When the modes are statistically independent, the mode correlation matrix $\underset{\sim}{\phi}$ is diagonal,

$$\phi_{mk} = \mathcal{N}_k \delta_{mk}, \tag{65}$$

where $\mathcal{N}_k$ is the average number of thermal photons in mode k and is given in terms of the frequency ω_k by (58) when the modes are in thermal equilibrium.

At this point, let us note that the P-representation of a density operator ρ is not unique. Instead of the coherent states $|\{a_k\}\rangle$, ρ can be expressed in the P-representation in terms of the right eigenvectors of a set of operators b_j, where

$$b_j = \sum_{k=1}^{\infty} V_{jk} a_k, \tag{66}$$

and the coefficients V_{jk} are elements of an unitary matrix V. That is,

$$\sum_{k=1}^{\infty} V_{jk} V_{kn}^+ = \sum_{k=1} V_{jk} V_{nk}^* = \delta_{jk}. \tag{67}$$

An easy algebraic manipulation shows that the operators b_j and their Hermitian adjoints b_j^+ satisfy the same commutation relations as the operators a_k and a_k^+. Hence, b_j and b_j^+ can be regarded as the annihilation and creation operators, respectively, of a set of new modes. The right eigenvectors $|\{\beta_k\}\rangle$ of the operators b_j

$$b_k|\{\beta_k\}\rangle = \beta_k|\{\beta_k\}\rangle$$

are coherent-state vectors spanning the same vector space as spanned by the vectors $|\{\alpha_k\}\rangle$. When the density operator ρ is expressed in terms of $|\{\beta_k\}\rangle$ in the P-representation

$$\rho = \int P'(\{\beta_k\})|\{\beta_k\}\rangle\langle\{\beta_k\}| \prod_{k=1}^{\infty} d^2\beta_k \tag{68}$$

the new weight function $P'(\{\beta_k\})$ can be obtained by substituting the relation

$$\alpha_k = \sum_{j=1}^{\infty} \beta_j V_{jk}^+ \tag{69}$$

in the weight function $P(\{\alpha_k\})$.

If the unitary matrix $\underset{\sim}{V}$ is such that the matrix $\underset{\sim}{V}\underset{\sim}{\phi}^{-1}\underset{\sim}{V}^+$ is diagonal, the density operator ρ in (63), when expressed in terms of eigenvectors of the operators b_j given by (66), is

$$\rho = \pi^{-\nu}\int \exp\left[-\sum_{k=1}^{\infty} |\beta_k-\mu_k'|^2/\mathcal{N}_k'\right]|\{\beta_k\}\rangle\langle\{\beta_k\}| \prod_k d^2\beta_k/\mathcal{N}_k', \tag{70}$$

where $\mathcal{N}_k'$ is the k^{th} diagonal element of the matrix $\underset{\sim}{V}\underset{\sim}{\phi}^{-1}\underset{\sim}{V}^+$, and

$$\mu_k' = \sum_{m=1}^{\infty} V_{km}\mu_m. \tag{71}$$

Therefore, for any particular coherent signal in Gaussian thermal noise, a set of normal modes can be chosen, by appropriately choosing the shape of the receiver cavity, to represent the received field so that the individual modes are uncorrelated.

When the frequency range of the signal is small so that for all k for which $\mu_k \neq 0$,

$$\mathcal{N}_k \approx \mathcal{N} \tag{72}$$

$\mathcal{N}'_k = \mathcal{N}$ in (70) can also be taken to be the average number of thermal photons in the new normal mode k. The density operator ρ can be expressed in terms of coherent states $|\{\gamma_k\}\rangle$ that are the right eigenvectors of the annihilation operators g_j

$$\rho = (\pi\mathcal{N})^{-\nu} \int \exp\left[-\left\{|\gamma_1-\mu|^2 - \sum_{k\neq 1} |\gamma_k|^2\right\}\Big/\mathcal{N}\right] \prod_k |\{\gamma_k\}\rangle\langle\{\gamma_k\}|\, d^2\gamma_k. \tag{73}$$

In this expression,

$$g_j = \sum_{m=1}^{\infty} V_{jm} b_m, \qquad \gamma_j = \sum_{m=1}^{\infty} V_{jm}\beta_m,$$

where the matrix elements V_{1m} are chosen to be

$$V_{1m} = \mu'_m/|\mu|, \qquad |\mu|^2 = \sum_{m=1}^{\infty} |\mu'_m|^2. \tag{74}$$

The other rows of the matrix $\underset{\sim}{V}$ are chosen so that $\underset{\sim}{V}$ is unitary. We see that only the mode with annihilation operator $g_1 = \sum_m \mu'_m b_m/|\mu|$ contains a coherent signal [18,45]. Hence without loss of generality, we often need to consider only one properly chosen normal mode of the received field.

Quantum-Mechanical Receiver

In an ideal receiver, the signal field accompanied by the chaotic thermal noise is admitted through the receiving aperture into a lossless cavity during the signaling interval (0, T). At the end of this interval, the aperture is closed, and measurements are made by the receiver on the received field, whose quantum-mechanical description has just been presented.

We shall assume that the measurement made by the receiver is ideal. Again, an ideal measurement is one in which the state of the field after the measurement does not depend on that before the measurement. It follows that the probability distribution of the outcome of any subsequent measurement does not depend on the transmitted input message. That is, any measurement made after an ideal measurement yields no information relevant to the optimum estimation of the transmitted message [46].

It will become apparent that the optimum performance of the system is independent of the time — after the receiving aperture is closed — at which the observation is made. The choice of the observables measured by the optimum receiver does depend, however, on the time of the observation.

The ideal receiver that we have discussed thus far might seem to be much too remote from an ordinary receiver to be relevant to a real optical communication system. A real optical receiver takes in light from the signal source, along with thermal radiation, through an aperture of fixed size, and processes this light by lenses, photodetectors, and possibly coherent heterodyning light generated by a local laser. The data upon which it bases its decisions are the values of observables of the electromagnetic field at the aperture during the interval (0, T). The field in the cavity of our ideal receiver is a linear functional of this aperture field. The optimum performance derived for the ideal receiver really sets a limit to the performance of any optical receiver processing the same aperture field.

When the signal radiation occupies a narrow band of frequencies and arrives from a narrow cone of directions, and when the background radiation is distributed broadly in frequency and angle, the quantum detection theory developed for the field in the ideal receiver can be applied to the aperture field itself. The important entities in that theory are the annihilation and creation operators for the mode fields and the Hilbert space spanned by their eigenvectors. Operators having the same properties can be defined for the aperture field by representing it as a superposition of spatio-temporal modes. Just as the mode functions for the cavity field are orthonormal with respect to integration over the three-dimensions of the cavity, these spatio-temporal modes are orthonormal for integration over the aperture and the observation interval (0, T). The eigenvectors of the associated annihilation and creation operators span a Hilbert space of state vectors to which the concepts and techniques that we have outlined can be applied [47, 48].

IV QUANTUM DETECTION THEORY

When the data source in Fig. 1 is digital, the input m to the transmitter in the signaling interval (0, T) is one of M messages denoted $m_1, m_2, \ldots m_M$. When the transmitted message is m_j, the electromagnetic field in the receiver cavity is in the statistical mixture of states specified by the density operator ρ_j. Therefore, in the time interval (0, T), the state of the received field is specified by one of M density operators $\rho_1, \rho_2, \ldots \rho_M$.

The output $\hat{m}$ of the system is also one of M messages, and is the estimate of the input message m. That is, the receiver decides in the time interval (0, T) among the M hypotheses $H_1, H_2, \ldots H_M$, of which the hypothesis H_j is that the message m_j is transmitted. The receiver is designed so that the probability of error

$$P_e = \Pr[\hat{m} \neq m] \tag{75}$$

is minimum.

Let $\underset{\sim}{X} = (X_1, X_2, \ldots X_L)$ denote the L-tuple of Hermitian operators corresponding to those observables chosen to be measured by the receiver. When these operators commute, a simultaneous measurement of the corresponding observables yields an L-tuple $\underset{\sim}{x}_n = (x_{1n}, x_{2n}, \ldots x_{Ln})$ of parameters, where x_{jn} is an eigenvalue of the operator X_j; j = 1, 2, ... L. For simplicity, we assume that the eigenspectra of the operators X_j and ρ_j are discrete. That this assumption imposes no real restriction has been pointed out in Section II.

Let $|\underset{\sim}{x}_n\rangle$ denote the simultaneous eigenvector of the commuting operators $X_1, X_2, \ldots X_L$ corresponding to the eigenvalue $\underset{\sim}{x}_n$. From (23), the conditional probability that the outcome of the measurement of $\underset{\sim}{X}$ is $\underset{\sim}{x}_n$, given that the message m_j is transmitted, is

$$P(\underset{\sim}{x}_n | m_j) = \langle \underset{\sim}{x}_n | \rho_j | \underset{\sim}{x}_n \rangle. \tag{76}$$

Let ζ_j be the prior probability of the message m_j, and p_{jn} be the probability that the receiver chooses the hypothesis H_j when the outcome of the measurement is $\underset{\sim}{x}_n$.

$$\sum_{j=1}^{M} p_{jn} = 1. \tag{77}$$

The probability of error in (75) is

$$P_e = 1 - \sum_{n} \sum_{j=1}^{M} \zeta_j p_{jn} \langle \underset{\sim}{x}_n | \rho_j | \underset{\sim}{x}_n \rangle . \tag{78}$$

The manner in which the optimum receiver processes the data obtained in the measurement is the same as that determined by the principles of classical detection theory [49]. Specifically, the receiver chooses the hypothesis H_j to minimize P_e when the observed value of $\underset{\sim}{X}$ is $\underset{\sim}{x}_n$ if the conditional probability

$$P(m_j | \underset{\sim}{x}_n) = \zeta_j P(\underset{\sim}{x}_n | m_j) \Big/ \sum_{i=1}^{M} \zeta_i P(\underset{\sim}{x}_n | m_i)$$

is maximum. In other words, the probability p_{jn} is 1 for all n such that

$$\zeta_j \langle \underset{\sim}{x}_n | \rho_j | \underset{\sim}{x}_n \rangle \geq \zeta_i \langle \underset{\sim}{x}_n | \rho_i | \underset{\sim}{x}_n \rangle, \qquad \text{all } i \neq \tag{79}$$

and all other p_{in} equal zero. This rule becomes ambiguous when the equality sign in (79) holds for some i. The ambiguity can be resolved, however, and the resultant minimum value of P_e is not affected by the way in which this ambiguity is resolved.

Let us define the operators Π_j as

$$\Pi_j = \sum_{n} p_{jn} | \underset{\sim}{x}_n \rangle \langle \underset{\sim}{x}_n | ; \qquad j = 1, 2, \ldots M. \tag{80}$$

Since the probabilities p_{in} are either one or zero, the operators Π_i are projection operators; therefore, they obey the defining equation (14). Moreover, it follows from (80), and that $p_{in}p_{jn}$ equals zero when $i \neq j$, that

$$\Pi_i \Pi_j = \Pi_i \delta_{ij} \tag{81}$$

and

$$\sum_{j=1}^{M} \Pi_j = \underset{\sim}{1}. \tag{82}$$

In terms of the projection operators Π_j, the probability of error P_e in (78) becomes

$$P_e = 1 - \sum_{j=1}^{M} \zeta_j \operatorname{Tr} [\rho_j \Pi_j]. \tag{83}$$

Therefore, the problem of finding the best receiver structure becomes that of finding the projection operators Π_j that satisfy the constraints (81) and (82) and minimize P_e.

It has been shown [50, 51] that a necessary condition for the set of projection operators Π_j satisfying constraints (81) and (82) to minimize P_e in (83) is

$$\sum_{j=1}^{M} \zeta_j \Pi_j \rho_j = \sum_{j=1}^{M} \zeta_j \rho_j \Pi_j. \tag{84}$$

This equation, together with the conditions

$$\left\{ \sum_{j=1}^{M} \zeta_i \Pi_i \rho_i - \zeta_j \rho_j \right\} \begin{array}{l} \text{are positive semidefinite} \\ \text{for all } j = 1, 2, \ldots M \end{array} \tag{85}$$

provides a sufficient condition for the set of projection operators Π_j to be an optimum solution.

When the projection operators Π_j have a complete set of eigenvectors in common, they can be taken as observables of the field, and the receiver measures them simultaneously. Hypothesis H_j is chosen when the observed value of Π_j is equal to 1. Because of (82), an optimum receiver in a binary communication system is specified simply by one projection operator Π_1.

When the receiver is allowed to make ideal measurements of incompatible observables in the sense discussed in Section II, the structure of the receiver is specified by the overcomplete set of measurement state vectors, $\{|\underset{\sim}{x}_n\rangle\}$, when the outcomes of the measurements are $\underset{\sim}{x}_n$. Since with $\{|\underset{\sim}{x}_n\rangle\}$ given, the probabilities p_{jn} are determined by the rule (79), the problem of finding the optimum receiver structure is that of finding the overcomplete set $\{|\underset{\sim}{x}_n\rangle\}$ to minimize

$$P_e = 1 - w \sum_n \max_{1 \leq j \leq M} \zeta_j \langle \underset{\sim}{x}_n | \rho_j | \underset{\sim}{x}_n \rangle \tag{86}$$

Very little is known about the solution of this minimization problem. When only orthonormal sets of measurement state vectors are allowed as solutions, this problem is equivalent to finding the operators Π_j, subject to constraints (81) and (82) to minimize P_e in (83). In general, the two maximization problems are not equivalent. [52].

To find optimum receivers in many communication systems of practical interest, there is no need to consider ideal measurements of incompatible observables. It can be shown from the completeness (32) of $\{|\underset{\sim}{x}_n\rangle\}$ that in all binary communication systems and in those M-ary systems in which the density operators ρ_j commute, optimum receivers measure observables corresponding to Hermitian commuting operators [53].

Binary Detection

We are concerned here with binary communication systems in which the input message in the time interval (0, T) is either the digit "1" or the digit "0". A digit "1" is represented by the presence of a signal pulse of duration T; a digit "0" is represented by the absence of the pulse. Therefore, in the time interval (0, T) the ideal receiver chooses between two hypotheses: (H_0) "the field in the cavity is due only to thermal radiation," and (H_1) "the field contains besides thermal radiation a signal of some specified form." The best receiver is one that enables the choice between the two hypotheses to be made with minimum probability of error.

Hypothesis H_1 represents a proposition that is either true or false. We have seen that such propositions are decided by measuring a projection operator Π. The outcome of the measurement is one of the two eigenvalues of Π, 0 or 1. If 1, hypothesis H_1 is adopted; if 0, H_0. The question remains, however, which of all of the projection operators Π that exist for the received field is best. It is answered in the same way as classical detection theory: We must use the operator for which the average probability of error is minimum [15, 18, 54].

Under hypothesis H_0 the normal modes of the receiver are excited only by random noise; they are in a mixture of states described by a density operator ρ_0, such as the one exhibited in (62). Under hypothesis H_1 the normal modes are in some other mixture of states described by a density operator ρ_1 such as the one in (63). Let ζ be the prior probability of hypothesis H_0. The average probability P_e of error in (83) can be rewritten

$$\begin{aligned} P_e &= \zeta \operatorname{Tr} [\rho_0 \Pi] + (1-\zeta)[1-\operatorname{Tr}[\rho_1 \Pi]] \\ &= (1-\zeta)\{1-\operatorname{Tr}[(\rho_1-\Lambda_0\rho_0)\Pi]\}, \qquad \Lambda_0 = \zeta/(1-\zeta). \end{aligned} \tag{87}$$

This quantity is to be minimized by properly choosing the projection operator Π. The minimizing operator we call the detection operator. We have put P_e in (87) in such a form that the problem of maximizing

$$\mathrm{Tr}\,[(\rho_1-\Lambda_o\rho_o)\Pi]$$

is clearly equivalent.

Let η_k be the eigenvalues, and $|\eta_k\rangle$ the associated eigenvectors of the operator $\rho_1 - \Lambda_o\rho_o$.

$$(\rho_1-\Lambda_o\rho_o)|\eta_k\rangle = \eta_k|\eta_k\rangle. \tag{88}$$

Then

$$\mathrm{Tr}\,[(\rho_1-\Lambda_o\rho_o)\Pi] = \sum_k \langle\eta_k|(\rho_1-\Lambda_o\rho_o)\Pi|\eta_k\rangle$$

$$= \sum_k \eta_k\langle\eta_k|\Pi|\eta_k\rangle. \tag{89}$$

This quantity is maximum for that projection operator for which $\langle\eta_k|\Pi|\eta_k\rangle = 1$ when $\eta_k \geqslant 0$, and $\langle\eta_k|\Pi|\eta_k\rangle = 0$ when $\eta_k < 0$, and the projection operator that fulfills the requirement is

$$\Pi = \sum_k U(\eta_k)|\eta_k\rangle\langle\eta_k|, \tag{90}$$

where U(x) is the unit step function. The average probability of error is then

$$P_e = (1-\zeta)\left[1 - \sum_k \eta_k U(\eta_k)\right]. \tag{91}$$

Prescription of the optimum receiver is simplest when the density operators ρ_o and ρ_1 commute. The vectors $|\eta_k\rangle$ are then identical with the eigenvectors $|\phi_k\rangle$ common to both ρ_o and ρ_1, which can now be written

$$\rho_j = \sum_k P_k^{(j)} |\phi_k\rangle\langle\phi_k|, \qquad j = 0, 1,$$

$$\sum_k P_k^{(j)} = 1. \tag{92}$$

In fact, the eigenvalues η_k of $\rho_1 - \Lambda_o\rho_o$ are

$$\eta_k = P_k^{(1)} - \Lambda_o P_k^{(0)}. \tag{93}$$

If the system is in such a state $|\phi_k\rangle$ that $\eta_k > 0$, or equivalently,

$$P_k^{(1)}/P_k^{(0)} > \Lambda_o,$$

hypothesis H_1 is chosen; otherwise H_o is adopted. This is just the classical likelihood-ratio test.

Suppose, for instance, that the signal has the same statistical properties as thermal noise, placing an average number N_s of photons into a single mode of the receiver and none into any of the others. Only that mode needs to be observed, and we assume that under hypothesis H_o it contains thermal noise with an average number $\mathcal{N}$ of photons. The density operators under the two hypotheses then, by (59), are

$$\rho_i = (1-v_i)\sum_{m=0}^{\infty} v_i^m |m\rangle\langle m|, \qquad i = 0, 1, \tag{94}$$

where

$$v_o = \mathcal{N}/(\mathcal{N}+1), \qquad v_1 = (\mathcal{N}+N_s)/(\mathcal{N}+N_s+1). \tag{95}$$

These density operators commute, and since both commute with the number operator a^+a, it suffices for optimum detection to count the number m of photons in the mode. The likelihood ratio is

$$P_m^{(1)}/P_m^{(0)} = \left(\frac{1-v_1}{1-v_o}\right)(v_1/v_o)^m \tag{96}$$

and it is decided that a signal is present whenever this ratio exceeds $\Lambda_o = \zeta/(1-\zeta)$; that is, when

$$m > \left[\ln \Lambda_o - \ln\left(\frac{1-v_1}{1-v_o}\right)\right] \Big/ \ln(v_1/v_o) = M.$$

The probability of error is then

$$P_e = \zeta v_o^{[M]+1} + (1-\zeta)\left(1-v_1^{[M]+1}\right), \tag{97}$$

where [M] is the greatest integer in the decision level M.

a. Choice between Pure States

When the density operators ρ_o and ρ_1 do not commute, it is necessary to solve the eigenvalue equation (88) before the structure of the optimum receiver can be specified and its performance assessed. This is generally very difficult. In a special case that can be solved exactly, the received field is in one pure state $|\psi_o\rangle$ under hypothesis H_o and in another, $|\psi_1\rangle$, under H_1. The density operators are

$$\rho_o = |\psi_o\rangle\langle\psi_o|, \qquad \rho_1 = |\psi_1\rangle\langle\psi_1|. \tag{98}$$

Unless $|\psi_o\rangle$ and $|\psi_1\rangle$ happen to be orthogonal, ρ_o and ρ_1 do not commute. An example is the detection of a coherent signal in the absence of any thermal radiation. The eigenvectors $|\eta_i\rangle$ are now linear combinations of $|\psi_o\rangle$ and $|\psi_1\rangle$,

$$|\eta_i\rangle = a_{io}|\psi_o\rangle + a_{i1}|\psi_1\rangle, \qquad i = 1, 2. \tag{99}$$

Only two eigenvalues differ from zero, and they are found by substituting (99) in (88) and setting the determinant of the coefficients in the resulting pair of simultaneous equations equal to zero. The minimum average probability of error is found to be [16, 55]

$$P_e = (1-\zeta)\left[\frac{1}{2}(1+\Lambda_o) - \mathcal{P}\right]$$

$$\mathcal{P} = \left\{\left[\frac{1}{2}(1-\Lambda_o)\right]^2 + \Lambda_o \mathcal{q}\right\}^{1/2}$$

$$\mathcal{q} = 1 - |\langle \psi_o | \psi_1 \rangle|^2 \qquad (100)$$

In the detection of a coherent signal occupying a single mode in the absence of thermal noise, the states $|\psi_o\rangle$ and $|\psi_1\rangle$ are, respectively, the coherent states $|0\rangle$ and $|\mu\rangle$, where $|\mu|^2 = N_s$ is the average number of signal photons. Then in (100), by (53),

$$\mathcal{q} = 1 - |\langle 0 | \mu \rangle|^2 = 1 - \exp(-N_s). \qquad (101)$$

The error probability P_e' is plotted in Fig. 2 against the signal-to-noise ratio

$$D = [4N_s/(2\mathcal{N}+1)]^{1/2} \qquad (102)$$

as the curve marked $\mathcal{N} = 0$; the prior probabilities were taken as $\zeta = 1 - \zeta = 1/2$, and $\Lambda_o = 1$. This signal-to-noise ratio D goes into the classical signal-to-noise ratio, $[2 N_s \hbar\omega/K\mathcal{T}]^{1/2}$, in the limit $\hbar\omega \ll K\mathcal{T}$.

b. Detection of a Coherent Signal in Thermal Noise

When the mode excited by a coherent signal also contains thermal noise of absolute temperature $\mathcal{T}$, the density operators ρ_o and ρ_1 take the forms in (57) and (60), respectively. An exact solution of the eigenvalue equation (88) with these density operators has not been obtained. It is possible to solve it approximately by using the matrix representation of ρ_1 in (61) and diagonalizing a truncated version of the infinite matrix by means of a digital computer. Figure 2 gives the error probabilities so obtained; the largest matrix found necessary used the first fifteen rows and columns of the infinite matrix $\langle n|(\rho_1-\Lambda_o\rho_o)|m\rangle$.

When the average number $\mathcal{N}$ of thermal photons is very large, the classical limit is approached. The classical detector of a signal of known phase in thermal noise corresponds in this model to measuring the component of the operator a along the phase of the signal. If we take, without loss of generality, the phase arg μ as zero, the classical detector measures the operator Q and compares it with a decision level. The probability density functions of the outcome q under hypotheses H_o and H_1, respectively, are [56]

$$P_o(q) = \langle \alpha|\rho_o|\alpha\rangle = (2\pi\sigma^2)^{-1/2} \exp[-q^2/2\sigma^2]$$

and

$$P_1(q) = \langle \alpha|\rho_1|\alpha\rangle = (2\pi\sigma^2)^{-1/2} \exp[-\{q-(2\hbar/\omega)^{1/2}\mu\}^2/2\sigma^2],$$

where μ now is real and $\sigma^2 = \hbar\left(\mathcal{N}+\frac{1}{2}\right)/\omega$.

The decision level q_o with which q is compared is determined from the likelihood-ratio formula,

$$P_1(q_o)/P_o(q_o) = \Lambda_o. \tag{103}$$

It is not hard to show that the probability of error by using this classical detector is

$$P_e = \zeta \operatorname{erfc}\left(\frac{1}{2}D + D^{-1}\ln\Lambda_o\right) + (1-\zeta)\operatorname{erfc}\left(\frac{1}{2}D - D^{-1}\ln\Lambda_o\right), \tag{104}$$

where D is given by (101) and

$$\operatorname{erfc} x = (2\pi)^{-1/2}\int_x^\infty \exp(-t^2/2)\,dt \tag{105}$$

is the error-function integral. For $\zeta = \frac{1}{2}$ the error probability is simply

$$P_e = \text{erfc}\left(\frac{1}{2}D\right), \tag{106}$$

which has been plotted against D in Fig. 2 as the curve marked $\mathcal{N} = \infty$. This classical detector can be used at any value of $\mathcal{N}$, but it suffers a higher probability of error than the optimum detector. The difference between the two detectors vanishes rapidly once the average number $\mathcal{N}$ of thermal photons in the mode becomes of the order of 1.

If the signal occupies many modes of the received field, the problem appears much more complicated. If it is a narrow-band signal, however, all of the modes excited by the signal will have nearly the same average number $\mathcal{N}$ of thermal photons. It is then possible to combine the amplitudes of those modes linearly in such a way that the resulting field amplitude contains the entire signal, in effect creating a new mode matched to the signal, as discussed in Section III. By means of other linear combinations a new set of modes orthogonal to the matched mode and containing no signal excitation is formed. The problem is then reduced to detection of the signal in a single mode and the results just derived apply [57].

If the absolute phase ϕ of the signal is unknown, as will happen if no attempt is made to maintain phase coherence between transmitter and receiver, the density operator ρ_1 must be averaged with respect to this phase. In the least favorable situation, the phase is a random variable uniformly distributed from 0 to 2π.

The elements of the matrix $\langle n|\rho_1|m\rangle$ specifying ρ_1 in the number representation are given in (61). When we average with respect to ϕ over $(0, 2\pi)$, all of the off-diagonal elements will be zero. The average density operator $\langle \rho_1 \rangle$ then, like ρ_o, is diagonal in the number representation

$$\langle \rho_1 \rangle = \sum_{k=0}^{\infty} P_k^{(1)} |k\rangle\langle k|, \tag{107}$$

with

$$P_k^{(1)} = (1-v)\, v^k \exp[-(1-v)N_s]\, L_n\left[-(1-v)^2 N_s/v\right.$$

$$N_s = |\mu|^2, \tag{108}$$

where $L_n(x)$ is the ordinary Laguerre polynomial.

The optimum receiver, therefore, simply counts the number of photons in the matched mode and compares it with a decision level determined by the likelihood-ratio formula (103). The average error probability can be computed from (108) [58]. Curves calculated for $\zeta = (1-\zeta) = \frac{1}{2}$ by digital computer are given in Figs. 3 and 4. In the limit $\mathcal{N} = \infty$ the error probability coincides with that of the classical detector of a signal of random phase.

M-ary Detection

In general, there may exist no solution to the problem of finding projection operators Π_j, satisfying the constraints (81) and (82), to minimize the probability of error P_e in (83). Here, our attention will be confined to the special case in which the density operators ρ_j commute. For this case, the projection operators Π_j, which satisfy the sufficient conditions given by (84) and (85), can be written

$$\Pi_j = \sum_n p_{jn} |\phi_n\rangle\langle\phi_n|, \tag{109}$$

where $|\phi_n\rangle$ are the simultaneous eigenvectors of the density operators ρ_j, j = 1, 2, ... M. Let $p_n^{(j)}$ denote the eigenvalue of the density operator ρ_j corresponding to the eigenvector $|\phi_n\rangle$. For any j = 1, 2, ... M and n = 1, 2, ..., the probability p_{jn} in (109) equals 1 if $\zeta_j p_n^{(j)} \geq \zeta_i p_n^{(i)}$ for all $i \neq j$, and equals zero otherwise.

For simplicity, we assume that the M messages $m_1, m_2, \ldots m_M$ occur with equal prior probabilities. Therefore, the information rate R, in nats per second, of the communication system is

$$R = [\ln M]/T. \tag{110}$$

a. Detection of Orthogonal Signals with Known Amplitudes but Random Phases

Let us envision an M-ary communication system in which each transmitted message is represented by a signal pulse of duration T with known classical amplitude, but an unknown absolute phase. Let z denote the direction of signal propagation. When the signal field is not modulated spatially and is linearly polarized, its classical waveform at the receiver can be described by the function

$$2\mathcal{R}\ell\,[S(t)\exp\{i\omega_o(z/c-t)+i\phi\}],$$

where S(t) is one of a set of complex time functions $\{S_j(t)\}$ depending on the transmitted message. The absolute phase ϕ of the signal is taken to be a random variable uniformly distributed over the interval $(0, 2\pi)$. In particular, we assume that the functions $S_j(t)$ are essentially narrow-band [59] orthogonal time functions over any signaling interval,

$$\int_0^T S_j(t)\,S_i^*(t)\,dt = E^2\delta_{ij}; \qquad i,j = 1,2,\ldots M. \tag{111}$$

To describe this signal set quantum-mechanically, we regard the received field to be a superposition of plane traveling-wave modes with normal mode functions $V^{-1/2}\exp[-i\omega_k z/c]$. When the transmitted message is m_j, the received field would be in one of the coherent states $\prod_k |\mu_{jk}e^{i\phi}\rangle$ in the absence of thermal radiation. The complex amplitudes μ_{jk} of the individual modes are known and are related to the waveforms $S_j(t)$ by

$$S_j(t) = \sum_n (\hbar\omega_n/2\epsilon_o V)^{1/2}\, i\mu_{jn}\exp[i(\omega_n-\omega_o)(z/c-t)], \qquad j = 1, 2, \ldots M.$$

In most practical systems, the width of the frequency range of interest is small compared with the carrier frequency ω_o; therefore, we may assume that for all n for which $\mu_{jn} \neq 0$,

$$\omega_n \approx \omega. \tag{112}$$

Since the complex functions $S_j(t)$ are orthogonal, the complex amplitudes μ_{jn} satisfy the condition

$$\sum_n \mu_{jn}\mu_{kn}^* = |\mu_j|^2 \delta_{jk}; \qquad j, k = 1, 2, \ldots M. \tag{113}$$

In the presence of an additive thermal-noise field, the received field is the mixed state specified by the density operator

$$\rho_j = \int_0^{2\pi} \frac{d\phi}{2\pi} \int \prod_k \exp\left[-|\alpha_k - \mu_{jk} e^{i\phi}|^2/\mathcal{N}_k\right] |\alpha_k\rangle\langle\alpha_k| d^2 \alpha_k/\pi\mathcal{N}_k \tag{114}$$

when the input message is m_j. Since by (112), the average numbers $\mathcal{N}_k$ of thermal photons in different modes are approximately equal, $\mathcal{N}_k \approx \mathcal{N}$, the density operators ρ_j are commutative [60]. For the reception of these M signals, the optimum receiver measures simultaneously the projection operators Π_j in (109). Equivalently, it measures the numbers of photons in the M modes associated with normal-mode functions $S_j(t)/E^2$ and with annihilation operators $b_j = \sum_n \mu_{jn} a_n / |\mu_j|$, where [61]

$$|\mu_j|^2 = \sum_n |\mu_{jn}|^2$$

When the outcome is $\underset{\sim}{n} = (n_1, n_2, \ldots n_M)$, hypothesis H_j, "the transmitted message is m_j," is chosen if the conditional probability $P(\underset{\sim}{n}|m_j)$ is the largest among all j. For M signals with disjoint frequency spectra, as in a frequency-position modulation system, the optimum receiver can be implemented as shown in Fig. 5. The function of the mode transformation filter is similar to that of an optical matched filter [62]. The optimum receiver for a set of pulse-position modulated signals is simply a direct detection system.

To compute the performance of this system, we shall assume for simplicity that

$$|\mu_j|^2 = pT. \tag{115}$$

Since $|\mu_j|^2$ is the average number of photons in the signal field when m_j is transmitted, p is the average number of signal photons arriving at the receiver per unit time, which by (115) is independent of the transmitted message. In this case, the probability of error, P_e, can be bounded [63] as

$$K_2 \exp[-T\, CE(R)] \leq P_e \leq K_1 \exp[-T\, CE(R)], \tag{116}$$

where the coefficients K_1 and K_2 are not exponential functions of T. The information rate R is given by (110), C is the channel capacity,[2] and the exponential factor E(R) is the system reliability function.

The channel capacity C is found to be

$$\begin{aligned} C &= p \ln (1+1/\mathcal{N}); \qquad \mathcal{N}_k \approx \mathcal{N} \\ &= \frac{P}{2\eta_z} \ln (1+2\eta_z/\eta_o), \end{aligned} \tag{117}$$

where $P = p\hbar\omega$ is the average received signal power, $\eta_z = \frac{1}{2}\hbar\omega$ is the so-called zero-point fluctuation energy, and $\eta_o = \mathcal{N}\hbar\omega$ is the average thermal energy in each mode of the received field. In the classical limit with $\mathcal{N} \gg 1$, the channel capacity C is approximately equal to $C_c = P/\eta_o$, the capacity of the classical additive white Gaussian channel with no bandwidth constraints [64] and with noise spectral density $\eta_o \approx K\mathcal{T}$.

For rates R in the range $0 \leq R \leq R_c$, where R_c is given by

$$R_c/C = \mathcal{N}(1+\mathcal{N})/(1+2\mathcal{N})^2, \tag{118}$$

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

$$E(R) \geq [(1+2\mathcal{N})\ln(1+1/\mathcal{N})]^{-1} - R/C. \tag{119}$$

For rates R in the range $R_c \leq R \leq C$, the reliability function E(R) is given by

$$E(R) = \{\mathcal{N}\ln(1+1/\mathcal{N})\}^{-1}\left\{\frac{1+2\mathcal{N}R/C - [1+4R\mathcal{N}(1+\mathcal{N})/C]^{1/2}}{1-[1+4R\mathcal{N}(1+\mathcal{N})/C]^{1/2}} - \ln\left\{\frac{1+2\mathcal{N}(1+\mathcal{N})R/C + [1+4R\mathcal{N}(1+\mathcal{N})/C]^{1/2}}{2R(1+\mathcal{N})^2/C}\right\}\right\} \tag{120}$$

The general behavior of the system reliability function for several different values of $\mathcal{N}$ is shown in Fig. 6. In contrast to the corresponding classical additive white Gaussian channel, E(R) depends not only on R/C, but also on the average noise level $\mathcal{N}$

When $\mathcal{N} = 0$, the channel capacity C in (117) becomes infinite, while E(R) approaches zero. The probability of error becomes

$$P_e \approx \exp[-Tp]. \tag{121}$$

The fact that P_e is independent of the number M of messages when $\mathcal{N}$ equals zero implies that an arbitrarily small probability of error can be achieved at an arbitrarily large information rate for a finite number p of photons per second in the signal field. Since signals are orthogonal, however, the average power in the signal field grows linearly with the number of input messages when p is being held constant. Hence the small probability of error is accomplished only by an accompanied increase in the power of the transmitted signal. It is more meaningful to derive the expressions of the channel capacity and the system reliability function under the assumption that the average power in the signal, $\sum_n |\mu_{jn}(\hbar\omega_k)^{1/2}|^2/T$, is held constant independently of the transmitted message. Unfortunately, analytic results cannot be obtained in this case.

It has been shown [65] heuristically, by use of the uncertainty principle, that, with no bandwidth constraints, the channel capacity has an upper limit $\ell(P/h)^{1/2}$, where P is the average signal power, and ℓ is a number that approximately equals 2.

b. Detection of Orthogonal Signals in a Rayleigh Fading Channel

We now consider a communication system in which the received field in the presence of thermal radiation is in the completely incoherent mixture of states described by the density operator

$$\rho_j = [\det \underset{\sim}{\phi}_j]^{-1} \int \ldots \int \exp\left[-\sum_m \sum_n \alpha_m^* (\underset{\sim}{\phi}_j)^{-1}_{mn} \alpha_n\right] \prod_m |\alpha_m\rangle\langle\alpha_m| \, d^2\alpha_m/\pi^{\nu M} \tag{122}$$

when the transmitted message is m_j. In this expression, ϕ_j is the mode correlation matrix whose elements are

$$[\underset{\sim}{\phi}_j]_{nm} = \mathrm{Tr}\left[\rho_j a_m^+ a_n\right] = \left[\underset{\sim}{\phi}_j^{(t)}\right]_{nm} + \mathcal{N}\delta_{mn}, \tag{123}$$

where $\mathcal{N}$ is the average number of thermal photons in each normal mode. $\left[\underset{\sim}{\phi}_j^{(t)}\right]_{mn}$ are elements of the mode correlation matrix in the absence of the thermal radiation. When the signals have orthogonal classical waveforms at the receiver, the mode correlation matrices $\underset{\sim}{\phi}_j$ commute, and therefore the density operators ρ_j commute. The commutativity of the matrices $\underset{\sim}{\phi}_j$ implies the existence of unitary transformation $\underset{\sim}{V}$ such that the matrices

$$\underset{\sim}{R}_j = \underset{\sim}{V}^+ \underset{\sim}{\phi}_j \underset{\sim}{V}, \qquad j = 1, 2, \ldots M, \tag{124}$$

are diagonal. When the elements of $\underset{\sim}{\phi}_j$ are given by (123), the k^{th} diagonal element of the matrix $\underset{\sim}{R}_j$ can be written

$$[\underset{\sim}{R}_j]_{kk} = \mathcal{N} + s_{jk}; \qquad j = 1, 2, \ldots M. \tag{125}$$

The system described above can serve as a quantum-mechanical model of a diversity transmission system in which signals are transmitted over a Rayleigh fading channel. When a signal pulse with classical waveform $S_j(t) \cos \omega_o t$ is transmitted through such a channel, the classical waveform of the received signal is $xS_j(t) \cos(\omega_o t + \phi)$, where x and ϕ are sample functions of random processes such that at any time t and any point $\underset{\sim}{r}$, $x(\underset{\sim}{r}, t)$, and $\phi(\underset{\sim}{r}, t)$ are Rayleigh-distributed and uniformly distributed random variables, respectively. (For simplicity, we suppose that in a given signaling interval and over the receiving aperture, they can be considered as constant random variables.) It has been shown [66, 67] that the probability of error in the reception of orthogonal signals transmitted over a Rayleigh fading channel can be reduced by making several transmissions for each input message. These transmissions are spaced either in time, space, or frequency so that fadings experienced by different transmissions are statistically independent. Such a system is called a diversity transmission system. The density operators ρ_j in (122) describe, quantum-mechanically, the received field in a diversity transmission system in which each input message is represented by several narrow-band signals with orthogonal classical waveforms [68]. In (125), s_{jk} is the average number of signal photons received at the end of the k^{th} diversity path when the message m_j is transmitted. Let ν denote the number of diversity transmissions; for each j, $s_{jk} \neq 0$ only for ν of the possible values of the subscript k.

Without loss of generality, we confine our discussion to frequency-diversity systems in which frequency-position modulation is used; the frequency spectra of the signals are disjoint. The optimum receiver for the reception of these signals measures simultaneously the observables corresponding to operators $b_n^+ b_n$, where $b_n = \sum_k V_{nk} a_k$; V_{nk} are the elements of the matrix $\underset{\sim}{V}$ in (124). The operators a_k are the annihilation operators in terms of whose right eigenvectors ρ_j is expressed in (122). Again, hypotheses are chosen by using the decision rule (79).

The performance of such a system has been evaluated when the number of signal photons s_{jk} are equal [69]

$$s_{jk} \equiv s \tag{126}$$

for $j = 1, 2, \ldots M$ and all k for which $s_{jk} \neq 0$. For the narrow-band signals considered here, (126) also implies that the signal energies in the ν diversity paths are equal, as in an equal-strength diversity system. For this special case, the structure of the optimum receiver simplifies to that shown in Fig. 7.

Just as in the case of known signals in thermal noise, the bounds of the probability of error P_e can be expressed as in (116). The quantity C given by (117) is the capacity of the system in which signals are in coherent states. The average number p of signal photons per second equals S/T in this case, where

$$S = \nu s \tag{127}$$

is the total number of photons transmitted through ν diversity paths. The system reliability function E(R) is the solution of the maximization problem,

$$E(R) = \max_{0 \leq \delta \leq 1} \{e(\delta, s) - \delta R/C\}$$

$$e(\delta, s) = \left\{ (1+\delta) \ln \left\{ 1 + (\mathcal{N}+s) \left[1 - \left(\frac{\mathcal{N}}{1+\mathcal{N}} \; \frac{1+\mathcal{N}+s}{s+\mathcal{N}} \right)^{\frac{\delta}{1+\delta}} \right] \right\} - \delta \ln \left[1+s/(1+\mathcal{N}) \right] \right\} / s \ln (1+1/\mathcal{N}). \tag{128}$$

As in the case of coherent signals, the reliability function depends not only on the signal-to-noise ratio $s/\mathcal{N}$, but also on the average noise level $\mathcal{N}$ and the number of diversity path ν. The optimum reliability function $E^o(R)$ is obtained by maximizing the function E(R) in (128) with respect to ν, or alternatively, with respect to s. Let s^o denote the value of s that maximizes the function $e(\delta, s)$; then

$$E^o(R) = \max_{0 \leq \delta \leq 1} \{e(\delta, s^o) - \delta R/C\} \tag{129}$$

if the value of s^o does not exceed S. When s^o is larger than S, we have

$$E^o(R) = \max_{0 \leq \delta \leq 1} \{e(\delta, S) - \delta R/C\}. \tag{130}$$

This maximization problem has been solved numerically. The results are shown in Fig. 8, where the optimum average number s^o of signal photons per diversity path is plotted as a function of R/C for $\mathcal{N} = 0.1$ and $\mathcal{N} = 10$. Also shown in Fig. 8 is the value of s^o in the classical limit [70]. The values of $s^o/\mathcal{N}$ for rates R less than R_c^o are independent of R/C, but they are functions of the average noise level $\mathcal{N}$. If the effective noise in the system is taken to be $\left(\mathcal{N} + \frac{1}{2}\right)$, the optimum ratio $s^o/\left(\mathcal{N} + \frac{1}{2}\right)$ is roughly 3 for $R \leq R_c^o$ independent of the value of $\mathcal{N}$.

For rates greater than R_c^o, the value of s^o increases rapidly with R/C. That is, for a fixed value of S, the optimum number of equal-strength diversity paths decreases at information rates higher than R_c^o. From Fig. 8, it is clear that for increasing R/C, s^o increases without bound. Hence, when the average number of transmitted photons is fixed at S, a point where the value of s^o is equal to S will eventually be reached. That is, the optimum value of ν is equal to one. The rate at which $S = s^o$ is called the threshold rate for the given value of S.

Let us assume for the moment that for any given value of R/C, S is large so that S/s^o is larger than 1. In this limiting case, the optimum value of $E^o(R)$ as a function of R/C is given by (130). The general behavior of $E^o(R)$ at rates above the threshold is given by Fig. 9 for different thermal noise levels. The reliability function for the optimum classical fading channel is also shown for comparison.

V. QUANTUM ESTIMATION THEORY

The communication systems treated thus far transmit information digitally by encoding it into symbols from an alphabet of two or more elements; the receiver is designed to make decisions among two or more corresponding hypotheses. It is also possible to transmit information on a continuous basis, by encoding it into the amplitude of the signal (pulse-amplitude modulation), into its carrier frequency (pulse-frequency modulation), or into its epoch (pulse-time modulation). The receiver is then confronted with a task not of decision, but of estimation. The information to be extracted appears as a set of <u>parameters</u> $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots \theta_M)$ of the received field.

In studying classical communication systems, we assume that the receiving aperture field can be sampled in as much detail as we wish in order to generate a set $\underset{\sim}{x} = (x_1, x_2, \ldots x_n, \ldots)$ of data having a known joint probability density function $P(\underset{\sim}{x}, \underset{\sim}{\theta})$, which depends on the information-bearing parameters $\underset{\sim}{\theta}$. Estimation theory is applied to determine an estimator $\hat{\underset{\sim}{\theta}}(\underset{\sim}{x})$ of the parameters that minimizes some cost function $C(\hat{\underset{\sim}{\theta}}, \underset{\sim}{\theta})$, which measures the cost to the experimenter of assigning a set of estimates $\hat{\underset{\sim}{\theta}}$ to the parameters when their true values are $\underset{\sim}{\theta}$. Examples of commonly used cost functions are the mean-square error and the absolute error [7.1]. For the purpose of finding the best estimator, it is usually necessary to provide a prior joint probability density function $z(\underset{\sim}{\theta})$ of the parameters, which represents the distribution of relative frequencies with which $\underset{\sim}{\theta}$ will appear in certain ranges of values in the communication system envisioned. As an example of the prescriptions derived from estimation theory, we mention the maximum-likelihood estimator, which assigns as the estimate of a parameter θ that value for which $P(\underset{\sim}{x}, \theta)$ is maximum, the prior probability density function $z(\theta)$ being assumed very broad.

When the communication system is described quantum-mechanically, the received field is described by a density operator $\rho(\underset{\sim}{\theta})$, which is a function of the unknown parameters $\underset{\sim}{\theta}$. If the receiver measures observables corresponding to commuting Hermitian operators

$\underset{\sim}{X} = (X_1, X_2, \ldots X_L)$, the joint conditional probability density function of the outcomes $\underset{\sim}{x} = (x_1, x_2, \ldots x_L)$ is given by

$$P(\underset{\sim}{x}|\underset{\sim}{\theta}) = \langle \underset{\sim}{x}|\rho(\underset{\sim}{\theta})|\underset{\sim}{x}\rangle. \tag{131}$$

Again, $|\underset{\sim}{x}\rangle$ are the simultaneous eigenvectors of the operators $\underset{\sim}{X}$ corresponding to the eigenvalues $\underset{\sim}{x}$. It follows that the joint probability density function of the parameters $\underset{\sim}{\theta}$ and the observed data $\underset{\sim}{x}$ is just

$$P(\underset{\sim}{x}, \underset{\sim}{\theta}) = z(\underset{\sim}{\theta})\langle \underset{\sim}{x}|\rho(\underset{\sim}{\theta})|\underset{\sim}{x}\rangle. \tag{132}$$

The optimum estimation function $\hat{\underset{\sim}{\theta}}(\underset{\sim}{x})$, which assigns the data $\underset{\sim}{x}$ to an estimate $\hat{\underset{\sim}{\theta}}$ of the parameters, can be determined from classical estimation theory. Our problem is, therefore, that of finding the operators $\underset{\sim}{X}$ whose orthonormal set of eigenvectors $\{|\underset{\sim}{x}\rangle\}$ is such that the average cost

$$\overline{C} = \iint z(\underset{\sim}{\theta})\ C(\hat{\underset{\sim}{\theta}}(x), \underset{\sim}{\theta})\langle \underset{\sim}{x}|\rho(\underset{\sim}{\theta})|\underset{\sim}{x}\rangle\ d\underset{\sim}{\theta} d\underset{\sim}{x} \tag{133}$$

is minimum.

Alternatively, "The values of the parameters $\underset{\sim}{\theta}$ lie between $\underset{\sim}{\theta}'$ and $\underset{\sim}{\theta}' + \Delta\underset{\sim}{\theta}$"' is a proposition of the kind described in Section II [72]. If the range of possible values of $\underset{\sim}{\theta}$ is broken up into contiguous, but nonoverlapping intervals $\Delta\underset{\sim}{\theta}$, and if the entire array of corresponding propositions is tested, one of them must be declared true and the rest false. The result is an estimate of $\underset{\sim}{\theta}$ within an uncertainty $\Delta\underset{\sim}{\theta}$, which in principle can be made as small as we like. When the observables measured by the receiver correspond to commuting Hermitian operators, each such proposition is associated with a projection operator, which we denote $\Delta E(\underset{\sim}{\theta}')$ for the range $(\underset{\sim}{\theta}', \underset{\sim}{\theta}'+\Delta\underset{\sim}{\theta}')$. These projection operators commute and add up to the identity operator

$$\sum_{\underset{\sim}{\theta}'} E(\underset{\sim}{\theta}') = \underset{\sim}{1}. \tag{134}$$

Passing to the limit $\Delta\underset{\sim}{\theta}' \to 0$, we speak of a <u>resolution of the identity</u> $E(\underset{\sim}{\theta}')$, with

$$\int dE(\underset{\sim}{\theta}') = \underset{\sim}{1}, \tag{135}$$

and this resolution is in effect an estimator of the parameters $\underset{\sim}{\theta}$.

The probability, $\underset{\sim}{\theta}$ being the true value of the parameters, that the estimates lie between $\underset{\sim}{\theta}'$ and $\underset{\sim}{\theta}' + d\underset{\sim}{\theta}'$ is $\mathrm{Tr}\,[\rho(\underset{\sim}{\theta})dE(\underset{\sim}{\theta}')]$, and the average cost is therefore [73]

$$\overline{C} = \iint z(\underset{\sim}{\theta})\; C(\hat{\underset{\sim}{\theta}}, \underset{\sim}{\theta})\; \mathrm{Tr}\,[\rho(\underset{\sim}{\theta})\, dE(\underset{\sim}{\theta}')]\; d\underset{\sim}{\theta}. \tag{136}$$

This average cost $\overline{C}$ is to be minimized by choosing the resolution of the identity, $dE(\underset{\sim}{\theta}')$, over the entire range of possible values of $\underset{\sim}{\theta}'$.

When the receiver makes an ideal measurement of incompatible observables as discussed in Section II, its structure is specified by a complete set of measurement state vectors $\{|\underset{\sim}{x}\rangle\}$, when the outcome of the measurement is $\underset{\sim}{x}$. For a given set of measurement state vectors $\{|x\rangle\}$, the manner in which the receiver assigns to the data $\underset{\sim}{x}$ an estimate $\hat{\underset{\sim}{\theta}}(\underset{\sim}{x})$ is prescribed by classical estimation theory. Our problem is to find the set of measurement state vectors $\{|\underset{\sim}{x}\rangle\}$ to minimize the average cost function in (133). Since very little is known about the solution to this problem, we shall not be concerned with it hereafter.

In the following discussion, the cost function with which we shall be solely concerned is the mean-square error of the estimate

$$\epsilon = \sum_{i=1}^{M} \underset{\sim}{E}[(\hat{\underset{\sim}{\theta}}_i - \underset{\sim}{\theta}_i)^2]. \tag{137}$$

The specification of the measured compatible observables $\underset{\sim}{X}$ and the estimate functions $\hat{\underset{\sim}{\theta}}_i(\underset{\sim}{x})$ can be combined in the specification of quantum-mechanical estimators $\hat{\underset{\sim}{\Theta}} = (\hat{\Theta}_1, \hat{\Theta}_2, \ldots \hat{\Theta}_M)$ of the parameters $\theta_1, \theta_2, \ldots \theta_M$, where

$$\hat{\Theta}_i = \int \theta_i(\underset{\sim}{x}) \, |\underset{\sim}{x}\rangle \langle \underset{\sim}{x}| \, d\underset{\sim}{x}$$

$$= \int \theta_i' \, dE(\underset{\sim}{\theta}'). \tag{138}$$

These operators are Hermitian and commutative, $|\underset{\sim}{x}\rangle$ being the simultaneous eigenvectors of $\underset{\sim}{X}$. A measurement of $\hat{\Theta}_i$ yields an estimate $\hat{\theta}_i$ of θ_i. The mean-square error ϵ in (137) becomes

$$\epsilon(\underset{\sim}{\Theta}) = \sum_{i=1}^{M} \int z(\underset{\sim}{\theta}) \, \mathrm{Tr}\, [\rho(\underset{\sim}{\theta})(\hat{\Theta}_i - \theta_i \underset{\sim}{1})^2] \, d\underset{\sim}{\theta} \tag{139}$$

when $\underset{\sim}{\theta}$ is a set of random parameters with prior probability density function $z(\underset{\sim}{\theta})$. When $\underset{\sim}{\theta}$ is a set of unknown parameters whose prior probability density functions are unknown, estimators are sought that have small or zero bias and at the same time have a small mean-square error over a broad range of true values of the parameters. The bias of an estimator $\hat{\Theta}_i$ of a parameter θ_i is

$$\langle \hat{\theta}_i \rangle - \theta_i = \mathrm{Tr}\, [\hat{\Theta}_i \rho(\underset{\sim}{\theta})] - \theta_i. \tag{140}$$

The mean-square error is

$$\epsilon_i = \mathrm{Tr}\, [\rho(\underset{\sim}{\theta})(\hat{\Theta}_i - \langle \hat{\theta}_i \rangle \underset{\sim}{1})^2]. \tag{141}$$

The problem of finding commuting Hermitian quantum-mechanical estimators $\hat{\Theta}_i$ to minimize the mean-square error $\epsilon(\underset{\sim}{\Theta})$ has not been solved for density operators $\rho(\underset{\sim}{\theta})$ that do not commute for different values of the parameters $\underset{\sim}{\theta}$. When the density operators commute, they possess a common set of orthonormal eigenvectors, $|\phi_n\rangle$. In other words,

$$\rho(\underset{\sim}{\theta}) = \sum_n P_n(\underset{\sim}{\theta}) \, |\phi_n\rangle \langle \phi_n| .$$

The quantum-mechanical estimators $\hat{\Theta}_1, \hat{\Theta}_2, \ldots \hat{\Theta}_M$ that minimize the mean-square error in (139) and (141) can be written

$$\hat{\Theta}_i = \sum_{n} \hat{\theta}_i(n) \, |\phi_n\rangle \langle \phi_n| . \tag{142}$$

$\hat{\theta}_i(n)$ is just the conditional mean

$$\hat{\theta}_j(n) = \int \theta_i z(\underset{\sim}{\theta}) \, P_n(\underset{\sim}{\theta}) \, d\underset{\sim}{\theta} \Big/ \left\{ \int z(\underset{\sim}{\theta}') \, P_n(\underset{\sim}{\theta}') \, d\underset{\sim}{\theta}' \right\} \tag{143}$$

when the parameters are random variables.

Suppose that we wish to estimate the signal strengths of different modes of the received field described by a density operator

$$\rho(\underset{\sim}{\theta}) = \pi^{-\nu} \int \prod_{k=1}^{\nu} \exp[-|\alpha_k|^2 / (\mathscr{N} + s_k)] \, |\alpha_k\rangle \langle \alpha_k| \, d^2\alpha_k / (\mathscr{N} + s_k).$$

The parameters $\underset{\sim}{\theta}$ are identified with the average numbers s_k of signal photons in the different modes. As discussed in Section IV, s_k can be considered as the signal strength in the k^{th} transmission through a Rayleigh fading channel in a diversity system. Since for all values of s_k, the eigenvectors of $\rho(\underset{\sim}{\theta})$ are $|n_1, n_2, \ldots n_\nu\rangle$, the simultaneous eigenvectors of the number operators $a_k^+ a_k$, it suffices to measure the operators $a_k^+ a_k$. The outcome $\underset{\sim}{n} = (n_1, n_2, \ldots n_\nu)$ of this measurement is a sufficient statistic [74] whose joint probability is

$$P(\underset{\sim}{n}, \{s_k\}) = \prod_{j=1} (1 - v_j) \, v_j^{n_j}$$

$$v_j = (\mathscr{N} + s_k) / (1 + \mathscr{N} + s_k). \tag{144}$$

In this equation, $\mathscr{N}$ is the average number of noise photons. The outcome of a measurement of the operator $a_k^+ a_k$ is an unbiased estimate of the sum $\mathscr{N} + s_k$. By subtracting the known value of $\mathscr{N}$, an unbiased estimate of s_k is obtained, which happens to have the smallest possible mean-square error.

It has been shown [75] that the quantum-mechanical estimator $\hat{\Theta}$ of a single random parameter θ with a prior probability density function $z(\theta)$

that minimizes the mean-square error, satisfies the operator equation

$$\Gamma\hat{\Theta} + \hat{\Theta}\Gamma = 2\eta. \tag{145}$$

when the operator

$$\Gamma = \int z(\theta)\, \rho(\theta)\, d\theta \tag{146}$$

is positive definite, and $\eta = \int \theta\rho(\theta)\, z(\theta)\, d\theta$. Moreover, the optimum estimator $\hat{\Theta}$ is uniquely given by

$$\hat{\Theta} = 2 \int_0^\infty e^{-\Gamma\alpha}\, \eta\, e^{\Gamma\alpha}\, d\alpha. \tag{147}$$

This result can be used to find the optimum quantum-mechanical mean-square-error estimate of the complex plane-wave envelope m(t) in double-sideband modulation [76]. Since m(t) can be represented by its Karhunen-Loève expansion

$$m(t) = \sum_k m_k s_k(t), \tag{148}$$

the problem of waveform estimation is reduced to the estimation of the random parameters m_k. When the bandwidth of m(t) is very small compared with the carrier frequency ω, the density operator $\rho(\{m_k\})$ describing the state of the received fields is

$$\rho(\{m_k\}) = \int \exp\left[-\sum_k \left|\beta_k - (2\epsilon_o V/\hbar\omega)^{1/2} m_k\right|^2 \Big/ \mathcal{N}\right] \prod_k |\beta_k\rangle\langle\beta_k|\, d^2\beta_k/\pi\mathcal{N};$$

where the vectors $|\beta_k\rangle$ are the right eigenvectors of the annihilation operators b_k associated with the modes with normal-mode functions $s_k(t)$. When the prior joint probability density function of the parameters m_k is

$$P(\{m_k\}) = \prod_k (2\pi\lambda_k)^{-1/2} \exp[-m_k^2/2\lambda_k],$$

the minimum mean-square-error estimator $\hat{\Theta}_k$ of m_k can be found from (147),

$$\hat{\Theta}_k = \left(b_k^+ + b_k\right) \lambda_k/2x \left(\tfrac{1}{2}\mathcal{N} + \lambda_k x^{-2} + \tfrac{1}{4}\right)$$

$$x = (\hbar\omega/2\epsilon V)^{1/2}$$

The mean-square error associated with the optimum extimator $\hat{\Theta}_k$ is

$$\epsilon_k = \underset{\sim}{E}[(\hat{\Theta}_k - m_k)^2] = \lambda_k\left(\mathcal{N} + \tfrac{1}{2}\right) \Big/ \left(2\lambda_k x^{-2} + \mathcal{N} + \tfrac{1}{2}\right)$$

which is the same as the classical estimation error in white noise, whose spectral density is $\left(\mathcal{N} + \frac{1}{2}\right)\hbar\omega$ [77].

Quantum-Mechanical Cramér-Rao Inequality

a. Single Unknown Parameter

Although the best estimator of a parameter θ, given the density operator $\rho(\theta)$, has not been found in general, a lower bound can be set to the mean-square error attainable by any estimator. It is the quantum-mechanical counterpart of the Cramér-Rao inequality of classical statistics [78, 79].

Let $\hat{\Theta}$ be an operator whose measurement yields an estimate $\hat{\theta}$ of the parameter θ. The expected value of the estimate $\hat{\theta}$, when the true value of the parameter is θ,

$$\underset{\sim}{E}[\hat{\theta}] = \mathrm{Tr}\,[\rho(\theta)\hat{\Theta}] = \langle\hat{\theta}\rangle, \tag{149}$$

and the mean-square error is

$$\epsilon = \underset{\sim}{E}[(\hat{\theta} - \langle\hat{\theta}\rangle)^2] = \mathrm{Tr}\,[\rho(\theta)(\hat{\Theta} - \langle\hat{\theta}\rangle)^2]. \tag{150}$$

If $\langle\hat{0}\rangle = \theta$, the estimate is said to be unbiased.

According to the quantum-mechanical form of the Cramér-Rao inequality ϵ cannot be smaller than [17, 80]

$$\epsilon \geq [\partial\langle \hat{\theta}\rangle/\partial\theta]^2 / [\mathrm{Tr}[\rho L^2]]. \tag{151}$$

where L is the symmetrizedlogarithmic derivative of $\rho(\theta)$ with respect to θ, defined by

$$\frac{\partial\rho(\theta)}{\partial\theta} = \frac{1}{2}(\rho L + L\rho). \tag{152}$$

Furthermore, the lower bound can be attained if and only if the symmetrized logarithmic derivative L has the form

$$L = k(\theta)(\hat{\Theta} - \theta), \tag{153}$$

where $k(\theta)$ is a function only of the true value θ of the parameter. The estimate $\hat{\theta}$ is then unbiased, for by (149) and (150), $k(\theta)(\langle\hat{\theta}\rangle - \theta) = \mathrm{Tr}(\rho L) = \mathrm{Tr}[\partial\rho(\theta)/\partial\theta] = \partial[\mathrm{Tr}\,\rho(\theta)]/\partial\theta = 0$, since $\mathrm{Tr}\,\rho(\theta) = 1$ independently of θ, whereupon $\langle\hat{\theta}\rangle = \theta$. The first factor $(\partial\langle\theta\rangle/\partial\theta)^2$ in (151) then equals 1, and we find that the mean-square error attains the minimum possible value

$$\epsilon_{min} = |k(\theta)|^{-1} \tag{154}$$

An example in which the lower bound is attained is the estimate of the amplitude A of a coherent signal in a single mode, corresponding to the state $|A\mu\rangle$, where μ is a known complex number, $|\mu| = 1$. Here $A^2 = N_s$ is the average number of signal photons in the mode. The noise is of the thermal variety, and the density operator $\rho(A)$ takes the form

$$\rho(A) = (\pi\mathcal{N})^{-1} \int \exp[-|\alpha - A\mu|^2/\mathcal{N}]\,|\alpha\rangle\langle\alpha|\, d^2\alpha, \tag{155}$$

where $\mathcal{N}$ is the average number of thermal photons per mode. The symmetrized logarithmic derivative L can be shown to be

$$L = 4(\mathscr{A} - A)/(2\mathcal{N} + 1), \tag{156}$$

with

$$\mathscr{A} = \frac{1}{2}(\mu a^{+} + \mu^{*} a) \tag{157}$$

an unbiased estimator of the amplitude A in terms of the annihilation and creation operators of the mode. The minimum mean-square error of this estimate, by (153) and (154), is

$$\epsilon_{min} = \frac{1}{4}(2\mathscr{N} + 1),$$

and the relative mean-square error is

$$\epsilon_{min}/A^2 = (2\mathscr{N} + 1)/4N_s = D^{-2}, \tag{158}$$

where D is the signal-to-noise ratio defined in (101) for detection of a coherent signal in thermal noise. The estimating operator $\mathscr{A}$ is proportional to the quasi-classical detection statistic described in Section IV [81].

b. Single Random Parameter

Let $\hat{\Theta}$ be the quantum-mechanical estimator of a random parameter θ with a prior probability density function $z(\theta)$. The quantum-mechanical form of the Cramér-Rao inequality is [82]

$$\epsilon = \int z(\theta)\, \mathrm{Tr}\left[(\Theta - \theta \underset{\sim}{1})^2 \rho(\theta)\right] d\theta$$

$$\geq R\ell \left\{ \int z(\theta)\, \mathrm{Tr}\left[\rho(\theta)\left[L(\theta) + \frac{d}{d\theta} \ln z(\theta)\right]^2\right] d\theta, \right.$$

where $L(\theta)$ is defined in (152). Equality holds if and only if

$$\hat{\Theta} - \theta \underset{\sim}{1} = k\left[\frac{d}{d\theta} \ln p(\theta) + L(\theta)\right.$$

for some constant k.

Footnotes

1. A note of caution is advisable here. While it is true that all hypotheses relating to the values of compatible observables are evaluated based on the outcomes of measurements of projection operators such as π_j, it is not obvious that all projection operators can be measured physically. Equivalently, while all observables correspond to Hermitian operators, it is not clear that all Hermitian operators with complete sets of eigenvectors correspond to observables, which in principle are measurable. This problem, of obvious practical importance, is beyond the scope of this paper, and we do in fact assume that any Hermitian operator that has a complete set of eigenvectors is an observable.

2. The channel capacity is the maximum rate at which the error probability P_e can be made arbitrarily small when constrained in signal power. Operation at a rate higher than capacity condemns the system to a high probability of error, regardless of the choice of signals and receiver.

References

[1] D. Middleton, An Introduction to Statistical Communication Theory, McGraw-Hill, New York, 1960.

[2] J. C. Hancock and P. A. Wintz, Signal Detection Theory, McGraw-Hill, New York, 1966.

[3] C. W. Helstrom, Statistical Theory of Signal Detection, Pergamon Press, Oxford, 2d edition, 1968.

[4] H. L. Van Trees, Detection, Estimation and Modulation Theory, vol. 1, John Wiley and Sons, Inc., New York, 1968.

[5] W. B. Davenport, Jr. and W. L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill, New York, 1958, pp. 185-189.

[6] C. W. Helstrom, op. cit., pp. 56-65.

[7] T. E. Stern, "Some Quantum Effects in Information Channels," IEEE Trans. on Information Theory, vol. IT-6, pp. 435-440, September, 1960.

[8] T. E. Stern, "Information Rates in Photon Channels and Photon Amplifiers," 1960 IRE International Convention Record, Pt. 4, pp. 182-188.

[9] J. P. Gordon, "Quantum Effects in Communications Systems," Proc. IRE, vol. 50, no. 9, pp. 1898-1908, September, 1962.

[10] T. Hagfors, "Information Capacity and Quantum Effects in Propagation Circuits," Technical Report 344, Lincoln Laboratory, M.I.T., Cambridge, Mass., January 24, 1964.

[11] C. Y. She, "Quantum Electrodynamics of a Communication Channel," IEEE Trans. on Information Theory, vol. IT-14, pp. 32-37, January, 1968.

[12] H. Takahasi, Advances in Communication Systems, vol. 1, A. V. Balakrishnan (ed.), Academic Press, New York, 1965, p. 227.

[13] G. Fillmore and G. Lachs, "Information Rates for Photocount Detection Systems," IEEE Trans. on Information Theory, vol. IT-15, pp. 465-468, July, 1969.

[14] 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, 1965.

[15] C. W. Helstrom, "Detection Theory and Quantum Mechanics," Inform. Contr., vol. 10, no. 3, pp. 254-291, March, 1968.

[16] C. W. Helstrom, "Detection Theory and Quantum Mechanics (II)," Inform. Contr., vol. 13, no. 2, pp. 156-171, August, 1968.

[17] C. W. Helstrom, "The Minimum Variance of Estimates in Quantum Signal Detection," Trans. IEEE, vol. IT-14, pp. 234-242, March, 1968.

[18] C. W. Helstrom, "Fundamental Limitations on the Detectability of Electromagnetic Signals," Int. J. Theoret. Phys., vol. 1, pp. 37-50, May, 1968.

[19] P. A. M. Dirac, Quantum Mechanics, Oxford University Press, London, 3rd edition, 1947.

[20] W. H. Louisell, Radiation and Noise in Quantum Electronics, McGraw-Hill, New York, 1964, Chap. 1, pp. 1-68.

[21] P. A. M. Dirac, op. cit., §5, pp. 14-18.

[22] P. R. Halmos, Finite-Dimensional Vector Space, D. Van Nostrand Company, Inc., Princeton, New Jersey, 1958, §7, pp. 10-11.

[23] P. A. M. Dirac, op. cit., Chap. III, pp. 53-83.

[24] Ibid., §12, pp. 45-48.

[25] Ibid., §13, pp. 49-52.

[26] W. H. Louisell, op. cit., Chap. I, pp. 47-53; P. A. M. Dirac, op. cit., Chap. IV, pp. 84-107.

[27] U. Fano, "Description of States in Quantum Mechanics by Density Matrix and Operator Techniques," Rev. Mod. Phys., vol. 29, pp. 74-93, January, 1957.

[28] W. H. Louisell, op. cit., Chap. 6, pp. 220-252.

[29] P. A. M. Dirac, op. cit., §27, pp. 108-111; W. H. Louisell, op. cit., §1.14 and 1.15, pp. 53-56.

[30] Ibid., §28, pp. 111-118; ibid., §1.16, pp. 57-61.

[31] R. Serber and C. H. Townes, "Limits in Electromagnetic Amplifications Due to Complementarity," Quantum Electronics, C. H. Townes (ed.), Columbia University Press, New York, 1960, pp. 233-255.

[32] J. P. Gordon and W. H. Louisell, "Simultaneous Measurements of Noncommuting Observables," in Physics of Quantum Electronics, P. L. Kelley, B. Lax, and P. E. Tannenwald (eds.), McGraw-Hill, New York, 1966, pp. 833-840.

[33] S. D. Personick, "Efficient Analog Communication over Quantum Channels," Technical Report 477, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., Sec. IV.

[34] W. H. Louisell, op. cit., §4.4, pp. 148-156.

[35] Ibid., Chap. 4, pp. 138-174.

[36] P. A. M. Dirac, op. cit., §34, pp. 136-139; W. H. Louisell, op. cit., pp. 71-85.

[37] R. J. Glauber, "Coherent and Incoherent States of the Radiation Field," Phys. Rev., vol. 131, pp. 2766-2788, September 15, 1963.

[38] Ibid., Eq. (3.7), p. 2769.

[39] Ibid., §IX, pp. 2781-2784.

40] R. J. Glauber, Quantum Optics and Electronics, Lecture Notes at Les Houches 1964 Session of the Summer School of Theoretical Physics, University of Grenoble, France.

41] R. J. Glauber, Phys. Rev., vol. 131, Eq. (3.32), p. 2771.

[42] W. H. Louisell, op. cit., §6.6, pp. 228-233.

[43] R. J. Glauber, Phys. Rev., vol. 131, §VIII, pp. 2779-2781.

[44] Equation (61) was obtained by applying Kummer's transformation to a form derived by R. Yoshitani (UCLA thesis, unpublished).

[45] Jane W. S. Liu, "Reliability of Quantum-Mechanical Communication Systems," Technical Report 468, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., December 31, 1968, Appendix A, pp. 78-81; also IEEE Trans. on Information Theory, vol. IT-16, May, 1970.

[46] J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering, John Wiley and Sons, Inc., New York, London, Sydney, 1965, pp. 220-222.

[47] A. S. Kuriksha, "Optimum Reception of Quantized Signals," Radio Engrg. and Electronic Phys., vol. 13, no. 10, pp. 1567-1575, October, 1968.

[48] C. W. Helstrom, "Modal Decomposition of Aperture Fields in Detection and Estimation of Incoherent Objects," J. Opt. Soc. Am., vol. 60, April, 1970.

[49] J. M. Wozencraft and I. M. Jacobs, op. cit., pp. 212-214.

[50] H. P. H. Yuen, "Theory of Quantum Signal Detection," Quarterly Progress Report No. 96, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., January 15, 1970, pp. 154-157.

[51] H. P. H. Yuen and M. Lax, "Theory of Quantum Signal Detection," submitted as a short paper to be presented at the Symposium on Information Theory, 1970.

[52] H. P. H. Yuen, private communications, 1969.

[53] Jane W. S. Liu, op. cit., pp. 17-20.

[54] C. W. Helstrom, "Quantum Detection and Estimation Theory," J. Statist. Phys., vol. 1, no. 2, pp. 231-252, 1969.

[55] P. A. Bakut, S. S. Shchurov, "Optimum Reception of a Quantum Signal," Problemy Pèredachi Informatsii, vol. 4, no. 1, pp. 77-82, 1968 (in Russian).

[56] R. J. Glauber, Phys. Rev., vol. 131, Eq. (3.29), p. 2771.

[57] C. W. Helstrom, Int. J. Theoret. Phys., vol. 1, §5, pp. 47-50.

[58] C. W. Helstrom, "Performance of an Ideal Quantum Receiver of a Coherent Signal of Random Phase," Trans. IEEE, vol. AES-5, pp. 562-564, May, 1969.

[59] H. L. Landau and H. O. Pollak, "Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty-III: The Dimension of the Space and Essentially Time- and Band-Limited Signals," Bell System Tech. J., vol. 41, pp. 1295-1336, July, 1962.

[60] Jane W. S. Liu, _op. cit._, Appendix F, pp. 96-98.

[61] _Ibid._, pp. 37-43.

[62] _Ibid._, Appendix G, pp. 99-102.

[63] _Ibid._, pp. 31-37; 91-95.

[64] J. M. Wozencraft and I. M. Jacobs, _op. cit._, (5.95b), p. 342.

[65] J. P. Gordon, "Information Capacity of Communications Channel in the Presence of Quantum Effects," _Advances in Quantum Electronics_, J. K. Singer (ed.), Columbia University Press, New York, 1961, pp. 509-519.

[66] J. M. Wozencraft and I. M. Jacobs, _op. cit._, §7.4, pp. 527-550.

[67] R. S. Kennedy, _Performance Limitations of Fading Dispersive Channels_, John Wiley and Sons, Inc., New York, London, Sydney, 1969, pp. 109-141.

[68] Jane W. S. Liu, _op. cit._, pp. 49-55.

[69] _Ibid._, pp. 55-75.

[70] R. S. Kennedy, _op. cit._, p. 123.

[71] C. W. Helstrom, Statistical Theory of Signal Detection, op. cit., Chap. VIII, pp. 249-289; H. L. Van Trees, op. cit., vol. 1, §2.4, pp. 52-86.

[72] J. von Neumann, Mathematical Foundations of Quantum Mechanics, R. T. Beyer (trans.), Princeton University Press, Princeton, N.J., 1955, see Chap. III, §1, pp. 196-200.

[73] C. W. Helstrom, Inform. Contr., vol. 10, op. cit., p. 265.

[74] C. W. Helstrom, Statistical Theory of Signal Detection, op. cit., p. 260.

[75] S. D. Personick, op. cit., Sec. 2.2.1.

[76] Ibid., pp. 71-75.

[77] H. L. Van Trees, op. cit., pp. 423-433.

[78] C. R. Rao, "Information and the Accuracy Attainable in the Estimation of Statistical Parameters," Bull. Calcutta Math. Soc., Vol. 37, pp. 81-91, 1945.

[79] H. Cramér, Mathematical Methods of Statistics, Princeton University Press, Princeton, N.J., 1946, pp. 473 ff.

[80] C. W. Helstrom, "Minimum Mean-Squared Error of Estimates in Quantum Statistics," Phys. Letters, vol. 25A, pp. 101-102, July 31, 1967; in (2) of this paper, $Tr(\rho L)^2$ should read $Tr(\rho L^2)$.

[81] C. W. Helstrom, Inform. Contr., vol. 13, §11, pp. 166-169.

[82] S. D. Personick, op. cit., Sec. 2.2.2.

Figure Captions

Fig. 1. Quantum-mechanical model of communication systems.

Fig. 2. Probability P_e of error in detection of known signal with prior probability $\zeta = 1/2$. D = signal-to-noise ratio $= [4N_s/(2\mathcal{N}+1)]^{1/2}$, N_s = average number of signal photons, $\mathcal{N}$ = average number of thermal photons.

Fig. 3. Average probability P_e of error vs average number S of signal photons, ideal quantum receiver of a coherent signal of random phase.

Fig. 4. Average probability P_e of error vs signal-to-noise ratio S/N, ideal quantum receiver of a coherent signal of random phase.

Fig. 5. Quantum-mechanical optimum receiver for narrow-band orthogonal signals with known classical amplitudes but random phases.

Fig. 6. System reliability function for orthogonal coherent signals.

Fig. 7. Optimum receiver for equal-strength orthogonal signals in a Rayleigh fading channel.

Fig. 8. Optimum number s^o of signal photons per diversity path/average number $\mathcal{N}$ of thermal noise photons vs R/C. (Classical limit taken from [67].)

Fig. 9. Optimum system reliability function for equal-strength orthogonal signals in Rayleigh fading channel. (Classical limit taken from [67].)

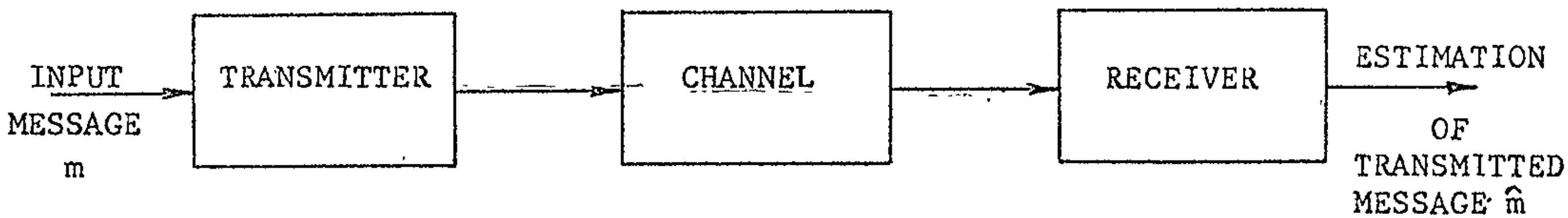
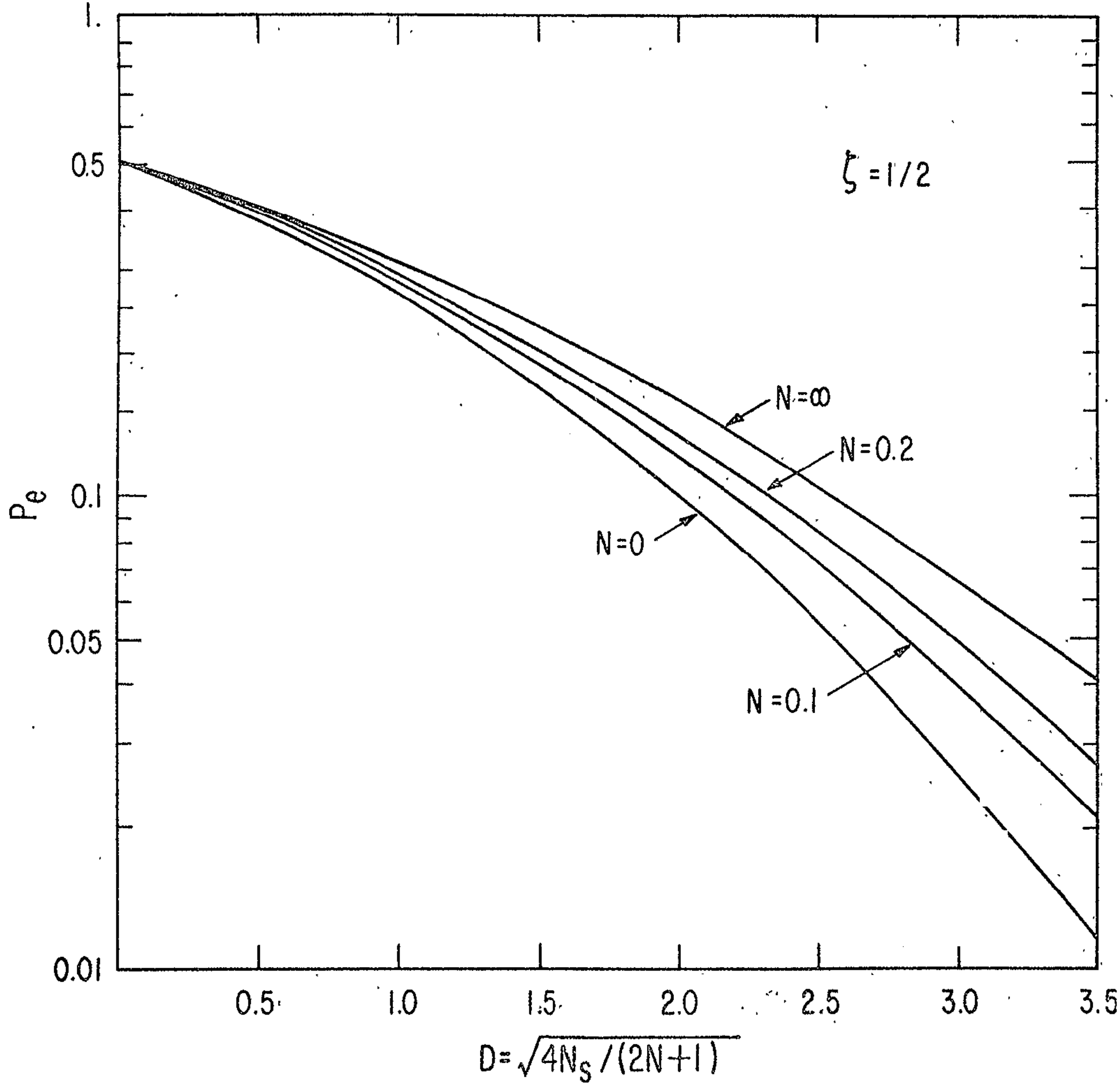

INPUT
MESSAGE
m
TRANSMITTER
CHANNEL
RECEIVER
ESTIMATION
OF
TRANSMITTED
MESSAGE $\hat{m}$

Figure 1



1.
0.5
0.1
0.05
0.01
P_e
ζ = 1/2
N = ∞
N = 0.2
N = 0
N = 0.1
0.5
1.0
1.5
2.0
2.5
3.0
3.5
D = $\sqrt{4N_S/(2N+1)}$

Figure 2

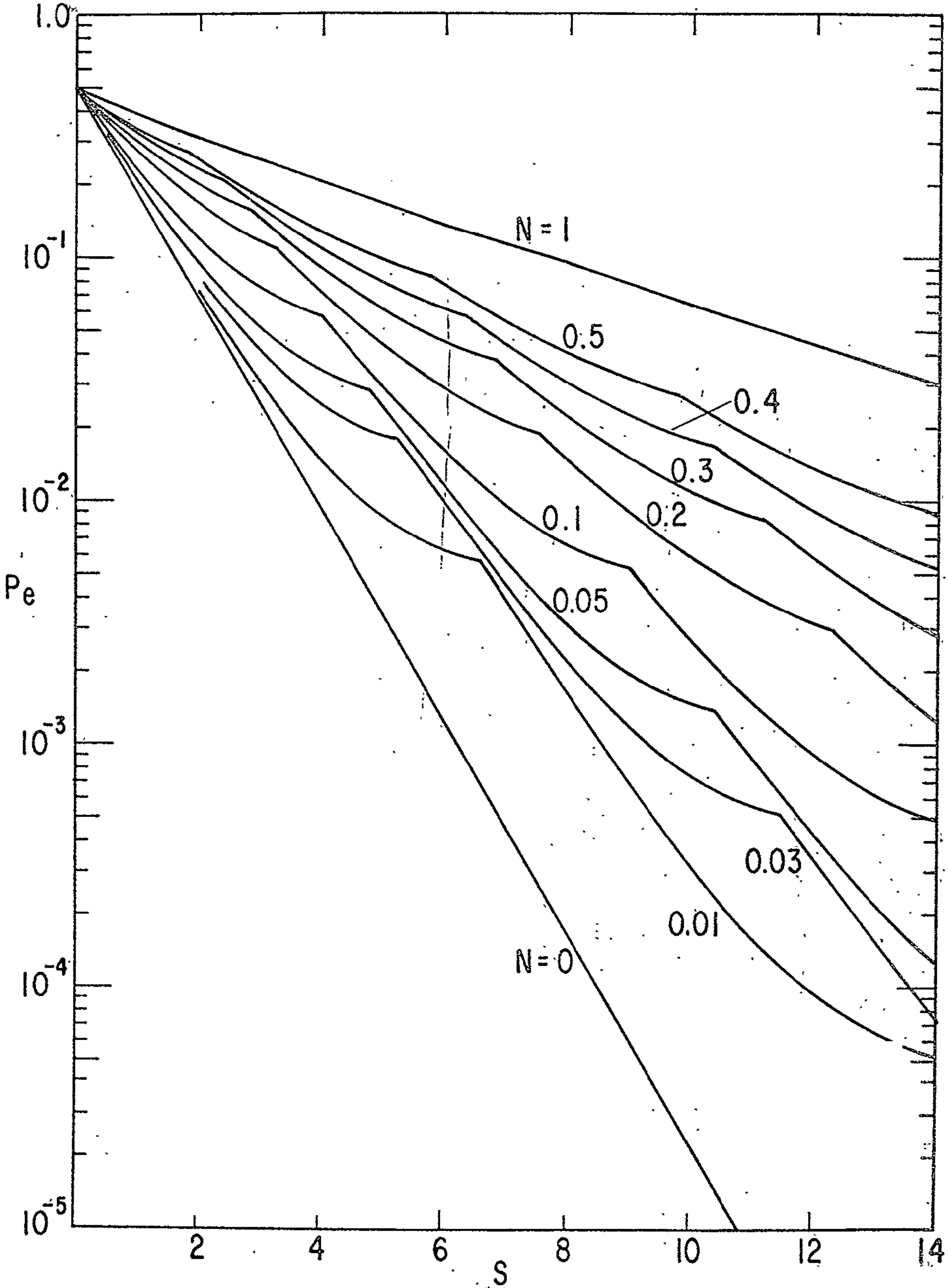

1.0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
P_e
2
4
6
8
10
12
14
S
N = 1
0.5
0.4
0.3
0.2
0.1
0.05
0.03
0.01
N = 0

Figure 3

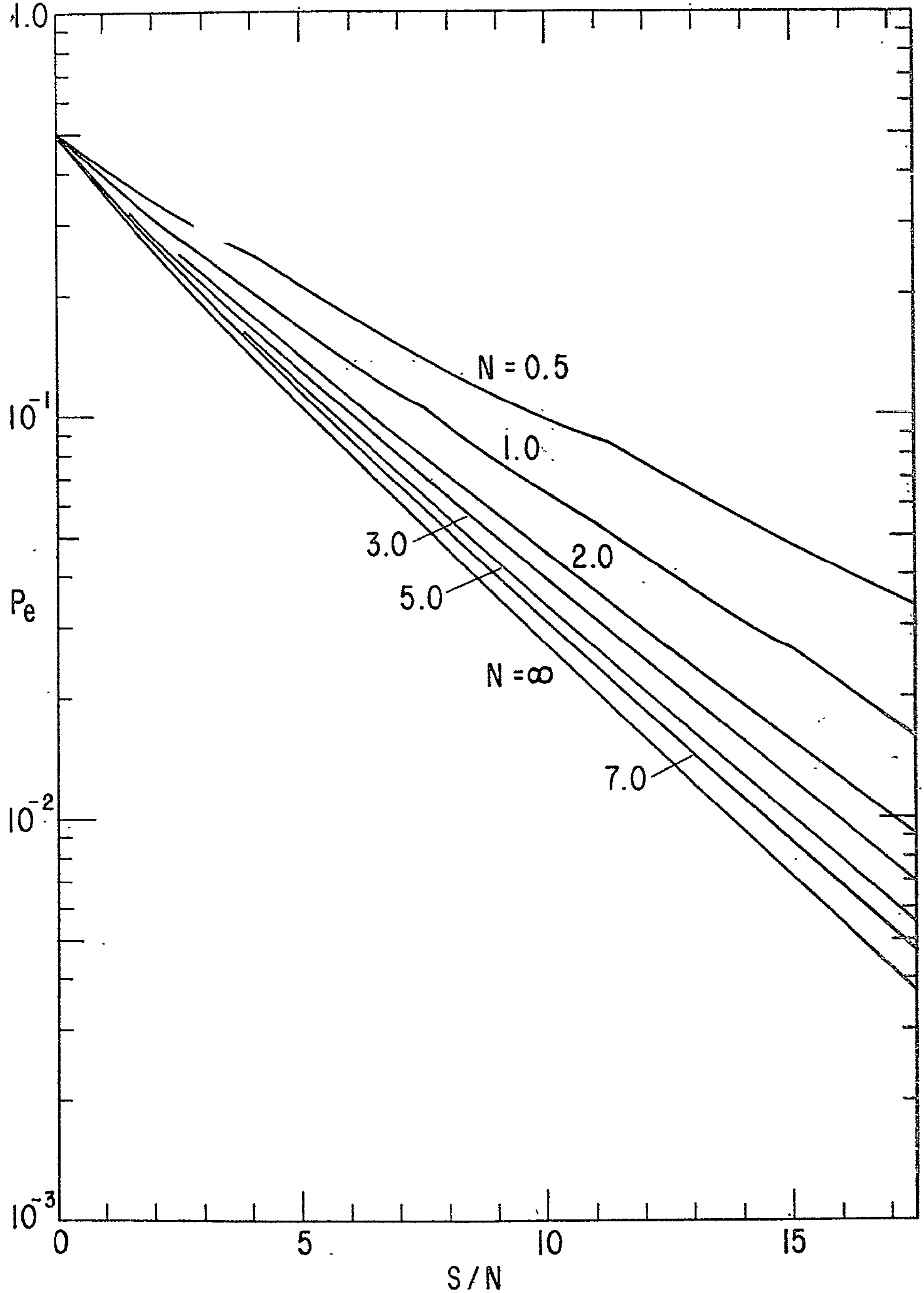
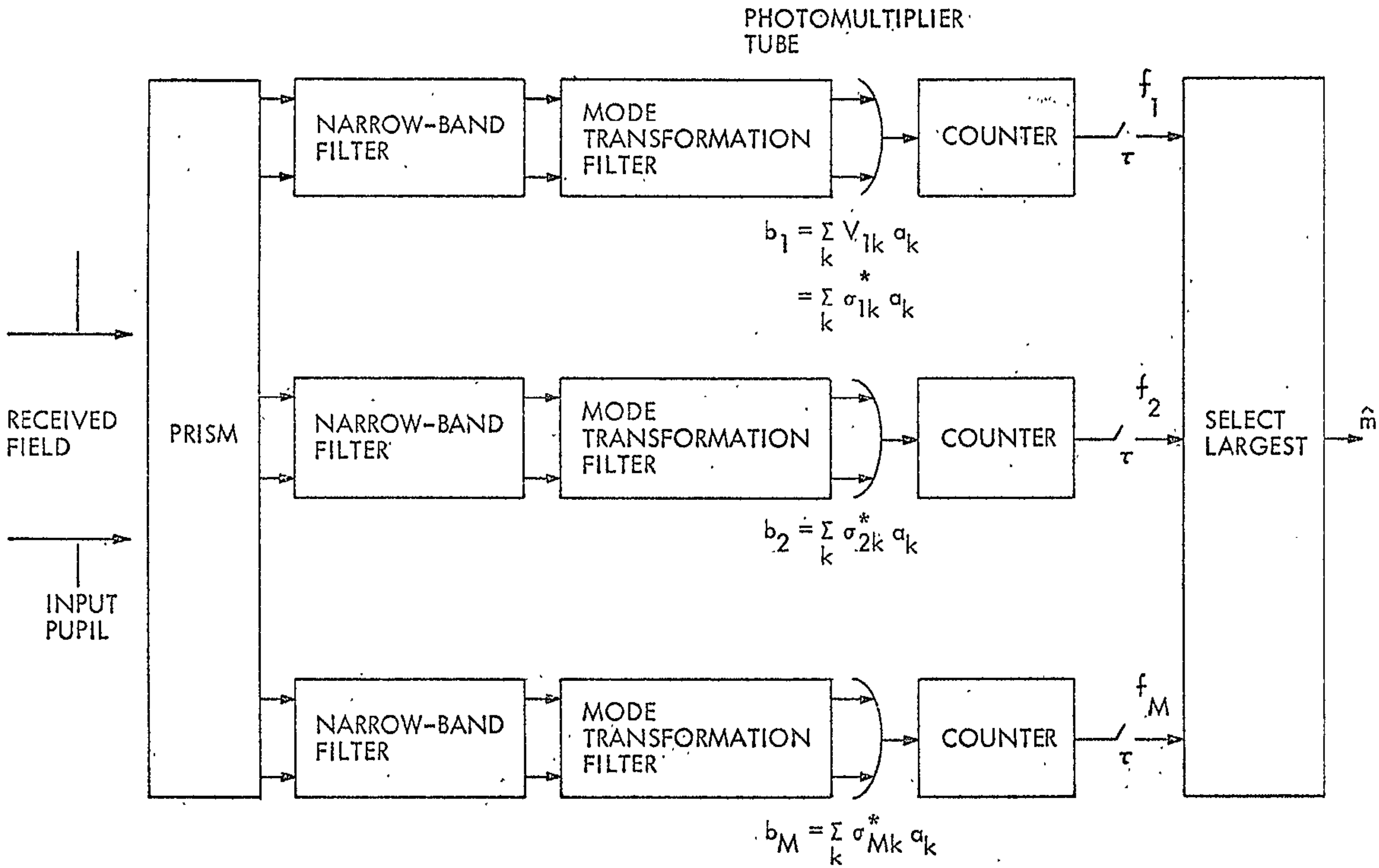


Figure 4



$$f_j = \sum_{r=0}^{j} \binom{n_j}{r} \frac{1}{r!} \left[|\mu_j|^2 / \mathcal{N}(1+\mathcal{N}) \right] \exp\left[-|\mu_j|^2 / (1+\mathcal{N}) \right]$$

Figure 5

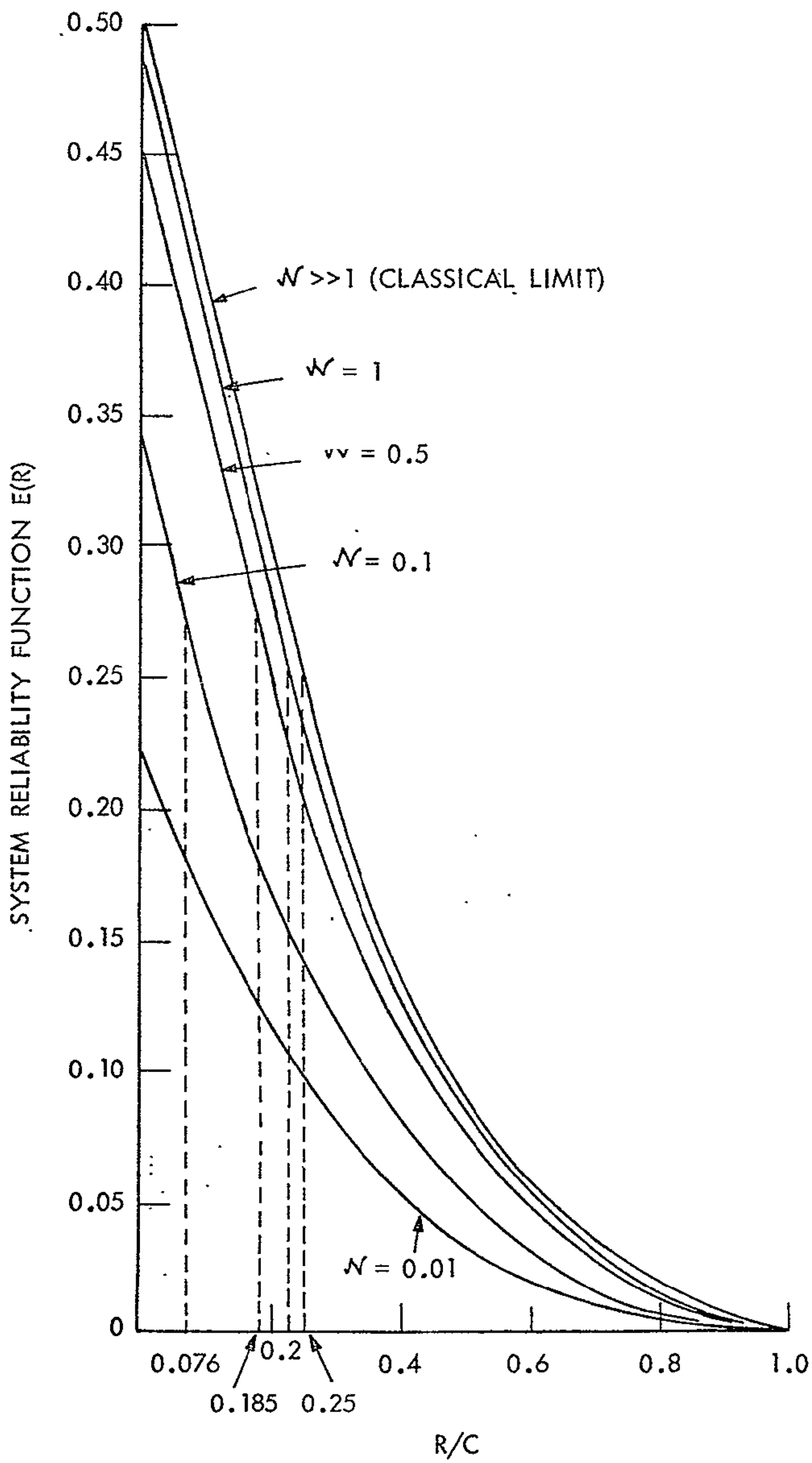

0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
SYSTEM RELIABILITY FUNCTION E(R)
𝒩 >>1 (CLASSICAL LIMIT)
𝒩 = 1
𝒩 = 0.5
𝒩 = 0.1
𝒩 = 0.01
0.076
0.2
0.185
0.25
0.4
0.6
0.8
1.0
R/C

Figure 6

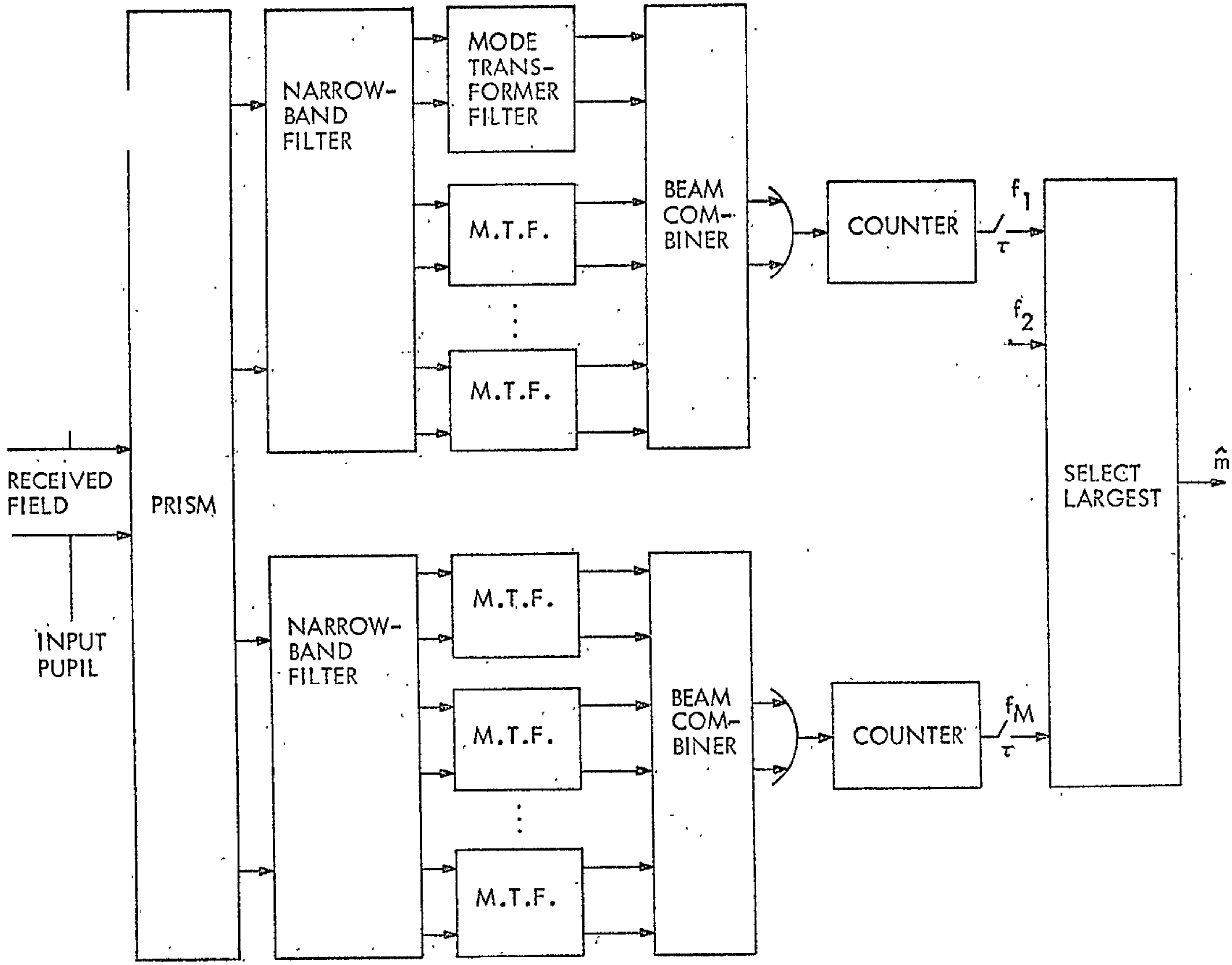

RECEIVED
FIELD
INPUT
PUPIL
PRISM
NARROW-
BAND
FILTER
MODE
TRANS-
FORMER
FILTER
M.T.F.
M.T.F.
BEAM
COM-
BINER
COUNTER
f_1
τ
f_2
NARROW-
BAND
FILTER
M.T.F.
M.T.F.
M.T.F.
BEAM
COM-
BINER
COUNTER
f_M
τ
SELECT
LARGEST
$\hat{m}$

Figure 7

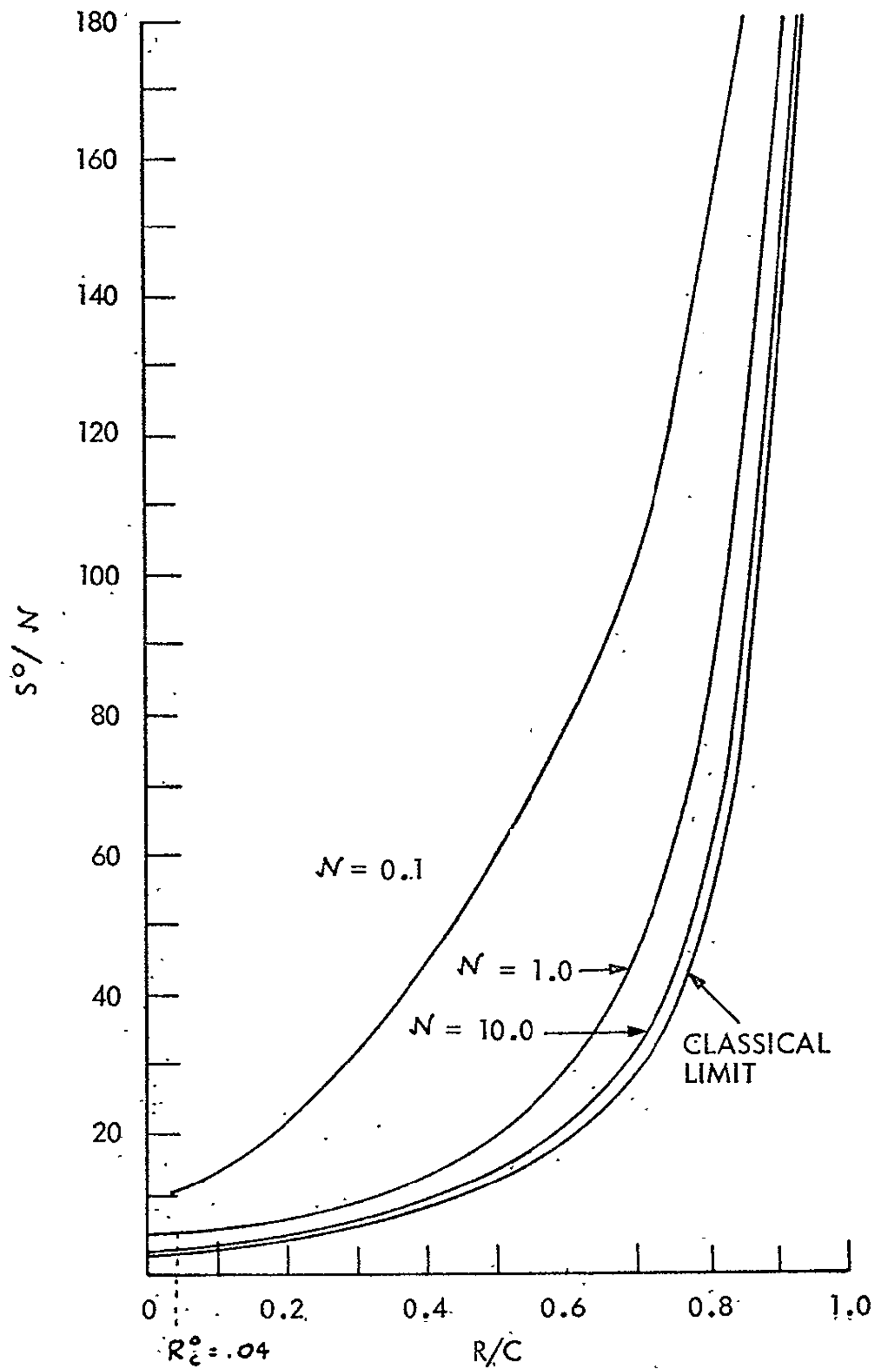

180
160
140
120
100
80
60
40
20
N/S_o
N = 0.1
N = 1.0
N = 10.0
CLASSICAL
LIMIT
0
0.2
0.4
0.6
0.8
1.0
R_c^o = .04
R/C

Figure 8

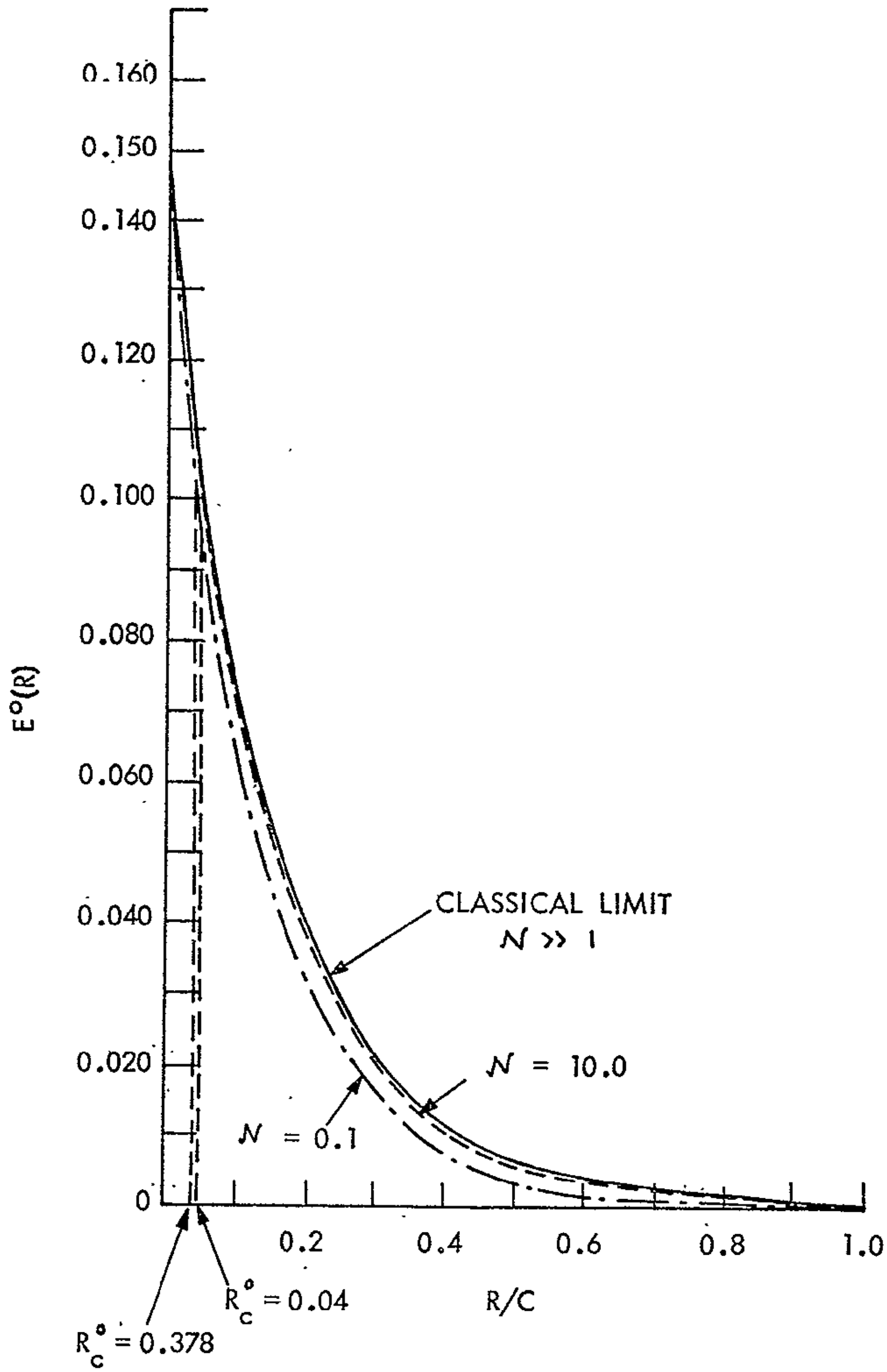

0.160
0.150
0.140
0.120
0.100
0.080
0.060
0.040
0.020
0
E°(R)
CLASSICAL LIMIT
𝒩 >> 1
𝒩 = 10.0
𝒩 = 0.1
0.2
0.4
0.6
0.8
1.0
R/C
$R_c^o = 0.04$
$R_c^o = 0.378$

Figure 9