

Quantum Theory of Partial Coherence*

E. C. G. SUDARSHAN

Department of Physics, Syracuse University, Syracuse, U. S. A.

ABSTRACT

The quantum theory of partial coherence is systematically developed. It is shown that the analytic coherence functions are completely determined by their boundary values for equal times; this result is valid both in classical theory and in quantum theory of coherence and for all coherence functions. The properties of reduced coherence functions of second order are used to decompose an arbitrary second order coherence function as a sum of coherence functions with unimodular reduced coherence functions. All results of the classical theory of second order partial coherence are recovered in the quantum theory formulation.

The higher order coherence functions and their relation to intensity correlations are discussed and the question of measurement of the coherence functions in terms of intensity correlation measurements is considered. The coherence indices of arbitrary order are defined as suitable normalized coherence functions. The coherence indices so defined are in general complex and of absolute value less than unity; the special case of unimodular coherence indices is studied in detail. It is found that the necessary and sufficient condition for the coherence index to be unimodular is that the corresponding coherence functions factorize. When the coherence functions of order (μ, ν) factorize so do all other coherence functions of higher order; more precisely, if the coherence function of order (μ', ν') factorizes of either μ' or ν' is larger than both μ and ν . Hence, except for a few lower order functions possibly, all the coherence functions factorize: the general density matrix compatible with this behaviour is obtained. Essentially all the excitation is in a single mode; the limiting case of "mode-pure" fields with excitation in only one mode is studied in detail. Several inequalities that have to be satisfied by the higher order "moments" defining a mode-pure fields are presented.

The method of higher order coherence functions is not adequate to describe all statistical wave fields; the higher order coherence functions may diverge. The proper treatment is in terms of the characteristic functional; this theory is developed in detail. The fundamental role played by the diagonal representation of the density operator is clearly brought out. It is shown that given the characteristic functional we can reconstruct the density matrix; the higher order coherence functions, when they exist can be obtained by a power series expansion of the characteristic functional.

An intermediate class of fields of illumination which include the mode-pure fields but less general than the general statistical field are considered and they are shown to be characterized by an auxiliary two-point function.

* Supported in part by the U. S. Atomic Energy Commission.

1. INTRODUCTION

The concept of partial coherence

The wave properties of light beams are well known. We know, for example, that light undergoes diffraction and can be made to undergo interference with other light beams. It obeys the general property of superposition. The predictions of the elementary theory of, say, double slit interference pattern or Fresnel diffraction furnish a quantitative interpretation of the observed pattern. Yet we know that the elementary theory cannot be a complete account of the phenomena involving light beams since it has no concept of partial coherence. We know that we do not obtain the familiar double slit interference pattern of equally spaced fringes with two arbitrary light sources but only with two secondary sources derived from the same source usually by using two slits in a screen kept at a sufficient distance from a narrow primary slit. We know from direct experimental evidence that the contrast between the intensity maxima and minima (the "visibility") becomes poorer as the distance of the primary source from the double slit is decreased. In this case we have a gradual transition from the fully coherent case (to which the elementary theory applies) through various stages of partial coherence to the case of complete incoherence. For a complete treatment of the problem, it is therefore necessary to discuss both the concept of partial coherence and the law of propagation of partial coherence. The questions have been systematically studied during the past few years within the framework of classical optics. The results of these investigations are of fundamental importance and are now classical [1] (in more than one sense!). One considers the light amplitudes $V(p_1, t_1)$, $V(p_2, t_2)$ at the points p_1 and p_2 at times t_1 and t_2 to be stochastic variables. The intensity of light at these two points are given by

$$(1.1) \quad \Gamma(p_j, t_j) = \langle V^*(p_j, t_j) V(p_j, t_j) \rangle$$

which is non-negative in all cases. The more general bilinear quantities

$$(1.2) \quad \Gamma(p_j, t_j, p_k, t_k) = \langle V^*(p_j, t_j) V(p_k, t_k) \rangle$$

are the (second order) coherence functions. For $p_j = p_k$, $t_j = t_k$ it reduces to the intensity function. One can also introduce the normalized degree of coherence

$$(1.3) \quad \gamma(p_j, t_j, p_k, t_k) = \frac{\Gamma(p_j, t_j, p_k, t_k)}{\sqrt{\Gamma(p_j, t_j, p_j, t_j) \Gamma(p_k, t_k, p_k, t_k)}}$$

which becomes unity when $p_j = p_k$, $t_j = t_k$. This quantity has a direct physical interpretation as a "visibility index": consider an opaque screen passing through the points p_1 and p_2 with pinholes at these two points. The two pinholes are not to be of equal size but the ratio of their sizes may be varied[†], they are however to be small enough so

[†] This differs somewhat from the standard discussion of the relation between visibility and the reduced coherence function.

that the variation of the mutual coherence function Γ as the points p_1 and p_2 are replaced by any other pair of points included in the area of the two pinholes respectively is negligible. We shall also assume that we have provision for shifting the phase of the light amplitude at either pinhole by any desired amount. This would of course involve ideal (nondispersive) phase plates. If we now superpose the light emerging from the two pinholes after their phases are shifted, the resulting amplitude is the stochastic variable

$$V(t) = a_1 e^{i\theta_1} V(p_1, t) + a_2 e^{i\theta_2} V(p_2, t)$$

where the real coefficients a_1 and a_2 depend on the sizes of the pinholes (and are to a good approximation proportional to their sizes). The intensity is now given by

$$I(t) = \langle V^*(t) V(t) \rangle = a_1^2 I(p_1, t) + a_2^2 I(p_2, t) + 2 a_1 a_2 \operatorname{Re} \left\{ \Gamma(p_1, t, p_2, t) e^{i(\theta_2 - \theta_1)} \right\}$$

By choosing

$$\frac{a_1}{a_2} = \pm \sqrt{\frac{I(p_2, t)}{I(p_1, t)}} \\ \theta_2 - \theta_1 = \arg \Gamma(p_2, t, p_1, t)$$

we get

$$I(t) = |a_1 a_2| \sqrt{I(p_1, t) I(p_2, t)} (1 \pm |\gamma|)$$

so that the ratio of the range of intensities by the average intensities gives[†]

$$(1.4) \quad \text{visibility} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = |\gamma|.$$

Thus the normalized degree of coherence is a measure of the optimum visibility of the field with the ratio of the transmission factors and the difference in the phases at the two slits at our command. If we keep the ratio a_2/a_1 different from the optimum ratio the visibility decreases. For the general case, as the phase difference $\theta_2 - \theta_1$ is increased monotonically the intensity associated with the superposed amplitude undergoes a periodic variation.

In practice ideal phase shifters ("phase plates") are not available; with light of relatively narrow variation in frequency it is often possible to choose a material such that the optical refractive index varies inversely as the frequency over the range of frequencies for which the light amplitude is appreciable. If the frequencies cover only a very narrow range (the "quasimonochromatic" case) we can approximate the phase shifter by propagation through free space. In this case we have the familiar optical arrangement of the double slit interference pattern. Even if the intensities at points P_1 and P_2 are not equal, by arranging the ratio of the sizes of the pinholes we can realize the maximum contrast.

[†] See foot note on page 122.

The law of propagation of the coherence functions

So far we have not specified the medium in which the light is propagating. If the medium is free space (or any passive linear medium) the propagation law is also linear; if the medium is non-linear so would the equations of motion be. The most important case is the propagation in free space; in this case the equations of motion for the wave disturbance is linear and homogeneous. It follows that since the ensemble averaging implicit in the definition of the coherence function commutes with space and time differentiation of the stochastic variables of the light amplitude, the coherence functions obey the "same" equations of motion as the wave amplitude. This is true whether we use a scalar amplitude or a solenoidal vector amplitude. The situation here is simpler than the case of correlation functions in hydrodynamic turbulence [2]. In the case of the correlation functions of turbulence the inertia term in the equations of motion of the fluid introduces a fundamental non-linearity which couples the correlation functions of different orders. In contrast, the equations of motion for the light amplitude relates the second order coherence function to itself.

The equations of motion for the light amplitudes

$$(1.5) \quad \left(\nabla^2 - \frac{\partial^2}{\partial t^2} \right) V = 0$$

is of second degree in the time derivative. Hence given the equal time coherence function and the derivative of the time-dependent coherence function evaluated for equal times (after the differentiation!) we can compute the time dependent coherence function for unequal times. The most important case is that for stationary states for which we have

$$(1.6) \quad \Gamma(p_1, t_1, p_2, t_2) = \Gamma(p_1, t_1 - t_2, p_2, 0) \equiv \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau).$$

In this case we can restrict attention to the wave equation

$$(1.7) \quad \frac{\partial^2}{\partial \tau^2} \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \nabla_1^2 \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \nabla_2^2 \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$$

where ∇_1^2 and ∇_2^2 are the Laplacians with respect to the space coordinates $\mathbf{r}_1, \mathbf{r}_2$ of the points P_1 and P_2 . We could rewrite this in the form

$$(1.8) \quad \frac{\partial}{\partial \tau} \begin{bmatrix} \Gamma \\ \frac{\partial \Gamma}{\partial \tau} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \nabla^2 & 0 \end{bmatrix} \begin{bmatrix} \Gamma \\ \frac{\partial \Gamma}{\partial \tau} \end{bmatrix}$$

which gives a linear equation for the two-component "vector" $\begin{bmatrix} \Gamma \\ \frac{\partial \Gamma}{\partial \tau} \end{bmatrix}$. Hence if these

quantities are specified at $\tau = 0$, the equations of motion determine it for all times. For a solenoidal vector amplitude the equations of motion are slightly more complicated but it is well-known that the Maxwell equations can be cast in the form

$$(1.9) \quad i \frac{\partial}{\partial t} (\mathbf{E} \pm i\mathbf{B}) = \pm \nabla \times (\mathbf{E} \pm i\mathbf{B})$$

so that the coherence functions obtained by considering $\mathbf{E} \pm i\mathbf{B}$ obey first order equations, but if we write it in terms of \mathbf{E} or \mathbf{B} we still have second order wave equations.

Analytic coherence functions

Instead of working with the real amplitudes, if we choose to work with the positive frequency part of the amplitude, the so-called analytic signal [4], then the second order equations of motion reduce to first order equations. For the scalar case we have

$$(1.10) \quad i \frac{\partial}{\partial \tau} \Gamma = + \sqrt{-\nabla^2} \Gamma$$

where $\sqrt{-\nabla^2}$ is an integral operator which has the form

$$(1.11) \quad \sqrt{-\nabla_1^2} \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = (2\pi)^{-3} \int d^3q |q| \exp(i\mathbf{q} \cdot \mathbf{r}_1) \times \int d^3r_1' \exp(-i\mathbf{q} \cdot \mathbf{r}_1') \Gamma(\mathbf{r}_1', \mathbf{r}_2, \tau).$$

Hence it follows that *given the equal time coherence function at all points, we can compute the unequal time coherence function also.* This result has the simple physical interpretation that the *frequency spectrum and the wave-number spectrum of natural light are essentially the same.* The complications introduced by considering the solenoidal vector amplitude are inessential. It is important however to note that while the temporal behaviour of the coherence function can be deduced from its spatial dependence, the polarization dependence must be specified explicitly.

The non-stationary case is not different in principle. We have for the coherence function with scalar analytic signals, the equations of motion

$$i \frac{\partial}{\partial t_1} \Gamma(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = \sqrt{-\nabla_1^2} \Gamma(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2)$$

$$i \frac{\partial}{\partial t_2} \Gamma(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = \sqrt{-\nabla_2^2} \Gamma(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2)$$

where $\sqrt{-\nabla_1^2}$ is an integral operator defined by

$$\sqrt{-\nabla_1^2} \Gamma(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = (2\pi)^{-3} \int d^3q |q| \exp(i\mathbf{q} \cdot \mathbf{r}_1) \times \int d^3r_1' \exp(-i\mathbf{q} \cdot \mathbf{r}_1') \Gamma(\mathbf{r}_1', t_1, \mathbf{r}_2, t_2)$$

Hence, if $\Gamma(\mathbf{r}_1, 0, \mathbf{r}_2, 0)$ is specified for all values of $\mathbf{r}_1, \mathbf{r}_2$ we can compute the general second order coherence function. The temporal behaviours are again completely determined in terms of the spatial dependences.

Properties of the reduced coherence functions

Because of the postulated requirement that the mapping of the stochastic variables preserve reality and nonnegativity conditions we can show that the reduced coherence function $\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ satisfies the inequalities [1]

$$0 \leq |\gamma(\mathbf{r}_1, \mathbf{r}_1, \tau)| \leq 1,$$

$$0 \leq |\gamma(\mathbf{r}_1, \mathbf{r}_2, 0)| \leq 1.$$

For the nonstationary case, we have, correspondingly

$$0 \leq |\gamma(\mathbf{r}_1, t, \mathbf{r}_2, t)| \leq 1,$$

$$0 \leq |\gamma(\mathbf{r}, t_1, \mathbf{r}, t_2)| \leq 1.$$

It can be shown [5] that if $|\gamma(\mathbf{r}, \mathbf{r}, \tau)| = 1$ for all τ and some one value \mathbf{R} of \mathbf{r} , then for this value \mathbf{R} we must have

$$\gamma(\mathbf{R}, \mathbf{R}, \tau) = \exp(-2\pi i \nu \tau)$$

for a suitable ν . We can further prove the following results [6]

(1) If for some pair of values $\mathbf{r}_1 = \mathbf{R}_1$, $\mathbf{r}_2 = \mathbf{R}_2$ and for all real values of τ the reduced coherence function is unimodular, then

$$\gamma(\mathbf{R}_1, \mathbf{R}_2, \tau) = \exp(i\beta - 2\pi i \nu \tau)$$

where β and ν are independent of τ and γ satisfies the factorization property

$$\gamma(\mathbf{r}_1, \mathbf{R}_2, \tau) = \gamma(\mathbf{R}_1, \mathbf{R}_2, 0) \gamma(\mathbf{r}, \mathbf{R}_1, \tau)$$

for every point \mathbf{r} .

(2) If for all values of \mathbf{r}_1 and \mathbf{r}_2 in a domain of space and for all real τ values the reduced coherence function $\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ is unimodular, then it must necessarily be of the form

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \exp\{i[\alpha(\mathbf{r}_1) - \alpha(\mathbf{r}_2) - 2\pi \nu \tau]\}$$

where $\alpha(\mathbf{r})$ is a real function and ν is a nonnegative constant. These theorems require only the nonnegativity property of the coherence function $\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)$ which implies that

$$\langle F^*(t) F(t) \rangle \geq 0$$

for all $F(t)$ of the form

$$F(t) = \int d^3 r' f(\mathbf{r}') V(\mathbf{r}', t + \tau(\mathbf{r}'))$$

and in particular

$$F(t) = \sum_{j=1}^n \beta_j V(\mathbf{x}_j, t + \tau_j).$$

The last one leads to the nonnegativity condition

$$\sum_{j=1}^n \sum_{k=1}^n \alpha_j^* \gamma(r_j, r_k, \tau_j - \tau_k) \alpha_k \geq 0$$

which forms the basis of the proof of the results stated above.

If we consider instead of the unimodularity for all τ , the case of the unimodularity for all points r_1, r_2 for some fixed τ we can again prove some similar results. We have for example the following result :

Given that for some times t_1, t_2 and all pairs of points r_1, r_2 the reduced coherence function $\gamma(r_1, t_1, r_2, t_2)$ is unimodular, it must factorize in the form

$$\gamma(r_1, t_1, r_2, t_2) = \exp \left\{ i \alpha(r_1, t) - i \alpha(r_2, t) \right\}$$

where $\alpha(r, t)$ is a real function. This result could be proved in essentially the same fashion by use of the non-negativity property of the reduced coherence function. We shall not prove it here but show that it is a special case of a general theorem about wave fields with unimodular coherence of any order. In the later part (Sec. 4) of this paper we shall prove the more general theorem in a fashion which applies equally well to quantum theory and classical theory.

Decomposition of a coherence function as a sum of coherence functions with unimodular reduced coherence functions

We now discuss a result which deals with the second order coherence functions for partially coherent fields, but which does not have an immediate generalization to higher order coherence functions. For this purpose we note that the coherence functions $\Gamma(r_1, t_1, r_2, t_2)$ constitute a convex cone [6] in the sense that, if $\Gamma_{(k)}(r_1, t_1, r_2, t_2)$ are a set of coherence functions and λ_k are a set of nonnegative numbers, then

$$\Gamma(r_1, t_1, r_2, t_2) = \sum_k \lambda_k \Gamma_{(k)}(r_1, t_1, r_2, t_2)$$

is also an allowed coherence function. We might now ask what are the generators of such a convex cone. To make the problem more tractable we shall restrict attention for the moment to coherence functions which belong to the Hilbert-Schmidt class of square integrable functions: $\int d^3r_1 \int d^3r_2 |\Gamma(r_1, t_1, r_2, t_2)|^2 < \infty$. From the discussion earlier in this section it is clear that we could without loss of generality, restrict attention to $t_1 = t_2 = 0$. Now it is clear that if

$$(1.12) \quad \Gamma(r_1, 0, r_2, 0) = \Gamma(r_1, r_2) = \sum_k \lambda_k \Gamma_{(k)}(r_1, r_2), \quad \lambda_k \geq 0$$

then without loss of generality we may choose

$$\int d^3r_1 \int d^3r_2 |\Gamma_{(k)}(\mathbf{r}_1, \mathbf{r}_2)|^2 = 1$$

for all k . Since the reduced coherence function is unimodular if and only if the coherence function factorizes it follows that unless all but one such λ_k are zero, the reduced coherence function obtained by the convex linear combination cannot be unimodular. We can now ask whether the generating elements of the convex cone of coherence functions are all unimodular; i.e., whether we could express every coherence function in the form (1.12) with all the $\Gamma_{(k)}(\mathbf{r}_1, \mathbf{r}_2)$ chosen such that the corresponding reduced coherence functions $\gamma_{(k)}(\mathbf{r}_1, \mathbf{r}_2, 0)$ being unimodular. The answer is in the affirmative, and applies to the wider class of "completely continuous" two-point functions (which, of course, includes the Hilbert-Schmidt class). We state the precise result in the following form:

Let the coherence function $\Gamma(\mathbf{r}_1, \mathbf{r}_2)$ considered as the integral operator

$$(1.13) \quad \Gamma \psi(\mathbf{r}) = \int d^3r' \Gamma(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}')$$

be completely continuous. Then the coherence function can be expressed as the convex linear (possibly infinite) sum of coherence functions with unimodular reduced coherence functions.

To prove this result it is sufficient to consider the eigenvector expansion for the "integral operator" $\Gamma(\mathbf{r}_1, \mathbf{r}_2)$

$$(1.14) \quad \Gamma(\mathbf{r}_1, \mathbf{r}_2) = \sum \lambda_k f_k^*(\mathbf{r}_1) f_k(\mathbf{r}_2).$$

Such an expansion always exists for a completely continuous operator [7]. For a nonnegative operator the λ_k are nonnegative. If we now identify $\Gamma_{(k)}(\mathbf{r}_1, \mathbf{r}_2)$ with $f_k^*(\mathbf{r}_1) f_k(\mathbf{r}_2)$ we have the desired result.

Earlier in this section we have shown that the $t_1 = t_2 = 0$ values of the coherence function determines the coherence function for all times by virtue of the equations of motion. But if the coherence function factorizes at $t_1 = t_2 = 0$, then it must factorize for all times in the form

$$(1.15) \quad \Gamma_{(k)}(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = f_k^*(\mathbf{r}_1, t_1) f_k(\mathbf{r}_2, t_2)$$

with

$$(1.16) \quad \begin{aligned} f_k(\mathbf{r}, t) &= \exp\left(i t \sqrt{-\nabla^2}\right) f_k(\mathbf{r}) \\ &= (2\pi)^{-3} \int d^3q \exp(i \mathbf{q} \cdot \mathbf{r} - i |q| t) \times \\ &\quad \int d^3r' \exp(-i \mathbf{q} \cdot \mathbf{r}') f_k(\mathbf{r}'). \end{aligned}$$

Hence we have the theorem:

The (second order) time-dependent coherence functions (which belong to the completely continuous class) constitute a convex cone. The generators of this convex

one are given by coherence functions whose equal time reduced coherence functions are unimodular.

The restriction of the discussion to the completely continuous class of coherence functions is not of any consequence for most physical situations since even the subclass of Hilbert-Schmidt integral operators are too wide a class: in most physical situations we are interested in the total energy in the field to be finite. With the electromagnetic field being characterized by the coherence functions of the electric or magnetic fields, this implies that

$$\int d^3r \Gamma(\mathbf{r}, t, \mathbf{r}, t) < \infty,$$

so that the nonnegative integral operator $\Gamma(\mathbf{r}_1, t, \mathbf{r}_2, t)$ belongs to the trace class. Now it is known that this class is a proper subset of the class of square integrable operators which in turn are contained in the class of completely continuous operators. So the decomposition theorem applies to all such cases. If instead of using the electric field we use the vector potential to construct the coherence functions the same decomposition holds in all cases where the coherence functions are bounded at infinity.

It has been pointed out to the author by C. L. Mehta that the blackbody coherence functions are an exception to this rule; in this case the trace diverges. But we could express the coherence function in this case as an integral over coherence functions with unimodular reduced coherence functions.

Outline of the contents of the following sections

In the following sections we give a systematic development of the quantum theory of partial coherence. In Sec. 2 we discuss the question of the determination of the (classical) unequal time coherence functions from their equal-time boundary values. A result of this study is the identification of the wave number spectrum and the frequency spectrum. In this section we also discuss the measurement of the higher coherence functions as a prescribed linear combination of intensity correlations. In Sec. 3 the complex variable description of a quantum-mechanical system is introduced and related to the definition of second order coherence functions. It is shown that all the known properties of second order coherence functions in classical optics is recovered in the quantum optics of second order coherence.

Quantum theory of coherence functions of arbitrary order is introduced in Sec. 4. For those states which have excitation in only one mode, the coherence functions factorize; such fields of illumination are referred to as "mode-pure." In terms of the coherence functions a new set of quantities called the "coherence indices" are defined; they are generalizations of the reduced coherence functions of second order. For mode-pure fields the coherence indices of all orders (μ, μ) are unimodular; the converse problem of determining the density matrix of a field for which some coherence index is unimodular is posed and solved; the states involve only some few photon states in addition to a mode-pure excitation. It is shown that all "higher" coherence functions are also factorizable.

The theory of mode-pure fields is taken up in Sec. 5. The coherence functions are defined in terms of a mode function and a sequence of moments. Several inequalities satisfied by these moments are derived.

The use of coherence functions to specify the statistical state of a system is not adequate to cover the general case. A method capable of dealing with arbitrary statistical state in terms of characteristic functionals is developed in Sec. 6 (for the mode-pure fields) and in Sec. 7. The problem of determining the density matrix from the characteristic functional is posed and solved; a fundamental role is played by the diagonal representation of statistical states in this reconstruction of the density matrix from the characteristic functional.

Fields of illumination which can be obtained by mixtures of mode-pure illuminations are studied in Sec. 8. In this section a necessary and sufficient criterion for such a field of illumination is presented; this involves the use of an auxiliary two-point function. Several theorems concerning such fields of illumination are stated and proved. In Sec. 9 it is shown that the only excited modes are those which occur in the eigenvector decomposition of the second order coherence function. This result has a simple form when re-expressed in terms of the characteristic functional.

The questions of equations of motion, conservation laws and the coupling with arbitrary time-dependent dynamical systems are not discussed in this paper. The author hopes to return to these questions in another paper.

2. HIGHER ORDER COHERENCE FUNCTIONS: CLASSICAL THEORY

In recent years experiments involving intensity correlations, photoelectric counting distributions and phenomena occurring in high intensity laser beams have made it necessary to consider processes in which higher order coherence functions have to be studied [8]. In elementary optics one is mainly interested in intensity distributions but the propagation law of intensity automatically involves the twopoint coherence function: on the other hand, the coherence function can be studied by studying intensity distributions on a screen illuminated by pinholes at the points of which the coherence function is being considered. Similarly even if one is interested only in two-point intensity correlations directly, we have to study the general four-point coherence functions and so on. We are thus led to consider the general coherence function of order (μ, ν) defined as [8]

$$(2.1) \quad \Gamma(x_1, \dots, x_\mu, y_1, \dots, y_\nu) = \langle V^*(x_1) \dots V^*(x_\mu) V(y_1) \dots V(y_\nu) \rangle$$

where $V(y)$ consists of the "analytic signal" (positive frequency) part of the field

$$(2.2) \quad V(x) = \sum_{\alpha} a_{\alpha} u_{\alpha}(x)$$

Here $u(x)$ are a complete set of positive energy mode functions. We shall adopt the convention that the label x stands for the space indices as well as polarization labels but not the time.

Determination of the unequal-time coherence function from equal-time coherence functions

In view of our remarks in the last two sections, for a free field the time dependent operator $V(x, t)$ will contain only positive frequencies and therefore it would obey a first order equation of the form

$$(2.3) \quad i \frac{\partial}{\partial t} V(x, t) = + \sqrt{-\nabla^2} V(x, t)$$

so that the time-dependent coherence functions can be constructed from the knowledge of the time independent coherence function of the same order (μ, ν) at all points (and all polarization indices). We have for example

$$(2.4) \quad \Gamma(x_1 t_1 x_2 t_2 \dots y_\nu t'_\nu) \\ = \exp\left(-it_1 \sqrt{-\nabla_1^2}\right) \Gamma(x_1 0 x_2 t_2 \dots y_\nu t'_\nu) \\ = \exp\left(+it'_1 \sqrt{-\nabla_1'^2}\right) \Gamma(x_1 t_1 x_2 t_2 \dots y_1 0 \dots y_\nu t'_\nu)$$

where $\exp\left\{-it_1 \sqrt{-\nabla_1^2}\right\}$ and $\exp\left\{+it'_1 \sqrt{-\nabla_1'^2}\right\}$ are

integral operators respectively with respect to x_1 and y_1 of the kind discussed in the previous section. Hence, we can restrict attention to equal time functions without any essential loss of generality.

Coherence functions and intensity correlations

The coherence functions of orders (μ, ν) with $\mu = \nu$ are closely related to μ -point intensity correlations and are therefore of most immediate interest. Considered as a matrix in the indices $x_1 \dots x_\mu, y_1 \dots y_\mu$ they constitute, by virtue of their definition a hermitian nonnegative matrix. By virtue of the integral propagation law, the propagation law for the μ -point intensity correlation involves the coherence function of order (μ, μ) . On the other hand we can study the $(\mu, \mu)^{th}$ order coherence function in terms of intensity correlations [9] by considering 2μ pinholes (of equal size) at the points $x_1 \dots x_\mu, y_1 \dots y_\mu$ with phase shifters at each of these points. Making use of the elementary identity

$$(2.5) \quad 4v^\dagger \omega = (v + \omega)^\dagger (v + \omega) - (v - \omega)^\dagger (v - \omega) \\ + i(v - i\omega)^\dagger (v - i\omega) - i(v + i\omega)^\dagger (v + i\omega)$$

we have the equality

$$4\kappa^{-2} \Gamma(x_1 t_1 x_2 t_2 \dots; y_1 t_1' y_2 t_2' \dots) \\ = \Gamma^{(0)}(\xi \tau x_2 t_2 \dots; \xi \tau y_2 t_2' \dots) \\ + i \Gamma^{(1)}(\xi \tau x_2 t_2 \dots; \xi \tau y_2 t_2' \dots) \\ - \Gamma^{(2)}(\xi \tau x_2 t_2 \dots; \xi \tau y_2 t_2' \dots) \\ - i \Gamma^{(3)}(\xi \tau x_2 t_2 \dots; \xi \tau y_2 t_2' \dots)$$

where $\Gamma^{(i)}(\xi \tau x_2 t_2 \dots; \xi \tau y_2 t_2' \dots)$ is the function obtained at a point ξ at time τ illuminated by two pinholes at x_1 and y_1 with phase plates furnishing phase differences

$$\theta_1' - \theta_1 = -j\pi/2$$

and a path difference equivalent to a time delay

$$\tau_1' - \tau_1 = t_1 - t_1'$$

The quantity κ is a measure of the size of the pinholes and the distance of the point ξ from the pinholes. Proceeding in this fashion we obtain

$$(2.6) \quad (2/\kappa)^{2\mu} \Gamma(x_1 t_1 x_2 t_2 \dots; y_1 t_1' y_2 t_2' \dots) \\ = \sum_{j_1 \dots j_\mu} (i)^{j_1 + j_2 + \dots + j_\mu} \Gamma^{(j_1 \dots j_\mu)}(\xi_1 \tau \xi_2 \tau \dots \xi_1 \tau \xi_2 \tau)$$

where each of the $\Gamma^{(j_1 \dots j_\mu)}$ is an intensity correlation at equal time obtained by combining at the points ξ_1 the illumination from the pinholes at x_1, y_1 with relative phase shift $\theta_1' - \theta_1 = -j_1 \pi/2$ and a path difference equivalent to $\tau_1' - \tau_1 = t_1 - t_1'$; at the point ξ_2 the illumination from the pinholes at x_2, y_2 with $\theta_2' - \theta_2 = -j_2 \pi/2$ and $\tau_2' - \tau_2 = t_2 - t_2'$ and so on. Hence the (μ, μ) order coherence function can be obtained as a linear combination of 2^μ equal time intensity correlations.

3. QUANTUM THEORY OF SECOND ORDER COHERENCE

The discussion in the previous section was within the framework of classical optics. While it provides us with a complete theory of second order partial coherence one would like to see the modifications introduced by quantum theory. The basic question is

of course the quantum mechanical definition of the light amplitude and the rule for associating the coherence function with a quantum mechanical state [10, 18]. There is no question of considering a complex value for the electric field but we can associate a complex function of space with the annihilation part of the quantum mechanical field operator. For a *free* field the annihilation part is associated with positive frequencies [11]* and in this case the complex function of space that is associated with the annihilation part of the field operator will be an analytic signal.

The complex variable description of a classical canonical system

To make this identification we consider a free harmonic oscillator with the equations of motion (we choose the mass = 1) :

$$\frac{d}{dt} q(t) = p(t)$$

$$\frac{d}{dt} p(t) = -\omega^2 q(t)$$

the general solution to which can be written as

$$p(t) = p(0) \cos \omega t - \omega q(0) \sin \omega t$$

$$q(t) = \omega^{-1} p(0) \sin \omega t + q(0) \cos \omega t$$

If we now choose the complex plane and choose

$$x = \text{Re } z = \sqrt{\omega/2} q$$

$$y = \text{Im } z = \sqrt{1/2\omega} p$$

then the vector

$$(3.1) \quad z = x + iy = (1/2\omega)^{\frac{1}{2}} (\omega q + ip)$$

may be used to denote the instantaneous state of a *classical* oscillator. The state represented by $z = 0$ corresponds to the lowest energy state. As a function of time the vector $z(t)$ rotates in the counterclockwise direction along a circle of radius

$$|z| = (2\omega)^{-\frac{1}{2}} (\omega^2 q^2 + p^2)^{-\frac{1}{2}}$$

so that the square of radius is the total energy of the oscillator divided by the angular frequency. The state labelled by the complex number z can be obtained by starting with the minimum state (labelled by $z = 0$) and displacing the complex dynamical variable $\omega q + ip$ from 0 to $\sqrt{2\omega} z$. Since the dynamical variable $(iq + \omega^{-1} p)$ is the generator of displacements in $(\omega q + ip)$ so that

In general it is essential to distinguish between the annihilation part of the field operator and the positive frequency part [11].

$$(3.2) \quad [\omega q + ip, iq + \omega^{-1} p]_{p, b} = -i,$$

so that the exponentiation of this transformation would generate the state z from the state 0. We can make these dynamical variables appear more symmetric if we consider the complex dynamical variables

$$(2\omega)^{-\frac{1}{2}} (\omega q \pm ip)$$

which have a Poisson bracket of i . It is relevant to note that the equations of motion for the free harmonic oscillator give

$$(3.3) \quad \{\omega q(t) \pm ip(t)\} = \exp(\mp i\omega t) \{\omega q(0) \pm ip(0)\}$$

so that we could identify the complex dynamical variable $(2\omega)^{-\frac{1}{2}} (\omega q + ip)$ corresponds to the positive frequency part of the real coordinate variable $q(t)$ (apart from the factor $\sqrt{2\omega}$). However, whether the equations of motion of the system corresponds to a free harmonic oscillator or not, as long as the system has one pair of canonical variables needed to describe it. We would describe it in terms of this complex canonical variable. In the general case the representative point will no longer move on a circle but describe a more complicated phase space trajectory.

The complex variable description of a quantum canonical system

It is remarkable that while in quantum theory we cannot specify the state by the values of q and p we can label states by a single complex number z and that these states are obtained from a standard state labelled by 0 by using a displacement operator [12]. In this connection we consider a pair of canonical variables q and p satisfying the commutation relation

$$(3.4) \quad ap - pq = i$$

(We choose units such that $\hbar = 1$.) Then we may construct the operators a and a^\dagger defined by

$$(3.5) \quad a = (2\omega)^{-\frac{1}{2}} (\omega q + ip)$$

$$a^\dagger = (2\omega)^{-\frac{1}{2}} (\omega q - ip)$$

which then satisfy the commutation relation

$$(3.6) \quad a a^\dagger - a^\dagger a = 1$$

We can now set up an infinite family of quantum mechanical states labelled by a single complex number which are eigenstates of the destruction operator as follows:

$$a |\Omega\rangle = 0 = 0 |\Omega\rangle$$

where $|\Omega\rangle$ is the ground state of the oscillator defined by

$$(3.8) \quad a^\dagger a |\Omega\rangle = 0.$$

The state corresponding to the eigenvalue z of a is constructed using the displacement operator a^\dagger to obtain

$$\exp(za^\dagger) |\Omega\rangle$$

which has the property

$$a \exp(za^\dagger) |\Omega\rangle = z \exp(za^\dagger) |\Omega\rangle$$

If we denote by $|z\rangle$ the normalized eigenstate of a with the eigenvalue z , then

$$(3.9) \quad |z\rangle = \exp(za^\dagger - \frac{1}{2}|z|^2) |\Omega\rangle$$

$$|0\rangle = |\Omega\rangle.$$

For a free harmonic oscillator these special states have a simple time dependence of the form

$$(3.10) \quad |z, t\rangle = \exp(-i\omega t a^\dagger a) |z\rangle = |ze^{-i\omega t}\rangle$$

so that just like for the classical oscillator the representative complex number moves along a circle at uniform speed in the clockwise direction. In the more general case of an arbitrary quantum mechanical system described by a pair of canonical variables, unlike the classical case, the time dependence of these states cannot be represented by the motion of representative point in the complex plane. It has been pointed out that as long as the equation of motion for the operator a contains only functions of the annihilation operator we could describe the time evaluation by a phase space trajectory. For a Hamiltonian system only such equations of motion are linear in the annihilation operator.

These states constitute an over-complete family of states; the completeness property [12] is expressed by the fundamental identity

$$(3.11) \quad \pi^{-1} \int d^2 z |z\rangle \langle z| = 1$$

so that, every state can be expanded in terms of these in the form

$$(3.12) \quad |\psi\rangle = \pi^{-1} \int d^2 z \langle z | \psi \rangle |z\rangle$$

and every operator can be expressed in the form

$$(3.13) \quad B = \pi^{-2} \int d^2 z \int d^2 \zeta \langle z | B | \zeta \rangle |z\rangle \langle \zeta|$$

Of particular interest in this case is the equation of motion for the oscillator with

a linear coupling to an external variable [13]*. The equations of motion can be written in the form

$$\begin{aligned}\frac{d}{dt} q^{(t)} &= p^{(t)} \\ \frac{d}{dt} p^{(t)} &= j - \omega^2 q^{(t)}\end{aligned}$$

and they can be reduced to the previous form by the change of variables

$$\begin{aligned}g(t) &\longrightarrow g(t) - \omega^{-2} j \\ p(t) &\longrightarrow p(t)\end{aligned}$$

The representative points are now displaced by $(2\omega^2)^{-\frac{1}{2}} j$ in the negative q direction. The time dependence of a state is now represented by motion in the clockwise direction along a circle centered at the point $x = -(2\omega^2)^{-\frac{1}{2}} j$, $y = 0$.

Complex variable description of the free field in quantum theory

The Hamiltonian for a free scalar field or the free electromagnetic field can be cast into the form of the sum of the Hamiltonians of an infinite collection of free (ideal) harmonic oscillators. Correspondingly we could expand the real field operator $\varphi(x)$ and the canonically conjugate real momentum density $\pi(x)$ for the scalar field case in the form

$$\begin{aligned}(3.14) \quad \varphi(x) &= \sum_{\alpha} \left\{ a_{\alpha} u_{\alpha}(x) + a_{\alpha}^{\dagger} u_{\alpha}^{*}(x) \right\} \\ \pi(x) &= \sum_{\alpha} \left\{ a_{\alpha} u_{\alpha}(x) - a_{\alpha}^{\dagger} u_{\alpha}^{*}(x) \right\}\end{aligned}$$

with the $u_{\alpha}(x)$ furnishing a complete set of positive frequency wave functions which satisfy the orthogonality and completeness relations

$$\begin{aligned}(3.15) \quad \int d^3 r u_{\alpha}(x) u_{\beta}(x) &= \delta_{\alpha\beta} \\ \sum_{\alpha} u_{\alpha}^{\dagger}(x_1) u_{\alpha}(x_2) &= \delta(x_1 - x_2)\end{aligned}$$

By virtue of these we have the inverse relations

$$\begin{aligned}(3.16) \quad a_{\alpha} &= \int d^3 x u_{\alpha}^{*}(x) \varphi(x) \\ a_{\alpha}^{\dagger} &= \int d^3 x u_{\alpha}(x) \varphi(x)\end{aligned}$$

* The most general case of motion which can be represented in terms of a trajectory in x -plane has been studied by Glauber [13].

The overcomplete family of states associated with the field operators is thus essentially the family of states associated with the operators $a_\alpha, a_\alpha^\dagger$. These states can be labelled by a set of complex numbers z_α one for each mode as follows :

$$(3.17) \quad a_\alpha | z_1, z_2, \dots \rangle = z_\alpha | z_1, z_2, \dots \rangle \\ = \exp \left\{ -\frac{1}{2} \sum_\alpha |z_\alpha|^2 \right\} \exp \left\{ \sum_\alpha z_\alpha a_\alpha^\dagger \right\} |0, 0, \dots \rangle$$

These states associated with the annihilation part $A(x)$ of the field operator the complex function

$$(3.18) \quad V(x) = \sum_\alpha z_\alpha u_\alpha(x)$$

Definition of the second order coherence functions in quantum theory

The fundamental definition of the quantum theory of partial coherence is the equation

$$(3.19) \quad \Gamma(x_1, x_2) = \text{tr} \{ A(x_2) \rho A^\dagger(x_1) \}$$

where ρ is the density matrix of the state and $A(x_2), A^\dagger(x_1)$ are respectively the parts involving a and a^\dagger in the expression for the field operator. We may rewrite this in the form

$$(3.20) \quad \Gamma(x_1, x_2) = \sum_{\alpha, \beta} u_\alpha^*(x_1) u_\beta(x_2) \text{tr} \{ a_\beta \rho a_\alpha^\dagger \}.$$

For the classical coherence function we would have in place of this equation the relation

$$(3.21) \quad \Gamma(x_1, x_2) = \sum_{\alpha, \beta} u_\alpha(x_1) u_\beta(x_2) \langle a_\alpha a_\beta \rangle \\ = \int d^2 z_1 \int d^2 z_2 \dots \phi(z_1, z_2, \dots) z_\alpha^* z_\beta$$

where z_α stands for the classical stochastic variable corresponding to the quantum dynamical variable a_α . The distribution $\phi(z_1, z_2, \dots)$ takes the place of the density matrix ρ . There is thus a formal correspondence between the classical entities and their quantum counterparts. This correspondence has formed the motivation for the statement of the Optical Equivalence Theorem; and for its rigorous mathematical formulation in

terms of distributions. In Sec. 6 below we would come across the quantum theory definition of the distribution $\phi(z)$ corresponding to the so-called "diagonal representation." This correspondence is discussed at length elsewhere, and it is shown that such a distribution can be defined unambiguously so that the expectation values of all (bounded) dynamical variables are uniquely specified.

Properties of second order coherence functions: Comparison with classical theory

For the present we need note however only the properties of the quantities $\text{tr} \left\{ a_\beta \rho a_\alpha^\dagger \right\}$. It follows from the non-negativity of the density matrix ρ that the matrix

$$\Gamma_{\alpha\beta} = \text{tr} \left\{ a_\beta \rho a_\alpha^\dagger \right\}$$

is nonnegative so that the quantum coherence function $\Gamma(x_1, x_2)$ is nonnegative. Further since the entire x_1, x_2 dependence is contained in the "mode functions", $u_\alpha^*(x_1)$ and $u_\beta(x_2)$ for the free field, the time development of $\Gamma(x_1, t_1, x_2, t_2)$ is entirely governed by the time development of the positive frequency mode functions. Since the time dependent mode functions $u_\alpha(x, t)$ are analytic signals (by definition) it follows that in the variables t_1, t_2 the coherence function satisfies first order differential equations. Hence their values at $t_1 = t_2 = 0$ for all points x_1, x_2 determine them for all times. Finally the nonnegativity properties automatically lead to factorization for unimodular reduced coherence functions; and to the theorem on the generators of the convex cone of coherence functions. Each of the generating elements of the set of coherence functions correspond to excitation of the field to a suitable extent in one mode only.

Thus all the results of modern classical theory of second order partial coherence are unaltered in the quantum theory formulation. (It is necessary to point out that if instead of using analytic signals in classical theory and annihilation part of the field in quantum theory we had worked with real fields in the classical theory and the hermitian field quantum theory this correspondence would have been different. However there is no need to work with real fields.)

Finally, the restriction to scalar fields can easily be removed. The only modification is to make the mode functions more complicated. For the case of the electromagnetic field the mode functions should contain either a two-valued polarization (helicity) label, or be expressed equivalently in terms of solenoidal vector wave functions. Since these complications are purely kinematical no harm is done by ignoring these complications in the present investigation.

4. HIGHER ORDER COHERENCE FUNCTIONS: QUANTUM THEORY

Having demonstrated the equivalence of the quantum theory and the classical theory of the second order coherence let us now return to the properties of the coherence function [10]

$$(4.1) \quad \Gamma(x; y) = \Gamma(x_1, \dots, x_\mu, y_1, \dots, y_\mu) = \text{tr} \left\{ \mathcal{A}^\dagger(y) \rho \mathcal{A}(x) \right\} \\ \equiv \text{tr} \left\{ A(y_1) \dots A(y_\mu) \rho A(x_1) \dots A(x_\mu) \right\}$$

of order (μ, μ) . By virtue of its definition, we have

$$(4.2) \quad \Gamma^*(x, y) = \Gamma^*(x_1, \dots, x_\mu, y_1, \dots, y_\mu) = \Gamma(y_1, \dots, y_\mu, x_1, \dots, x_\mu) = \Gamma(y, x)$$

and, hence, for the time dependent functions

$$(4.3) \quad \Gamma^*(x_1, t_1, \dots, y_\mu, t_\mu) = \Gamma^*(x, t, y, t') = \Gamma(y - t', x - t)$$

Considered as a matrix in the indices x_1, \dots, x_μ and y_1, \dots, y_μ , $\Gamma(x, y)$ is a non-negative definite matrix in the sense that

$$(4.4) \quad \int dx_1 \dots \int dx_\mu f(x_1, \dots, x_\mu) \Gamma(x_1, \dots, x_\mu, y_1, \dots, y_\mu) \int dy_1 \dots \int dy_\mu f^*(y_1, \dots, y_\mu) \Gamma(x, y) \\ = \int dx f(x) \int dy f^*(y) \Gamma(x, y) \geq 0$$

From its definition it follows that $\Gamma(x; y)$ is symmetric under any permutations of the variables x_1, x_2, \dots, x_μ and under any permutation of the variables y_1, \dots, y_μ . The nonnegativity of $\Gamma(x; y)$ is a consequence of the nonnegativity of the density matrix ρ since the expression above is simply the trace of the nonnegative matrix

$\mathcal{A}^\dagger \rho \mathcal{A}$ where

$$\mathcal{A} = \int dx f(x) \mathcal{A}(x) = \int dx_1 \dots \int dx_\mu f(x_1, \dots, x_\mu) A(x_1) A(x_2) \dots A(x_\mu)$$

By considering the operator

$$(4.5) \quad D(x, y) = \mathcal{A}(x) - \frac{\Gamma(x, y)}{\Gamma(y, y)} \mathcal{A}(y)$$

and the nonnegativity of the trace of the nonnegative matrix $D^\dagger \rho D$ we find that the optical discriminant of order (μ, μ) given by

$$(4.6) \quad \Delta(x, y) = \Gamma(x, y) \Gamma(y, x) - \Gamma(y, y) \Gamma(x, x)$$

is nonpositive: $\Delta(x, y) \leq 0$.

Let us define the coherence index of order (μ, μ) by [14, 17]

$$(4.7) \quad S(x, y) = \sqrt{\frac{\Gamma(x, y)}{\Gamma(x, x) \Gamma(y, y)}}$$

* We call these quantities coherence indices rather than normalized (or normalized) coherence functions since there are several such definitions already in literature [17].

The Coherence Index so defined coincides with the reduced (or normalized) coherence function for $\mu = 1$ (second order); since for higher order other authors [14] have defined different reduced functions, to avoid confusion, we shall refer to these reduced quantities as coherence indices in the sequel. Then the nonpositivity of the optical discriminant implies that the absolute value of the coherence index cannot exceed unity :

$$(4.8) \quad 0 \leq |S(x, y)| \leq 1$$

We have the following relation between the optical discriminant $\Delta(x, y)$ and the absolute value of the coherence index :

$$\Delta(x, y) = \{ |S(x, y)|^2 - 1 \} \Gamma(x, x) \Gamma(y, y)$$

The optical discriminant (4.6) vanishes if and only if the coherence index is unimodular. In this case the nonnegative matrix $D \rho D^\dagger$ must vanish :

$$D \rho D^\dagger = 0 \quad \begin{array}{c} \longrightarrow \\ \longleftarrow \end{array} \quad |S| = 1$$

We shall see later that this implies a very special form of the density matrix specifying the state of the system.

For the more general case of the coherence function of order (μ, ν) we may write

$$(4.9) \quad \Gamma^{(\mu, \nu)}(x, y) = \Gamma(x_1, \dots, x_\mu; y_1, \dots, y_\nu)$$

We now define the optical discriminant for this case by

$$\Delta^{(\mu, \nu)}(x, y) = \Gamma^{(\mu, \nu)}(x, y) \Gamma^{(\nu, \mu)}(y, x) - \Gamma^{(\mu, \mu)}(x, x) \Gamma^{(\nu, \nu)}(y, y)$$

which can again be shown to be nonpositive by considering the operators

$$(4.10) \quad D(x, y) = \mathcal{G}_t^{(\mu)}(x) - \frac{\Gamma^{(\nu, \mu)}(y, x)}{\Gamma^{(\nu, \nu)}(y, y)} \mathcal{G}_t^{(\nu)}(y)$$

and the nonnegative matrix $D \rho D^\dagger$

where

$$\mathcal{G}_t^{(\mu)}(x) = A(x_1) \dots \dots \dots A(x_\mu)$$

$$\mathcal{G}_t^{(\nu)}(y) = A(y_1) \dots \dots \dots A(y_\nu)$$

We define the coherence index of order (μ, ν) by

$$(4.11) \quad S^{(\mu, \nu)}(x; y) = \frac{\Gamma^{(\mu, \nu)}(x, y)}{\sqrt{\Gamma^{(\mu, \mu)}(x, x) \Gamma^{(\nu, \nu)}(y, y)}}$$

Then we have the relations

$$\Delta^{(\mu, \nu)}(x, y) = \left\{ |S^{(\mu, \nu)}(x, y)|^2 - 1 \right\} \Gamma^{(\mu, \mu)}(x, x) \Gamma^{(\nu, \nu)}(y, y) \leq 0$$

$$0 \leq |S^{(\mu, \nu)}(x, y)| \leq 1$$

Further the vanishing of Δ implies unimodularity of $S^{(\mu, \nu)}$ and the vanishing of the matrix $D^\dagger \rho D$. We shall see later that this also implies a very special form for the density matrix.

Given the density matrix ρ for the system one can attempt to calculate all the moments $\Gamma(x, y)$. We may therefore attempt to characterize [15] the nature of the optical field by all the coherence functions of all orders (μ, ν) . In many cases such a characterization is possible but there are two drawbacks to this scheme. First, not all the moments may exist; they may diverge! Second, even if they all exist there is no guarantee that the density matrix is completely characterized by them. There is a more satisfactory method of characterizing a general density matrix in terms of the characteristic functional [16]. We shall content ourselves here with noting this circumstance, but this topic is discussed in more detail in a later section.

Time dependence of the higher order coherence functions

It is instructive to rewrite the coherence function of order (μ, ν) in the following form

$$(4.12) \quad \Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu, y_1, \dots, y_\nu)$$

$$= \sum_{\substack{\alpha_1 \dots \alpha_\mu \\ \beta_1 \dots \beta_\nu}} u_{\alpha_1}^*(x_1) \dots u_{\alpha_\mu}^*(x_\mu) u_{\beta_1}(y_1) \dots u_{\beta_\nu}(y_\nu) \times$$

$$\times \text{tr} \left\{ a_{\beta_1} \dots a_{\beta_\nu} \rho a_{\alpha_1}^\dagger \dots a_{\alpha_\mu}^\dagger \right\}$$

It follows the *unequal time coherence functions can be computed if the equal time coherence functions are known at all sets of points.*

"Mode-pure" fields of illumination.

There exists a special class of fields in which there is excitation only in one mode, say $u(x)$: in this case the coherence functions factorize in the form [17]

$$\begin{aligned}
 (4.13) \quad & \Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu, y_1, \dots, y_\nu) \\
 &= u^*(x_1) \dots u^*(x_\mu) u(y_1) \dots u(y_\nu) \operatorname{tr} \left\{ a^\nu \rho a^{\dagger \mu} \right\} \\
 &= g^{(\mu, \nu)} u^*(x_1) \dots u^*(x_\mu) u(y_1) \dots u(y_\nu)
 \end{aligned}$$

for all μ, ν . In this case all the optical discriminants $\Delta^{(\mu, \nu)}(x; y)$ vanish and all the coherence indices $S^{(\mu, \nu)}(x; y)$ are unimodular. In fact

$$S^{(\mu, \nu)}(x, y) = \frac{u^*(x_1)}{u(x_1)} \dots \frac{u^*(x_\mu)}{u(x_\mu)} \frac{u(y_1)}{u^*(y_1)} \dots \frac{u(y_\nu)}{u^*(y_\nu)}$$

Such an optical field may be called "mode pure" and may be characterized by specifying the mode function $u(x)$ and the sequence of moments $g^{(\mu, \nu)}$ if they exist. These fields are treated in greater detail in Section 5.

Unimodular coherence indices and the factorizability of coherence functions

We now wish to show that the converse of this result holds in a slightly weakened form. We wish to show that if $|S^{(\mu, \nu)}(x; y)| = 1$, so that $\Delta^{(\mu, \nu)}(x; y) = 0$, then the (μ, ν) order coherence function must factorize. We first of all observe* that since $\Delta^{(\mu, \nu)}(x; y)$ is the trace of a nonnegative matrix, whenever Δ vanishes we must have

$$D \rho D^\dagger = 0 \quad \text{where, as before,}$$

$$D(x, y) = C_t^{(\mu)}(x) - \frac{\Gamma^{(\mu, \nu)}(y, x)}{\Gamma^{(\nu, \nu)}(y, y)} C_t^{(\nu)}(y).$$

But the density operator ρ is a nonnegative completely continuous operator with the eigenvector decomposition

$$(4.14) \quad \rho = \sum_\lambda \rho(\lambda) \psi(\lambda) \psi^\dagger(\lambda); \quad \rho(\lambda) \geq 0.$$

Hence

$$D \rho D^\dagger = 0 = \sum_\lambda \rho(\lambda) \{D \psi(\lambda)\} \{D \psi(\lambda)\}^\dagger.$$

* The technique of proof follows Titulaer and Glauber [17].

This is possible if and only if

$$(4.15) \quad D \psi(\lambda) = 0, \quad \rho(\lambda) \neq 0$$

for all the eigenvectors $\psi(\lambda)$ with nonzero eigenvalue $\rho(\lambda)$. It then follows that all these states $\psi(\lambda)$ have the property

$$(4.16) \quad G_t(x) \psi(\lambda) = \frac{\Gamma(y, x)}{\Gamma(y, y)} G_t^{\dagger}(y) \psi(\lambda)$$

This implies, in particular, the relations

$$G_t(x) \rho = \frac{\Gamma(y, x)}{\Gamma(y, y)} G_t^{\dagger}(y) \rho$$

$$\rho G_t^{\dagger}(y) = \frac{\Gamma(x', y)}{\Gamma(x', x')} \rho G_t^{\dagger}(x')$$

which imply

$$\begin{aligned} \text{tr} \left\{ G_t(x) \rho G_t^{\dagger}(y') \right\} &= \frac{\Gamma(x, y)}{\Gamma(y, y)} \text{tr} \left\{ G_t(y) \rho G_t^{\dagger}(y') \right\} \\ &= \frac{\Gamma(x', y')}{\Gamma(x', x')} \text{tr} \left\{ G_t(x) \rho G_t^{\dagger}(x') \right\}. \end{aligned}$$

In other words, we have

$$\begin{aligned} \Gamma^{(\mu, \nu)}(x, y) &= \frac{\Gamma^{(\mu, \nu)}(x, y')}{\Gamma^{(\nu, \nu)}(y', y')} \Gamma^{(\nu, \nu)}(y', y) \\ &= \frac{\Gamma^{(\mu, \nu)}(x', y)}{\Gamma^{(\mu, \nu)}(x', x')} \Gamma^{(\mu, \mu)}(x, x') \end{aligned}$$

Making use of the definition of the coherence indices we can rewrite these relations in the form

$$(4.17) \quad \begin{aligned} S^{(\mu, \nu)}(x, y) &= S^{(\mu, \nu)}(x, y') S^{(\nu, \nu)}(y', y) \\ &= S^{(\mu, \mu)}(x, x') S^{(\mu, \nu)}(x', y) \end{aligned}$$

These equations are valid if and only if

$$S^{(\mu, \mu)}(x, x') = \sigma^{(\mu)}(x') / \sigma^{(\mu)}(x)$$

$$S^{(\nu, \nu)}(y, y') = \sigma^{(\nu)}(y') / \sigma^{(\nu)}(y)$$

$$S^{(\mu, \nu)}(x, y) = \sigma^{(\nu)}(y) / \sigma^{(\mu)}(x).$$

Now the complex conjugation relations

$$\{S^{(\mu, \nu)}(x, y)\}^* = S^{(\nu, \mu)}(y, x)$$

and other similar relations are automatically satisfied. Note that $\sigma^{(\mu)}(x)$ and $\sigma^{(\nu)}(y)$ are at the present time two independent functions in the relevant variables. They must however satisfy some additional constraints by the Field operators which we now proceed to derive as follows: We can substitute these results back into the equation

$$G_t^{(\mu)}(x) \rho = \frac{\Gamma^{(\nu, \mu)}(y, x)}{\Gamma^{(\nu, \nu)}(y, y)} G_t^{(\nu)}(y) \rho$$

to derive

$$(4.18) \quad G_t^{(\mu)}(x) \rho = \alpha^{(\mu)}(x) \mathcal{B} \rho$$

where \mathcal{B} is a suitable operator of μ^{th} degree in annihilation operators and $\alpha(x)$ is defined by

$$\alpha^{(\mu)}(x) = \sigma^{(\mu)}(x) \sqrt{\Gamma^{(\mu, \mu)}(x, x)}$$

Now consider

$$\begin{aligned} A(\xi) A(x_1) \dots A(x_\mu) \rho &= A(x_1) \dots A(x_\mu) A(\xi) \rho \\ &= A(\xi) \alpha^{(\mu)}(x_1 \dots x_\mu) \mathcal{B} \rho \\ &= A(x_\mu) \alpha^{(\mu)}(x_1 \dots \xi) \mathcal{B} \rho \end{aligned}$$

Hence,

$$(4.19) \quad A(\xi) \mathcal{B} \rho = \frac{\alpha^{(\mu)}(x_1 \dots x_{\mu-1} \xi)}{\alpha^{(\mu)}(x_1 \dots x_{\mu-1} x_\mu)} A(x_\mu) \mathcal{B} \rho.$$

It follows that $\alpha^{(\mu)}(x)$ must factorize in the form :

$$\alpha^{(\mu)}(x_1 \dots x_\mu) = c \cdot \alpha^{(\mu-1)}(x_1 \dots x_{\mu-1}) u(x_\mu)$$

with a suitable mode function $u(x_\mu)$ and c is a suitable constant. But by reason of symmetry of $\alpha^{(\mu)}(x)$ in the μ variables it follows that we can factorize $\alpha^{(\mu)}(x)$ in the form

$$\alpha^{(\mu)}(x) = c_\mu u(x_1) \dots u(x_\mu)$$

where c_μ is some constant depending on the nature of ρ . Similarly

$$\mathcal{C}^{(\nu)}(y) \rho = \alpha^{(\nu)}(y) \mathcal{C} \rho ; \quad \alpha^{(\nu)}(y) = c_\nu u(y_1) \dots u(y_\nu)$$

where c_ν is some other constant. We have therefore established the result that the coherence functions of orders (μ, ν) , (μ, μ) , (ν, ν) factorize in the fashion

$$(4.20) \quad \Gamma^{(\mu, \nu)}(x \dots x_\mu ; y_1 \dots y_\nu) = g^{(\mu, \nu)} u^*(x_1) \dots u(y_\nu)$$

$$\Gamma^{(\mu, \mu)}(x_1 \dots x_\mu ; x_1' \dots x_\mu') = g^{(\mu, \mu)} u^*(x_1) \dots u(x_\mu')$$

$$\Gamma^{(\nu, \nu)}(y_1' \dots y_\nu' ; y_1 \dots y_\nu) = g^{(\nu, \nu)} u^*(y_1') \dots u(y_\nu)$$

with

$$g^{(\mu, \nu)} = c_\mu^* c_\nu^* \text{tr} \left\{ \mathcal{C} \rho \mathcal{B}^\dagger \right\}$$

$$g^{(\mu, \mu)} = |c_\mu|^2 \text{tr} \left\{ \mathcal{B} \rho \mathcal{B}^\dagger \right\} ; \quad g^{(\nu, \nu)} = |c_\nu|^2 \text{tr} \left\{ \mathcal{C} \rho \mathcal{C}^\dagger \right\}.$$

The density matrix of states with a unimodular coherence index

We can now determine the most general density matrix ρ consistent with such a factorization. For this purpose we note that for every one of the states $\psi(\lambda)$ for which $\rho(\lambda) \neq 0$

$$A(x) \psi = \alpha^{(\mu)}(x) \mathcal{B}^{(\mu)} \psi$$

Hence the states $\mathcal{B} \psi$ have the property that

$$(4.21) \quad A(\xi) \mathcal{B}^{(\mu)} \psi = u(\xi) a \mathcal{B}^{(\mu)} \psi$$

where

$$(4.22) \quad a = \int d\xi u^*(\xi) A(\xi)$$

Hence only one mode is excited in the state $\mathcal{B}^{(\mu)} \psi$, namely that one with the mode function $u(\xi)$. We represent this state symbolically by an operator function of a^\dagger acting in the vacuum state Ω :

$$(4.23) \quad \mathcal{B}^{(\mu)} \psi = F_1(a^\dagger) \Omega$$

But we are interested in the state ψ and not $\mathcal{B} \psi$. Now $\mathcal{B}^{(\mu)}$ is an operator of degree μ in all the annihilation operators which would annihilate every state with excitation of all modes together less than μ quanta. Hence the general solution to ψ is given by

$$(4.24) \quad \psi = N \left\{ F(a^\dagger) + P^{(\mu)}(A^\dagger) \right\} \Omega$$

where $P(A^\dagger)$ is a polynomial functional of the creation part of the field of degree less than μ , $F(a^\dagger)$ is a suitable function of the creation operator a^\dagger associated with the mode $u(\xi)$ and N is a suitable normalization constant. The general density matrix ρ is hence given by

$$\rho = \sum_{\lambda} \rho(\lambda) N^2(\lambda) \left\{ F_{\lambda}(a^\dagger) + P_{\lambda}(A^\dagger) \right\} \Omega \Omega^\dagger \left\{ F_{\lambda}(a^\dagger) + P_{\lambda}(A^\dagger) \right\}^\dagger$$

with $\rho(\lambda)$ nonnegative and summing to unity. By considering the corresponding relations obtained by starting from the relation

$$A(y) \psi = \alpha^{(\nu)}(y) \mathcal{B}^{(\nu)} \rho$$

we note that the degree of the polynomial functionals $P^{(\mu)}(A^\dagger)$ must be less than ν also. Hence the polynomials may be denoted as $P(A^\dagger)(A^\dagger)$ with degree less than $\min.(\mu, \nu)$. With this understanding we have the general expression for the density matrix consistent with unimodular coherence index of order (μ, ν) :

$$(4.25) \quad \rho = \sum_{\lambda} \rho(\lambda) \left\{ F_{\lambda}(a^\dagger) + P_{\lambda}(A^\dagger) \right\} \Omega \Omega^\dagger \left\{ F_{\lambda}(a^\dagger) + P_{\lambda}(A^\dagger) \right\}^\dagger$$

with the normalization constants absorbed into the definition of the operators F_{λ} , P_{λ} . We have thus proved the following result:

THEOREM 2. If the coherence index of order (μ, ν) is unimodular then the coherence functions of orders (μ, μ) , (μ, ν) and (ν, ν) factorize in the form (4.20). The most general density matrix consistent with such a constraint is given by (4.25).

In view of this structure of the density matrix it is straightforward to observe that all coherence functions $\Gamma^{(\mu', \nu')}$ of order (μ', ν') must also factorize provided only the mode-pure components $F_{\lambda} \Omega$ of the states Ψ contribute to these coherence functions. This happens when $\max(\mu', \nu') \geq \min(\mu, \nu)$. We have thus the following result :

COROLLARY 2.1. If the coherence index of order (μ, ν) is unimodular, the coherence functions of order (μ', ν') will factorize provided

$$\max(\mu', \nu') \geq \min(\mu, \nu).$$

The coherence indices $S^{(\mu', \nu')}(x, y)$ will not in general be unimodular but will be constants (independent of x, y) and all coherence functions of order (μ', ν') , $\mu' \geq \min(\mu, \nu)$ will be unimodular. In particular, if the second order coherence index is unimodular all coherence functions factorize ; none of the coherence function of order (μ, μ) can vanish unless the coherence functions of orders (μ, ν) vanish for $\nu < \mu$.

Mode-pure illumination and factorizable coherence functions in classical theory

It is now possible to see that the results of the classical theory of high order coherence functions are stronger*. In this case the result

$$\Gamma^{(\mu, \nu)}(x_1 \dots x_{\mu}, y_1 \dots y_{\nu}) = g^{(\mu, \nu)} u^*(x_1) \dots u^*(x_{\mu}) u(y_1) \dots u(y_{\nu})$$

implies that the probability measure for the stochastic variable $V(\xi)$ must be concentrated on such fields for which

$$V(\xi) = \frac{u(\xi)}{u(0)} V(0)$$

with the variable $V(0)/u(0)$ being distributed as a single stochastic variable. Hence it follows that the coherence function of order (μ, ν) must be of the form

$$g^{(\mu, \nu)} u^*(x_1) \dots u^*(x_{\mu}) u(y_1) \dots u(y_{\nu})$$

with $g^{(\mu, \nu)}$ being given by the expectation value of the quantity $\left(V^*(0)/u^*(0) \right)^{\mu} \times \left(V(0)/u(0) \right)^{\nu}$. It is worthwhile pointing out that the first part of Theorem (2) and Corollary (2.1) are valid for the classical theory ; the corollary can be strengthened to state that if the coherence index of any nontrivial order (i.e. $\mu + \nu > 0$) is unimodular, then the coherence functions of all orders (μ, μ) are unimodular and coherence indices of all orders $S^{(\mu, \nu)}(x; y)$ will be constants independent of x, y . In other words, the field is "mode-pure."

* Compare Matha [14].

5. COHERENCE FUNCTIONS OF MODE-PURE FIELDS OF ILLUMINATION

We now return to the quantum theory of mode-pure optical fields. We have already seen that such a field has a characteristic mode $u(\xi)$ (which includes the specification of the space and polarization dependence) and a sequence of the "normal-ordered" moments $g^{(\mu, \nu)}$ given by the expectation values [17]

$$(5.1) \quad g^{(\mu, \nu)} = \text{tr} \left\{ a^\nu \rho a^{\dagger \mu} \right\}$$

If the density matrix ρ is expressed in the occupation number representation

$$\rho = \sum_{m, n} \rho(m, n) |m\rangle \langle n|$$

these moments are given by

$$(5.2) \quad \sum_{l \geq \mu, \nu} \rho(l - \mu, l - \nu) \frac{l!}{\sqrt{(l - \mu)!(l - \nu)!}} = g^{(\mu, \nu)}$$

which need not exist unless $\rho(m, n)$ fall off sufficiently fast for $m, n \rightarrow \infty$. As an example we may consider a simple case

$$\rho(m, n) = \begin{cases} 0 & m \neq n \\ \frac{6}{\pi^2 n^2} & m = n \end{cases}$$

In this case $g^{(\mu, \nu)}$ vanishes for $\mu \neq \nu$ while $g^{(0, 0)}$ and $g^{(1, 1)}$ exist all higher moments $g^{(\mu, \mu)}$, $\mu \geq 2$ diverge.

Properties of the sequence of moments for "intense" illuminations

We consider first the wavefields for which $g^{(1, 1)}$ is larger than some integer $\mu + \lambda$. Since the mode function $u(x)$ is normalized, the magnitude of $g^{(1, 1)}$ is a direct measure of the intensity of illumination. We may refer to these cases with $g^{(1, 1)}$ larger than some integer $\mu + \lambda$ as fields of intense illumination. There is no guarantee that the quantities $g^{(\mu, \nu)}$ contain increasing numbers as μ, ν increase. For the special states with the number of photons limited, sufficiently high moments vanish [17]. Let us first consider the some simple inequalities satisfied by the moments $g^{(\mu, \nu)}$. Since the density matrix ρ is nonnegative

$$\text{tr} \left\{ (a^\mu - K a^\nu) \rho (a^{\dagger \mu} - K^* a^{\dagger \nu}) \right\} \geq 0$$

so that [17]

$$(5.2) \quad g^{(\mu, \mu)} g^{(\nu, \nu)} \geq |g^{(\mu, \nu)}|^2$$

Now if

$$g^{(\mu, \mu)} = \text{tr} \left\{ a^\mu \rho a^{\dagger \mu} \right\} = 0$$

then the nonnegative matrix $a^\mu \rho a^{\dagger \mu}$ itself should vanish; and hence, all the moments $g^{(\lambda, \lambda)}$, $\lambda > \mu$ would also vanish. It then follows the vanishing of $g^{(\lambda, \lambda)}$ implies the vanishing of all $g^{(\mu, \nu)}$ with $\max(\mu, \nu) \geq \lambda$. Since the $g^{(\mu, \nu)}$ are defined as the expectation values of normal ordered operators, as μ, ν increase the moments may increase or decrease. In particular $g^{(\mu, \mu)}$ may vanish for large μ , or it may increase as fast as $\mu!$. It is therefore difficult to exhibit simple inequalities satisfied by these moments [18]* Consider, however, the identity

$$(a^\dagger)^{\mu+1} a^{\mu+1} + \mu (a^\dagger)^\mu a^\mu = (a^\dagger)^\mu a^\mu a^\dagger a$$

Since the probability densities $\rho(n, n)$ are nonnegative we have the inequality

$$\text{tr} \left\langle \rho \left\{ (a^\dagger)^\mu a^\mu \right\} \right\rangle \cdot \text{tr} \left\langle \rho \left\{ a^\dagger a \right\} \right\rangle \leq \text{tr} \left\langle \rho \left\{ (a^\dagger)^\mu a^\mu \right\} \left\{ a^\dagger a \right\} \right\rangle$$

Hence we get

$$\mu g^{(\mu, \mu)} + g^{(\mu+1, \mu+1)} \geq g^{(\mu, \mu)} g^{(1, 1)}$$

so that

$$(5.3) \quad \frac{g^{(\mu+1, \mu+1)}}{g^{(\mu, \mu)}} \geq (g^{(1, 1)} - \mu)$$

Consequently, if $g^{(1, 1)} > \mu + \lambda$ we have

$$g^{(\mu+\lambda, \mu+\lambda)} \geq g^{(\mu, \mu)} \frac{(g^{(1, 1)} - \mu)!}{(g^{(1, 1)} - \mu - \lambda)!} \geq g^{(\mu, \mu)} \left(g^{(1, 1)} - \mu - \frac{\nu}{2} \right)^\nu$$

where the factorials are defined for nonintegral arguments in terms of the gamma function. We could transcribe the last inequality in a slightly different form:

$$(5.4) \quad \frac{g^{(\mu+\lambda, \mu+\lambda)}}{g^{(\mu-\lambda, \mu-\lambda)}} \geq \left(g^{(1, 1)} - \mu \right)^{2\lambda}; \quad g^{(1, 1)} \geq \mu + \lambda$$

* Inequalities that hold for states with positive definite diagonal representation distributions (according to the optical equivalence theorem stated in [18]) are discussed by Titulaer and Glauber [17].

Thus as long as $g^{(1,1)}$ is sufficiently large compared with unity, the moments increase in the beginning. Using the relation

$$\frac{g^{(\mu+\lambda, \mu+\lambda)}}{g^{(\mu, \mu)}} \geq \left(g^{(1,1)} - \mu - \frac{\nu}{2}\right)^\nu; \quad g^{(1,1)} \geq \mu + \lambda$$

we could derive

$$\frac{\left(g^{(\mu+\lambda, \mu+\lambda)}\right)^2}{g^{(\mu, \mu)} g^{(\nu, \nu)}} \geq \left(g^{(1,1)} - \mu - \frac{\nu}{2}\right)^\nu \left(g^{(1,1)} - \nu - \frac{\mu}{2}\right)^\mu; \quad (g^{(1,1)} \geq \lambda + \mu)$$

which could be transcribed into the form

$$(5.5) \quad \frac{\left\{g^{(\mu+\lambda, \mu+\lambda)}\right\}^2}{g^{(2\mu-\lambda, 2\mu-\lambda)} g^{(2\lambda-\lambda, 2\lambda-\mu)}} \geq \left(g^{(1,1)} - \frac{3}{2}\mu\right)^{2\lambda-\mu} \cdot \left(g^{(1,1)} - \frac{3}{2}\lambda\right)^{2\mu-\lambda}$$

provided

$$g^{(1,1)} \geq \mu + \lambda; \quad 2\lambda \geq \mu \geq \frac{\lambda}{2}$$

Absolute inequalities for the sequence of moments

A large number of inequalities between the moments $g^{(\mu, \mu)}$ which hold absolutely (i.e., independent, say, of the magnitude of $g^{(1,1)}$) may be derived by relating those moments to the expectation values of the powers of the number operator $a^\dagger a$. For this purpose we make use of the identity

$$(5.6) \quad (a^\dagger a)^\lambda = \sum_{\mu=0}^{\lambda} d_{\lambda \mu} (a^\dagger)^\mu a^\mu; \quad d_{\lambda \mu} = \sum_{l=0}^{\mu} \frac{(-1)^{\mu-l} l^\lambda}{l! (\mu-l)!}$$

This identity can be established by repeated use of the commutation relation using mathematical induction but it can be established by equating the expectation values of both sides in the state $|z\rangle$ which is an eigenstate of the annihilation operator and comparing coefficients of powers of $|z|^2$. We have

$$\begin{aligned} \langle z | (a^\dagger a)^\lambda | z \rangle &= \sum_{N=0}^{\infty} N^\lambda \frac{|z|^{2N}}{N!} e^{-|z|^2} \\ &= \sum_{r=0}^{\infty} (-1)^r \frac{|z|^{2\gamma}}{\gamma!} \sum_{N=0}^{\infty} \frac{|z|^{2N}}{N!} N^\lambda \\ &= \sum_{\mu=0}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^l |z|^{2\mu}}{l! (\mu-l)!} (\mu-l)^2 \end{aligned}$$

while

$$\sum_{\mu} d_{\lambda\mu} \langle z | (a^\dagger)^\mu a^\mu | z \rangle = \sum_{\mu} d_{\lambda\mu} |z|^{2\mu}$$

Comparing coefficients of powers of $|z|^2$ gives the required result. Let us now write

$$(5.7) \quad h_\lambda = \sum_{\mu} d_{\lambda\mu} g^{(\mu, \mu)}$$

Since these modified moments h_λ are moments of powers of an operator, a multitude of inequalities satisfied by them may be derived.

Let us first show that h_λ form a nondecreasing sequence. We note first of all that $h_0=1$ and unless all $h_\lambda = 0, \lambda > 0$ none of the h_λ can vanish. We shall exclude the trivial case $h_\lambda = 0, \lambda > 0$ and assume from now on that all the h_λ are positive. Since the operator

$$\left\{ (a^\dagger a)^\lambda - K (a^\dagger a)^\mu \right\}^2$$

is positive for all values of k , we have the inequality

$$(5.8) \quad \left\{ h_{\lambda\lambda} - K h_{\lambda+\mu} \right\} \geq K \left\{ h_{\lambda+\mu} - K h_{\lambda\mu} \right\}$$

$$h_{\mu+1} - h_\mu \geq h_\mu - h_{\mu-1}$$

But we can also see that $h_\lambda \geq h_1$ (though h_1 may or may not be larger than $h_0 = 1$) since

$$h_\lambda = \sum_{N=1}^{\infty} p(N) N^\lambda \geq \sum_{N=1}^{\infty} p(N) N \geq h_1$$

where $p(n) \geq 0$ are the nonnegative density matrix elements $p(n, n)$.

Hence, it follows not only that h_λ increase with λ but

$$h_{\lambda+1} \geq (\lambda - 1) (h_\lambda - h_1).$$

There is a slightly stronger result

$$h_{\lambda+1} \geq \max \left\{ (\lambda - \mu) (h_{\mu+1} - h_\mu) \right\}; \mu \geq 1$$

Thus the moments cannot decrease; they must increase unless $h_\lambda - h_1 = 0$ and this occurs if only if $p(n) = 0, n > 1$. In this case all the h_λ except h_0 are equal.

We can show that if the moments are not equal so that $h_2/h_1 > 1$ then they must increase even faster. By a suitable application of the Schwartz identity we can see that

$$\text{tr} \left\{ \rho (a^\dagger a)^{2\lambda} \right\} \text{tr} \left\{ \rho (a^\dagger a)^{2\mu} \right\} \geq \left(\text{tr} \left\{ \rho (a^\dagger a)^{\lambda + \mu} \right\} \right)^2$$

This could be transcribed into the inequality

$$(5.10) \quad \frac{h_{\mu + \lambda}}{h_\mu} \geq \frac{h_\mu}{h_{\mu - \lambda}}$$

and, in particular,

$$\frac{h_{\mu + 1}}{h_\mu} \geq \frac{h_\mu}{h_{\mu - 1}}$$

Since $h_2/h_1 > 1$ it follows that $h_{\mu + 1}$ is bounded below by $b_1 (h_2/h_1)^{\mu - 1}$

$$(5.11) \quad h_{\mu + 1} \geq \frac{h_2^{\mu - 1}}{h_1^{\mu - 2}}$$

The equality obtains only if all the ratios $h_{\mu + 1}/h_\mu$ for $\mu > 0$ are equal; and corresponds to a system for which $p(0)$ and any one $p(n)$ are nonvanishing; i.e. all the eigenstates of the density matrix are superpositions of the vacuum state and states with n_0 photons where n_0 is integral and given by h_2/h_1 . Thus in general the higher moments must increase at least exponentially with the exponent $\log(h_2/h_1)$. But, they may increase faster, even becoming infinite. If h_μ is infinite, $h_{\mu + \lambda}$ must also be infinite.

6. THE CHARACTERISTIC FUNCTIONAL OF A MODE-PURE FIELD OF ILLUMINATION

In case any of the moments become infinite, the characterization of the mode-pure optical field in terms of higher order coherence functions becomes inadequate. On the other hand if the moments all exist and increase not too fast the moments can be used to characterize the statistical state of the system. (We shall see below the precise meaning of the term "not too fast.") This circumstance is of course not peculiar to mode-pure fields, since the higher order coherence functions need not exist for a general optical field in which components with large photon numbers contribute. A somewhat different method of characterizing the statistical wave field is available [16] which employs a modified version of the method of generating functionals. We discuss the specialization of this method to mode-pure fields here; a discussion of the general theory is given in the next section.

The characteristic function for single mode excitation

For an arbitrary complex number

$$(6.1) \quad \zeta = \frac{1}{2} (\xi + i \eta)$$

consider the unitary operator

$$(6.2) \quad \begin{aligned} W(\xi, \eta) = W(\zeta) &= \exp \left\{ i (\zeta^* a + \zeta a^\dagger) \right\} = e^{-\frac{1}{2} |\zeta|^2} e^{i \zeta a^\dagger} e^{+i \zeta^* a} \\ &= e^{+\frac{1}{2} |\zeta|^2} e^{+i \zeta^* a} e^{+i \zeta a^\dagger} \end{aligned}$$

and its expectation value [16]

$$\omega(\xi, \eta) = \omega(\zeta) = \text{tr} \left[\rho \exp \left\{ i (\zeta^* a + \zeta a^\dagger) \right\} \right]$$

Since $W(\zeta)$ is unitary the function $\omega(\zeta)$ always exists and has an absolute magnitude no greater than unity :

$$(6.3) \quad |\omega(\zeta)| = \left| \text{tr} \left\{ \rho W(\zeta) \right\} \right| \leq \text{tr} \left\{ \rho |W(\zeta)| \right\} \leq \text{tr} \rho = 1$$

We can make use of the completeness identity [12]

$$\frac{1}{\pi} \int d^2 z |z\rangle \langle z| = 1$$

Satisfied by the over-complete family of eigenstates of the annihilation operator the re-express $\omega(\zeta)$ in the form [16, 19]

$$(6.4) \quad \begin{aligned} \omega(\zeta) &= \frac{1}{\pi} e^{\frac{1}{2} |\zeta|^2} \int d^2 z \text{tr} \left\{ \rho e^{i \zeta^* a} |z\rangle \langle z| e^{i \zeta a^\dagger} \right\} \\ &= \frac{1}{\pi} e^{\frac{1}{2} |\zeta|^2} \int d^2 z \langle z| e^{i \zeta a^\dagger} \rho e^{i \zeta^* a} |z\rangle \\ &= e^{\frac{1}{2} |\zeta|^2} \int d^2 z \exp \left\{ i (\zeta z^* + \zeta^* z) \right\} \rho_A(z) \end{aligned}$$

where

$$\rho_A(z) = \frac{1}{\pi} \langle z| \rho |z\rangle = \frac{1}{\pi} e^{-|z|^2} \sum_{m, N} \frac{\rho(m, N) z^*{}^m z^N}{\sqrt{m! N!}}$$

is uniquely defined as soon as the density matrix is given. We note that the quantity $\rho_A(z)$ is real, nonnegative, and bounded above by $1/\pi$. If we write $z = x + iy$, and consider

$$\rho_A(z) = \rho_A(x, y)$$

as a function of two variables x, y we can show [16] that $\rho_A(x, y)$ is both integrable and square integrable. From the expression for $\rho_A(x, y)$ in terms of the elements of the density matrix, it can be seen that $\rho_A(x, y)$ is the boundary value of an entire analytic function $\rho_A(z_1, z_2)$ in two complex variables z_1, z_2 when z_1 and z_2 become real. This property holds not only for the function associated with the density matrix but for the function associated with any bounded operator.

In terms of $\rho_A(z)$ we define the characteristic functional

$$(6.5) \quad N(\zeta) = e^{\frac{1}{2}|\zeta|^2} \omega(\zeta) = \text{tr} \left\{ e^{i\zeta^* a} \rho e^{i\zeta a^\dagger} \right\}.$$

This quantity always exists for all ζ . Since $\omega(\zeta)$ is bounded above by unity it follows that

$$(6.6) \quad |N(\zeta)| \leq e^{\frac{1}{2}|\zeta|^2}$$

For $\zeta = 0$, $\omega(\zeta)$ is unity and hence

$$(6.7) \quad N(0) = 1.$$

Since every bounded operator function of a and a^\dagger can be expressed as a linear combination (and approximated arbitrarily closely by a finite linear combination) of the unitary operators $W(\zeta)$, it follows that the expectation value of any bound operator can be determined as soon as $\omega(\zeta)$ or $N(\zeta)$ is given [16].

In general there is no guarantee that $N(\zeta)$ or $\omega(\zeta)$ possesses power series expansions in ζ and ζ^* in the neighbourhood of the origin. However, if it does then we can write down the Maclaurin expansion

$$N(\zeta) = e^{\frac{1}{2}|\zeta|^2} \sum_{m, N} \frac{\partial^m + N}{(\partial \zeta^*)^m (\partial \zeta)^N} \omega(\zeta) \Big|_{\zeta=0} \cdot \frac{(\zeta^*)^m \zeta^N}{m! N!}$$

which can then be rewritten in the form

$$(6.8) \quad N(\zeta) = \sum_{\mu, \nu} g^{(\mu, \nu)} \frac{(i\zeta)^\mu (i\zeta^*)^\nu}{\mu! \nu!}$$

where

$$g^{(\mu, \nu)} = \text{tr} \left(a^\nu \rho a^{\dagger \mu} \right)$$

Hence the moments $g^{(\mu, \nu)}$, when they exist, can be computed as soon as the characteristic function $N(\zeta)$ of the mode-pure field is given.

*Determination of the density matrix from the characteristic function:
The diagonal representation*

We now consider the question of whether the state is completely specified by the characteristic function $N(\zeta)$. For this purpose let us consider the integral equation [16, 20]

$$(6.9) \quad N(\zeta) = \int d^2 z \varphi(z) \exp \{i(\zeta^* z + \zeta z^*)\}$$

By rewriting this equation in terms of the real and imaginary parts of z and ζ we may exhibit this integral equation as a real double Fourier transformation:

$$(6.10) \quad N(\xi, \eta) = \iint dx dy \varphi(x, y) \exp \{2i(x\xi + y\eta)\}$$

One might attempt to find solutions $\varphi(x, y)$ to this integral equation in a suitable class of functions, say absolutely square integrable functions. We have seen above that $|N(\zeta)|$ is bounded by $\exp(\frac{1}{2}|\zeta|^2)$; there is no guarantee that considered as a function of ξ, η it does not increase or increases no faster than a polynomial. It is possible to construct density matrices for which $N(\xi, \eta)$ increases exponentially or even as the exponential of a quadratic form. (It can increase no faster, since we have already seen above an upper bound for $|N(\zeta)|$). Therefore if we need a general solution to the quantity $\varphi(z)$ we may have to go outside the class of square integrable functions, or even of tempered distributions. But we can always consider $\varphi(z) = \varphi(x, y)$ as a distribution [16] that maps the exponential function into $N(\zeta) = N(\xi, \eta)$ in the form

$$(6.11) \quad \exp \{2i(x\xi + y\eta)\} \xrightarrow{\mathcal{P}} N(\xi, \eta)$$

There is an alternate characterization of the distribution φ which makes use of the analytic function $\rho_A(z) = \rho_A(x, y)$. Since using the identity

$$e^{i\zeta a^\dagger} e^{i\zeta^* a} = e^{-\frac{1}{2}|\zeta|^2} e^{i\zeta^* a} e^{i\zeta a^\dagger}$$

we have established

$$\begin{aligned} \exp(\frac{1}{2}|\zeta|^2) \int d^2 z \exp \{i(\zeta z^* + \zeta^* z)\} \rho_A(z) \\ = \omega(\zeta) = e^{-\frac{1}{2}|\zeta|^2} N(\zeta) \end{aligned}$$

we may rewrite the defining integral equation for the distribution in the form [16, 19]

$$(6.12) \quad \int d^2 z \varphi(z) \exp \{i(\zeta^* z + \zeta z^*)\} = \exp(-\frac{1}{2}|\zeta|^2) \int d^2 z \rho_A(z) \exp \{i(\zeta^* z + \zeta z^*)\}$$

or, in terms of real double Fourier integrals, in the form

$$(6.13) \quad \int \int dx dy \varphi(x, y) \exp \{2i(x\xi + y\eta)\} \\ = \exp(\xi^2 + \eta^2) \int \int dx dy \rho_A(x, y) \exp \{2i(x\xi + y\eta)\}$$

The class of distributions φ so defined is not one of the more familiar classes of distributions. However, the mapping (6.11) is a rigorous mathematical characterization of the distribution [21]. We note, in passing, that the distribution φ so defined is essentially unique: if two distributions φ_1 and φ_2 mapped all exponentials according to (6.11) into the same function $N(\xi, \eta)$ then it will map all bounded functions of x and y into zero. Hence, as long as we are interested in *bounded* functions of x, y (or, more generally, in tempered distributions in x, y) we may put $\varphi_2 - \varphi_1 = 0$.

The importance of the distribution in the theory of mode-pure fields is that we can reconstruct the density matrix for the system in terms of φ by the following diagonal representation [18]:

$$\rho = \int d^2z \varphi(z) |z\rangle\langle z|$$

More precisely, what we can show is the following: Consider the operator $P[\varphi]$ defined by

$$(6.14) \quad P[\varphi] = \int d^2z \varphi(z) |z\rangle\langle z|$$

Then all unitary operators $W(\zeta)$ and hence all bounded operator functions $F(a, a^\dagger)$ have the same expectation value in the state ρ as in the state $\rho[\varphi]$. To prove this result we compute the expectation value of $W(\zeta)$ to obtain

$$\begin{aligned} \text{tr} \{ (P[\varphi] W(\zeta)) \} &= \int d^2z \varphi(z) e^{-\frac{1}{2}|\zeta|^2} \text{tr} (e^{i\zeta^* a} |z\rangle\langle z| e^{i\zeta a^\dagger}) \\ &= \exp(-\frac{1}{2}|\zeta|^2) \int d^2z \varphi(z) \langle z | \exp(i\zeta a^\dagger) \exp(i\zeta^* a) | z \rangle \\ &= \exp(-\frac{1}{2}|\zeta|^2) \int d^2z \varphi(z) \exp\{i(\zeta^* z + \zeta z^*)\} = \omega(\zeta). \end{aligned}$$

This verifies our assertion that $\rho[\varphi]$ gives the same expectation value as ρ for all unitary operators $W(\zeta)$ and hence for all bounded operators. Even if we consider *unbounded* operators, as long as they are self-adjoint (or, more generally, a complex linear combination of self-adjoint operators) *every* spectral projection will be mapped into zero. We may then identify $\rho[\varphi]$ to coincide with ρ and thus complete the question of the uniqueness of the state once the characteristic functional is specified. We have demonstrated elsewhere [18] the fundamental role played by the diagonal representation in coherence theory.

Alternate method of reconstruction of the density matrix from the characteristic function

In view of the importance of the question of reconstruction of the density matrix from the characteristic functional for the state we give an entirely different proof [16] of the existence of a unique density matrix for every characteristic functional. This method exploits the completeness identity satisfied by the over-complete family of states and the property of $\rho_A(x, y)$ being the boundary value of an entire analytic function $\rho_A(z_1, z_2)$ of two complex variables. We proceed as follows: Given the characteristic functional $N(\zeta)$ consider the integral equation

$$(6.16) \quad \int \int dx dy A(x, y) \exp \{2i(x\xi + y\eta)\} = \exp(-\xi^2 - \eta^2) N(\xi, \eta)$$

where x, y, ξ, η are all considered as real variables. Since $|N(\xi, \eta)| = |N(\zeta)|$ is bounded above by $\exp\{+\frac{1}{2}(\xi^2 + \eta^2)\}$ it follows that the right-hand side of (6.16) has an absolute magnitude which is bounded by $\exp\{-\frac{1}{2}(\xi^2 + \eta^2)\}$. Consequently the integral equation has the solution

$$A(x, y) = \frac{1}{\pi^2} \int \int d\xi d\eta \exp\{-2i(x\xi + y\eta)\} \exp(-\xi^2 - \eta^2) N(\xi, \eta)$$

which has the bound

$$|A(x, y)| \leq \frac{1}{\pi^2} \int \int d\xi d\eta \exp\{-\frac{1}{2}(\xi^2 + \eta^2)\} = \frac{1}{\pi}.$$

Further the function $A(x, y)$ is the boundary value of an analytic function $A(z_1, z_2)$ defined by

$$A(z_1, z_2) = \frac{1}{\pi^2} \int \int d\xi d\eta \exp(-\xi^2 - \eta^2) N(\xi, \eta) \exp\{-2i(z_1\xi + z_2\eta)\}$$

If the imaginary parts of z_1 and z_2 are denoted by ω_1 and ω_2 we have

$$\begin{aligned} |A(z_1, z_2)| &\leq \frac{1}{\pi^2} \int \int d\xi d\eta \exp\{-\frac{1}{2}(\xi^2 + \eta^2)\} \exp\{2(\omega_1\xi + \omega_2\eta)\} \\ &= \frac{1}{\pi} \exp\{-2(\omega_1^2 + \omega_2^2)\} \end{aligned}$$

Hence $A(z_1, z_2)$ is bounded for all (finite) complex values of z_1 and z_2 and defines an entire analytic function of order two in either variable. Such a function has an absolutely convergent double power series expansion.

$$A(z_1, z_2) = \sum_{m=0, N=0}^{\infty} \rho(m, N) z_1^m z_2^N.$$

Let us now construct the operator

$$\begin{aligned}
 (6.17) \quad \rho [A] &= \frac{1}{\pi} \int \int d^2 \zeta_1 d^2 \zeta_2 A \left(\frac{\zeta_1 + \zeta_2}{2}, i \frac{\zeta_1 - \zeta_2}{2} \right) \\
 &\quad \times \exp \left\{ + \zeta_1^* \zeta_2 - \frac{1}{2} |\zeta_1|^2 - \frac{1}{2} |\zeta_2|^2 \right\} |\zeta_1^* \rangle \langle \zeta_2| \\
 &= \frac{1}{\pi} \int \int d^2 \zeta_1 d^2 \zeta_2 A \left(\frac{\zeta_1^* + \zeta_2}{2}, i \frac{\zeta_1^* - \zeta_2}{2} \right) \\
 &\quad \times \exp \left\{ + \zeta_1^* \zeta_2 - \frac{1}{2} |\zeta_1|^2 - \frac{1}{2} |\zeta_2|^2 \right\} |\zeta_1 \rangle \langle \zeta_2|
 \end{aligned}$$

We now assert that $\rho [A]$ so constructed is the required density matrix corresponding to the characteristic functional $N(\zeta)$.

Proof of the reconstruction theorem

As a first step in proving this result we observe that the Kernel

$$(6.18) \quad k(\zeta, z) = \frac{1}{\pi} \exp(\zeta z^* - |z|^2)$$

has the remarkable property that for any entire function $f(z)$ we have the identity

$$(6.19) \quad \int d^2 z k(\zeta, z) f(z) = f(\zeta).$$

This result [22] can be verified by writing

$$f(z) = f(r e^{i\theta}) = \sum_{l=0}^{\infty} f_l r^l e^{il\theta}$$

and evaluating the integral

$$\begin{aligned}
 \int d^2 z k(\zeta, z) f(z) &= \frac{1}{\pi} \sum_{l=0}^{\infty} f_l \int r dr \int d\theta \exp\{-r^2 + \zeta r e^{-i\theta}\} r^l e^{il\theta} \\
 &= \frac{1}{\pi} \sum_{l=0}^{\infty} r dr \int d\theta e^{-r^2} r^l e^{il\theta} \sum_{k=0}^{\infty} \frac{(\zeta r)^k}{k!} e^{-ik^2} \\
 &= \int 2r dr e^{-r^2} \sum_{l=0}^{\infty} \frac{(\zeta r^2)^l}{l!} f_l = \sum_{l=0}^{\infty} f_l \zeta^l \frac{1}{l!} \int du u^l e^{-u} \\
 &= \sum_{l=0}^{\infty} f_l \zeta^l = f(\zeta)
 \end{aligned}$$

Using the definition of $\rho [A]$ we can construct

$$\begin{aligned}
 (6.20) \quad \langle z | \rho [A] | z \rangle &= \frac{1}{\pi} \int \int d^2 \zeta_1 d^2 \zeta_2 A \left(\frac{1}{2} (\zeta_1^* + \zeta_2), \frac{1}{2} i (\zeta_1^* - \zeta_2) \right) \\
 &\times \exp \{ \zeta_1^* \zeta_2 - \frac{1}{2} |\zeta_1|^2 - \frac{1}{2} |\zeta_2|^2 \} \langle z | \zeta_1 \rangle \langle \zeta_2 | z \rangle \\
 &= \frac{1}{\pi} \int \int d^2 \zeta_1 d^2 \zeta_2 A \left(\frac{1}{2} (\zeta_1^* + \zeta_2), \frac{1}{2} i (\zeta_1^* - \zeta_2) \right) \\
 &\times \exp \{ \zeta_1^* \zeta_2 - |\zeta_1|^2 - |\zeta_2|^2 - |z|^2 + z^* \zeta_1 + z \zeta_2^* \}.
 \end{aligned}$$

Using the fact that fact $A \left(\frac{1}{2} (\zeta_1^* + \zeta_2), \frac{1}{2} i (\zeta_1^* - \zeta_2) \right)$ is analytic and entire in ζ_1^* and ζ_2 and the property of the Kernels $K(z, \zeta_2)$ we have successively

$$\begin{aligned}
 (6.21) \quad \langle z | \rho [A] | z \rangle &= \int d^2 \zeta_2 \exp (-|z|^2 - |\zeta_2|^2 + z \zeta_2^*) \\
 &\times \frac{1}{\pi} \int d^2 (\zeta_1^*) \exp (-|\zeta_1^*|^2 + z^* \zeta_1) \left\{ A \left(\frac{\zeta_1^* + \zeta_2}{2}, \frac{i(\zeta_1^* - \zeta_2)}{2} \right) \exp (\zeta_2 \zeta_1^*) \right\} \\
 &= \pi \cdot \frac{1}{\pi} \int d^2 \zeta_2 \exp (-|\zeta_2|^2 + z \zeta_2^*) \\
 &\times \left\{ A \left(\frac{z^* + \zeta_2}{2}, \frac{i(z^* - \zeta_2)}{2} \right) \exp (\zeta_2 z^* - |z|^2) \right\} \\
 &= \pi A \left(\frac{z^* + z}{2}, \frac{i(z^* - z)}{2} \right)
 \end{aligned}$$

Hence $\rho [A]$ has the "diagonal" matrix elements

$$\langle z | \rho [A] | z \rangle = \pi A(x, y) \quad ; \quad z = x + iy$$

Hence the characteristic functional corresponding to $\rho [A]$ is given by

$$\begin{aligned}
 M(\zeta) &= \text{tr} \{ \exp (i \zeta^* a) \rho (A) \exp (-i \zeta a^\dagger) \} \\
 &= \exp (+|\zeta|^2) \text{tr} \{ \exp (i \zeta a^\dagger) \rho [A] \exp (+i \zeta^* a) \} \\
 &= \exp (|\zeta|^2) \frac{1}{\pi} \int d^2 z \text{tr} \{ |z\rangle \langle z| \exp (i \zeta a^\dagger) \rho [A] \exp (i \zeta^* a) \} \\
 &= \exp (|\zeta|^2) \frac{1}{\pi} \int d^2 z \langle z | \rho [A] | z \rangle \exp \{ i (\zeta z^* + \zeta^* z) \} \\
 &= \exp (|\zeta|^2) \iint dx dy A(x, y) \exp \{ 2i(x\xi + y\eta) \} = N(\zeta)
 \end{aligned}$$

Hence $\rho [A]$ has the given functional as the characteristic function. We have therefore proved the following basic result :

THEOREM 3 Given the characteristic function $N(\zeta)$ we can compute all the normal ordered moments, whenever they exist, by a power series expansion of $N(\zeta)$ in the neighbourhood of the origin. The characteristic function $N(\zeta)$ always exists for all complex values of ζ and is bounded $\exp(\frac{1}{2} |\zeta|^2)$. Given a characteristic function which satisfies this requirement we can reconstruct the density matrix ρ by the formula

$$(6.14) \quad \rho[\varphi] = \int d^2 z \phi(z) |z\rangle \langle z|$$

with $\varphi(z)$ defined as the distribution solution of the integral equation

$$(6.22) \quad \int d^2 z \varphi(z) \exp\{i(z\zeta^* + z^*\zeta)\} = N(\zeta).$$

We have the alternate formula

$$(6.17) \quad \rho = \frac{1}{\pi} \int d^2 z \int d^2 \zeta \exp\{z^*\zeta - \frac{1}{2}|z|^2 - \frac{1}{2}|\zeta|^2\} \times \\ A\left(\frac{z^* + \zeta}{2}, \frac{i}{2}(z^* - \zeta)\right) |z\rangle \langle \zeta|$$

where the entire analytic function $A(z_1, z_2)$ of two complex variables is the solution of the integral equation

$$(6.23) \quad A(z_1, z_2) = \frac{1}{\pi^2} \int \int d\xi d\eta \exp(-\xi^2 - \eta^2) N(\xi, \eta) \exp\{-2i(\xi z_1 + \eta z_2)\}$$

7. THE CHARACTERISTIC FUNCTIONAL OF A GENERAL FIELD OF ILLUMINATION

In the previous section we have discussed the method of specifying the statistical state of a mode-pure field in terms of the characteristic function $N(\zeta)$ which is more general than the specification in terms of the higher order moments. We now wish to generalize this to arbitrary optical wavefields. We shall start with the quantum theory rule [10] for computing the higher order coherence functions :

$$\Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu; y_1, \dots, y_\nu) = \text{tr} \left\{ V(y_1) \dots V(y_\nu) \rho V^\dagger(x_1) \dots V^\dagger(x_\mu) \right\}$$

where $V(x)$ is the annihilation part of the field

$$V(x) = \sum_{\alpha} a_{\alpha} u_{\alpha}(x).$$

We shall continue to use the convention that for the transverse vector field, we will use the label x to stand for the space coordinates and polarization labels but not the time parameters. Of course for a transverse vector field, the mode functions $u_\alpha(x)$ would be transverse vector functions. We have also seen that the coherence functions could be rewritten in the form

$$(7.1) \quad \Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu; y_1, \dots, y_\nu) = \sum_{\substack{\alpha_1 \dots \alpha_\mu \\ \beta_1 \dots \beta_\nu}} u_{\alpha_1}^*(x_1) \dots u_{\alpha_\mu}^*(x_\mu) u_{\beta_1}(y_1) \dots u_{\beta_\nu}(y_\nu) \times \\ \text{tr} \left\{ a_{\beta_1} \dots a_{\beta_\nu} \rho a_{\alpha_1}^\dagger \dots a_{\alpha_\mu}^\dagger \right\}$$

Definition of the characteristic functional

Given the quantities

$$\text{tr} \left\{ a_{\beta_1} \dots a_{\beta_\nu} \rho a_{\alpha_1}^\dagger \dots a_{\alpha_\mu}^\dagger \right\} = g^{(\mu, \nu)}(\alpha; \beta)$$

we could construct the auxiliary quantities

$$(7.2) \quad g^{(\mu, \nu)}(\zeta) = \frac{i^{\mu+\nu}}{\mu! \nu!} \zeta_{\alpha_1} \dots \zeta_{\alpha_\mu} \zeta_{\beta_1}^* \dots \zeta_{\beta_\nu}^* \text{tr} \left\{ a_{\beta_1} \dots a_{\beta_\nu} \rho a_{\alpha_1}^\dagger \dots a_{\alpha_\mu}^\dagger \right\} \\ = \frac{i^{\mu+\nu}}{\mu! \nu!} \text{tr} \left\{ [B(\zeta)]^\nu \rho [B^\dagger(\zeta)]^\mu \right\}$$

where

$$B(\zeta) = \sum_{\alpha} \zeta_{\alpha}^* a_{\alpha} \quad ; \quad B^\dagger(\zeta) = \sum_{\alpha} \zeta_{\alpha} a_{\alpha}^\dagger.$$

In terms of these define the functional $\chi(\zeta)$ by the construction

$$(7.3) \quad \chi(\zeta) = \sum_{\mu, \nu=0}^{\infty} g^{(\mu, \nu)}(\zeta) = \text{tr} \left\{ e^{iB(\zeta)} \rho e^{iB^\dagger(\zeta)} \right\}$$

We can write down the expression for $B(\zeta)$ in terms of a function $\zeta(x)$ defined by

$$(7.4) \quad \zeta(x) = \sum_{\beta} \zeta_{\beta} u_{\beta}(x)$$

in the form

$$\begin{aligned}
 B(\zeta) &= \sum_{\alpha, \beta} \zeta_{\beta}^* \delta_{\alpha\beta} a_{\alpha} \\
 &= \sum_{\alpha, \beta} \zeta_{\beta}^* \int dx u_{\beta}^*(x) u_{\alpha}(x) a_{\alpha} \\
 &= \int dx \zeta^*(x) V(x)
 \end{aligned}$$

If the sequence of numbers ζ_{α} satisfy $\sum_{\alpha} |\zeta_{\alpha}|^2 < \infty$

then we have, automatically,

$$(7.5) \quad \|\zeta\| = \int |\zeta(x)|^2 dx < \infty.$$

We shall restrict attention to such sequences ζ_{α} and such functions $\zeta(x)$. We may now write the functional $\chi(\zeta)$ of the sequences $\{\zeta\}$ as a functional of the square integrable function $\zeta(x)$ in the form

$$(7.6) \quad N[\zeta(\cdot)] = \text{tr} \left[\exp \left\{ i \int dy \zeta^*(y) V(y) \right\} \rho \exp \left\{ + i \int dy \zeta(y) V^{\dagger}(y) \right\} \right]$$

This is the characteristic functional of the state [23] specified by the density matrix ρ .

Properties of the characteristic functional; Relation to coherence functions

By virtue of the fact that $\zeta(x)$ is square integrable it is possible to show that the functional $N[\zeta(x)]$ always exists. For this purpose we observe that

$$(7.7) \quad W[\zeta(\cdot)] = \exp \left\{ i \int dx \left[\zeta^*(x) V(x) + \zeta(x) V^{\dagger}(x) \right] \right\}$$

is a unitary operator-valued functional. We can deduce by the use of the commutation relations

$$[V(x), V^{\dagger}(y)] = \delta(x, y)$$

that

$$\begin{aligned}
 (7.8) \quad \exp \left\{ + i \int dx \zeta(x) V^{\dagger}(x) \right\} \exp \left\{ i \int dy \zeta^*(y) V(y) \right\} \\
 = \exp \left(+ \frac{1}{2} \|\zeta\|^2 \right) W[\zeta(\cdot)]
 \end{aligned}$$

Hence the number-valued functional $N[\zeta(\cdot)]$ is a complex number of absolute value bounded by $\exp \left(- \frac{1}{2} \|\zeta\|^2 \right)$.

If we define the complex number-valued functional

$$\omega[\zeta(\cdot)] = \text{tr} \{ \rho W[\zeta(\cdot)] \}$$

then

$$(7.9) \quad \omega[\zeta(\cdot)] = \exp(-\frac{1}{2} \|\zeta\|^2) N[\zeta(\cdot)]$$

In general there is no guarantee that $N[\zeta(\cdot)]$ can be expanded in terms of multiple integrals over multilinear forms in $\zeta(\cdot)$. But if the functional is an analytic functional, we may evaluate the various coherence functions as the coefficient functions in this expansion. In any case we can formally write down

$$(7.10) \quad N[\zeta(\cdot)] = \sum_{\mu, \nu=0}^{\infty} \frac{i^{\mu+\nu}}{\mu! \nu!} \int dx_1 \dots \int dx_{\mu} \int dy_1 \dots \int dy_{\nu} \Gamma^{(\mu, \nu)}(x_1 \dots x_{\mu}; y_1 \dots y_{\nu}) \times \zeta(x_1) \dots \zeta(x_{\mu}) \zeta^*(y_1) \dots \zeta^*(y_{\nu}).$$

In the differential from this may be written

$$(7.11) \quad \Gamma^{(\mu, \nu)}(x_1, \dots, x_{\mu}, y_1, \dots, y_{\nu}) = (-i)^{\mu+\nu} \frac{\delta^{\mu+\nu} N[\zeta(\cdot)]}{\delta \zeta(x_1) \dots \delta \zeta(x_{\mu}) \delta \zeta^*(y_1) \dots \delta \zeta^*(y_{\nu})}$$

Thus, if we are given the density matrix we can calculate the characteristic functional $N[\zeta(\cdot)]$. The various order coherence functions, *when they exist*, can be obtained from the "power series" expansion of the functional in terms of multiple integrals of multiple products of ζ and ζ^* . But the characteristic function exists for all $\zeta(\cdot)$ with $\|\zeta\| < \infty$, whether the higher coherence functions exist or not.

Reconstruction of the density matrix from the characteristic functional

We can now raise the question of whether the characteristic functional $N[\zeta(\cdot)]$ uniquely determines the statistical state of the system. For this purpose we consider a distribution over the unitary number-valued functionals

$$\exp \left\{ i \int dx [\psi(x) \zeta^*(x) + \psi^*(x) \zeta(x)] \right\}$$

which maps them into $N[\zeta(\cdot)]$. We could write this symbolically as a functional integral:

$$(7.12) \quad \exp \left\{ i \int dx [\psi(x) \zeta^*(x) + \psi^*(x) \zeta(x)] \right\} \xrightarrow{\varphi} \\ = \int d^2[\psi(\cdot)] \varphi[\psi(\cdot)] \exp \left\{ i \int dx [\psi(x) \zeta^*(x) + \psi^*(x) \zeta(x)] \right\} \\ = N[\zeta(\cdot)]$$

where $d^2[\psi(\cdot)]$ is the natural measure in the space of complex functions $\psi(\cdot)$. Then we could write down the functional integral representation for the density matrix in the form

$$(7.13) \quad \rho = \int d^2[\psi(\cdot)] \varphi[\psi(\cdot)] |\psi(\cdot)\rangle \langle \psi(\cdot)|$$

where the (state) vector-valued functional $|\psi(\cdot)\rangle$ is defined by the simultaneous eigenstate of the annihilation operators according to the equation

$$\int dx f^*(x) V(x) |\psi(\cdot)\rangle = \int dx f^*(x) \psi(x) |\psi(\cdot)\rangle$$

In the above equations we have used the *diagonal representation* for the field density matrix [18].

The method of characteristic functionals that we have outlined here realizes its proper role when one considers the equations of motion for higher order coherence functions; we hope to return to this topic in another paper.

8. MIXTURES OF MODE-PURE ILLUMINATIONS

In an earlier section we have discussed mode-pure fields for which the various coherence functions factorize. In this case the problem can essentially be reduced to the study of the density matrix of a single oscillator. But in most cases the illumination will not be mode-pure. In the study of the second order coherence functions we saw that they form a convex cone and the generators had coherence functions for which the (equal-time) reduced coherence functions were unimodular. This theorem does not have a direct extension to higher order coherence functions in general. For mode-pure fields the theorem is trivial. In this section we study the mode-mixtures which belong to the convex cone generated by fields almost all whose coherence functions factorize; and the necessary and sufficient condition for a field to belong to this convex cone. As a corollary we will be able to demonstrate that not all fields of illumination can be so constructed.

Coherence functions as infinite partitioned matrices

As a preliminary to the study of these questions we note that an infinite partitioned matrix Γ may be constructed out of the coherence functions in which the (μ, ν) th block of the partition consists of the coherence function

$$\Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu; y_1, \dots, y_\nu) \text{ with the "coordinates" } \{x; y\}$$

acting as continuous matrix indices; sum over repeated matrix indices would then be replaced by integration over these continuous indices. This matrix is hermitian and nonnegative. The first principal block consists of $\Gamma^{(1, 1)} = 1$. Since this matrix is hermitian and nonnegative, every principal submatrix of it must also be hermitian and nonnegative. All those properties are preserved under the operation of homogeneous

convex combinations. If $\Gamma_{(k)}$, $1 \leq k \leq N$ are admissible coherence matrices, then the homogeneous convex combinations

$$(8.1) \quad \Gamma = \sum_k \lambda_k \Gamma_{(k)}, \quad \sum \lambda_k = 1, \quad \lambda_k \geq 0$$

are also admissible coherence matrices. The external elements of this homogeneous convex set are pure states :

$$\Gamma = \psi \psi^\dagger$$

which satisfy

$$\Gamma^2 = \Gamma$$

where ψ is an infinite column vector with appropriate partitions.

It is interesting to note that the characteristic functionals $N[\zeta(\cdot)]$ also constitute a homogeneous convex set. These functionals have also the property of being closed under homogeneous convex [6] combinations

$$(8.2) \quad N = \sum_k \lambda_k N_{(k)}, \quad \sum \lambda_k = 1, \quad \lambda_k \geq 0.$$

The extremal elements correspond to pure states whose characteristic functions satisfy a convolution property (which we shall not discuss here!). Define a column vector $z[\zeta(\cdot)]$ whose n^{th} block entry $0 \leq n \leq \infty$ is given by the n^{th} Kronecker power of $\zeta(\cdot)$ multiplied by $(i^n/n!)$. We may now write

$$(8.3) \quad N = z^\dagger \Gamma z$$

(provided the functional $N[\zeta(\cdot)]$ has a power series expansion) so that the two convex sets are simply related.

From the last expression for $N[\zeta(\cdot)]$ in terms of $z[\zeta(\cdot)]$ for those cases in which $N[\zeta(\cdot)]$ has a power series expansion we may read off the structure of the characteristic functional of a mode-pure field. In this case the coherence functional of order (μ, ν) is the product of the constant $g^{(\mu, \nu)}$ by μ factors of $u^*(\cdot)$ and ν factors of $u(\cdot)$. The multiplication by z^\dagger and z provides μ factors of ζ and ν factors of ζ^* as well as a factor $i^{\mu + \nu} / \mu! \nu!$. Hence we finally obtain

$$(8.4) \quad N[\zeta(\cdot)] = \sum_{\mu, \nu=0}^{\infty} \frac{g^{(\mu, \nu)} i^{\mu + \nu}}{\mu! \nu!} \left[\int dx \zeta^*(x) u(x) \right]^\nu \left[\int dy \zeta(x) u^*(x) \right]^\mu \\ = \sum_{\mu, \nu} \frac{g^{(\mu, \nu)} i^{\mu + \nu}}{\mu! \nu!} (\omega^*)^\nu \omega^\mu$$

where

$$(8.5) \quad \omega = \int d x \zeta^*(x) u(x).$$

It thus reduces to a function of the quantity ω . While the above derivation assumed that $N[\zeta(\cdot)]$ had a power series expansion the result that $N[\zeta(\cdot)]$ is a function of ω above is true for all the general mode pure fields.

Properties of mixtures of mode-pure fields: The auxiliary two-point function

For a general field which is a finite combination of mode-pure fields the characteristic functional would be a homogeneous convex combination of functions of one variable. The question we would like to ask is if it would be possible to find a sequence of orthogonal modes, $k = 1, 2, \dots$ such that

$$N[\zeta(\cdot)] = \sum_k \lambda_k N(\omega_k); \quad \lambda_k \geq 0; \quad \sum_k \lambda_k = 1.$$

Let us assume that we are considering such a state which is a mixture of mode-pure fields. Then the coherence functions of order (μ, ν) could be written in the form:

$$(8.6) \quad \begin{aligned} \Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu; y_1, \dots, y_\nu) &= \sum_k \lambda_k \Gamma_{(k)}^{(\mu, \nu)}(x_1, \dots, x_\mu; y_1, \dots, y_\nu) \\ &= \sum_k \lambda_k g_{(k)}^{(\mu, \nu)} u_k^*(x_1) \dots u_k^*(x_\mu) u_k(y_1) \dots u_k(y_\nu) \end{aligned}$$

By assumption, the modes u_k are orthonormal. For the $\mu = \nu$ coherence function we have the eigenvector decomposition

$$\Gamma(x; y) = \sum_k \gamma_k f_k^*(x) f_k(y)$$

where

$$\gamma_k = \lambda_k g_{(k)}^{(\mu, \mu)}; \quad \int d x f_k^*(x) f_k(x) = \delta_{kk'}$$

$$f_k(x) = u_k(x_1) \dots u_k(x_\mu)$$

Let us now construct the function

$$(8.7) \quad E_k(\eta, \xi) = \int d x_1 \dots \int d x_\mu f_k^*(\xi, x_1, \dots, x_\mu) f_k(\eta, x_1, \dots, x_\mu).$$

Then we can easily see that

$$(8.8) \quad E_k(\eta, \xi) = u_k(\eta) u_k^*(\xi)$$

Characterization of mixtures in terms of the auxiliary two-point function

We have seen above that if a field is composed of the mixture of a set of mode-pure fields with orthonormal mode functions the auxiliary two point functions $E_k(\xi, \eta)$ are of rank one (that is, they have only one eigenvector whose eigenvalue is nonzero). The question is whether this is also a sufficient condition. We shall see below that it is also sufficient.

Let us consider the complete orthonormal set of functions $u_{rk}(\xi)$ of one variable which are eigenfunctions of $E_k(\eta, \xi)$ so that

$$(8.9) \quad \int d\xi E_k(\eta, \xi) u_{rk}(\xi) = e_k(r) u_{rk}(\eta); e_k(r) \geq 0$$

$$\int d\xi u_{rk}^*(\xi) u_{sk}(\xi) = \delta_{rs}$$

$$\sum_r u_{rk}(\xi) u_{rk}^*(\eta) = \delta(\xi, \eta).$$

Because of the normalization of f_k we have $\sum_r e_k(r) = 1$. Define the coefficient functions $h_{rk}(x_2, \dots, x_\mu)$ by

$$f_k(\xi, x_2, \dots, x_\mu) = \sum_r u_{rk}(\xi) h_{rk}(x_2, \dots, x_\mu)$$

so that

$$h_{rk}(x_2, \dots, x_\mu) = \int d\xi u_{rk}(\xi) f_k(\xi, x_2, \dots, x_\mu).$$

Then, the different functions h_{rk}, h_{sk} are orthogonal when $r \neq s$ since

$$\int d x_2 \dots \int d x_\mu h_{rk}^*(x_2, \dots, x_\mu) h_{sk}(x_2, \dots, x_\mu)$$

$$= \int d \xi \int d \eta u_{rk}(\xi) u_{sk}^*(\eta) \int d x_2 \dots \int d x_\mu f_k^*(\xi, x_2, \dots, x_\mu) f_k(\eta, x_2, \dots, x_\mu)$$

$$= \int d \xi \int d \eta u_{rk}(\xi) u_{sk}^*(\eta) E_k(\eta, \xi) = e_k(r) \delta_{rs}.$$

Hence for all the values r for which $e_k(r) = 0$, the coefficient functions $h_{rk}(x_2, \dots, x_\mu)$ vanish: while for all other values of r the $h_{rk}(x_2, \dots, x_\mu)$ for fixed k constitute an orthogonal set (which is in general not complete). For the special case of only a single nonvanishing eigenvalue for $E_k(\eta, \xi)$ that eigenvalue must be unity. Denoting the corresponding eigenfunction by $u_k(\xi)$ we have all the coefficient functions $h_{kr}(x_2, \dots, x_\mu)$

vanishing except the one corresponding to $u_k(\xi)$. We may then write the simple factorization:

$$f_k(\xi, x_2, \dots, x_\mu) = u_k(\xi) h_k(x_2, \dots, x_\mu).$$

But the function f_k is symmetric in its μ arguments and is normalized. Hence we have the explicit expression

$$f_k(x_1, \dots, x_\mu) = u_k(x_1) \dots u_k(x_\mu)$$

This leads finally to the decomposition of the coherence function

$$\Gamma^{(\mu, \mu)}(x; y) = \sum \gamma_k u_k^*(x_1) \dots u_k^*(x_\mu) u_k(y_1) \dots u_k(y_\mu).$$

By virtue of the orthogonality relation for the eigenvector $f_k(x)$ of $\Gamma(x; y)$ which states

$$\int dx f_k^*(x) f_\lambda(x) = \delta_{k\lambda}$$

we have, in the present case

$$\left[\int d\xi u_k^*(\xi) u_\lambda(\xi) \right]^\mu = \delta_{k\lambda}$$

so that the different mode functions $u_k(\xi)$ are orthonormal. Hence we have proved the following result:

The necessary and sufficient condition for the coherence function $\Gamma(x; y)$ of order (μ, μ) to be expressible as a nonnegative linear combination of coherence functions of order (μ, μ) for mode-pure fields is that if we write the eigenvector decomposition of $\Gamma(x; y)$ in the form

$$\Gamma(x; y) = \sum \gamma_k f_k^*(x) f_k(y); \quad \gamma_k \geq 0$$

then the auxiliary functions

$$\int dx_2 \dots \int dx_\mu f_k^*(\xi, x_2, \dots, x_\mu) f_k(\eta, x_2, \dots, x_\mu) = E_k(\eta; \xi)$$

must be of the form

$$(8.8) \quad E_k(\eta, \xi) = u_k(\eta) u_k^*(\xi)$$

When this condition is satisfied one has

$$(8.9) \quad f_k(x_1, \dots, x_\mu) = u_k(x_1) \dots u_k(x_\mu)$$

with an orthonormal set of mode functions $u_k(\xi)$.

Some theorems on mixtures of mode-pure illuminations

The above result can be used to lead us to a very remarkable property of fields of illumination which have coherence functions of any order (μ, μ) , $\mu \geq 2$, that can be expressed as a nonnegative linear combination of mode pure fields. For recognizing this property let us define the annihilation and creation operators

$$(8.10) \quad \begin{aligned} a_\alpha &= \int d\xi A(\xi) u_\alpha^*(\xi) \\ a_\alpha^\dagger &= \int d\xi A^\dagger(\xi) u_\alpha(\xi) \end{aligned}$$

which satisfy

$$[a_\alpha, a_\beta^\dagger] = \delta_{\alpha\beta}$$

and where $u_\alpha(\xi)$ are mode functions associated with the (μ, μ) order coherence function of this class of fields of illumination. Then we have by virtue of the orthonormality of the $u_\alpha(\xi)$ the result

$$(8.11) \quad \text{tr} \left\{ a_{\alpha_1} a_{\alpha_\mu} \rho a_{\beta_1}^\dagger \dots a_{\beta_\mu}^\dagger \right\} = \begin{cases} \gamma_k & \text{if } \alpha_1 = \alpha_2 \dots = \alpha_\mu = \beta_1 \dots = \beta_\mu = k \\ 0 & \text{otherwise.} \end{cases}$$

We note, in passing, that

$$(8.12) \quad \gamma_k = \text{tr} \left\{ a_k^\mu \rho (a_k^\dagger)^\mu \right\}$$

Of course for many values of k , γ_k vanishes but not all γ_k can vanish unless the (μ, μ) order coherence function itself vanishes identically. By choosing $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2, \dots, \alpha_\mu = \beta_\mu$ we can see that we must have identically for every eigenvector ψ of the density matrix ρ the relation

$$a_{\alpha_1} \dots a_{\alpha_\mu} \psi = \begin{cases} (a_k)^\mu \psi & \alpha_1 = \alpha_2 \dots = \alpha_\mu = k \\ 0 & \text{otherwise} \end{cases}$$

Now consider the matrix

$$\sum_{\alpha_{\mu+1}} u_{\alpha_{\mu+1}}(\xi) A_{\alpha_{\mu+1}} a_{\alpha_1} \dots a_{\alpha_\mu} \psi \psi^\dagger.$$

This must vanish, by virtue of the symmetry in the indices and the above result unless all the indices are the same so that

$$A(\xi) a_{\alpha_1} \dots a_{\alpha_\mu} \psi = \begin{cases} u_k(\xi) (a_k)^\mu \psi; & \alpha_1 = \dots = \alpha_\mu = k \\ 0 & \text{otherwise} \end{cases}$$

Hence we have

$$A(\xi) A(x_1) \dots A(x_\mu) = \sum_k u_k(\xi) u_k(x_1) \dots u_k(x_\mu) (a_k)^{\mu+1} \psi.$$

It follows that provided $\max(\lambda, \nu) \geq \mu$

$$\Gamma^{(\lambda, \nu)}(x_1, \dots, x_\lambda; y_1, \dots, y_\nu) = \sum_k \gamma_k^{(\lambda, \nu)} u_k^*(x_1) \dots u_k^*(x_\lambda) u_k(y_1) \dots u_k(y_\nu)$$

where

$$(8.13) \quad \gamma_k^{(\lambda, \nu)} = \text{tr} \left\{ (a_k)^\lambda \rho (a_k^\dagger)^\nu \right\}.$$

The general solutions to the density matrix ρ and the eigenvectors ψ of the density matrix ρ are given by

$$(8.14) \quad \rho = \sum_k \gamma_k \rho_k; \quad \rho_k = \sum_r k(r) \psi_{kr} \psi_{kr}^\dagger$$

$$\psi_{kr} = N \left\{ F_{(kr)} (a_k^\dagger) + P_{(kr)} (A^\dagger) \right\}$$

where the $F_{(kr)}$ are arbitrary operator-valued functions of the creation operator for the mode k . We have thus proved the following result.

THEOREM 4. If the coherence function of order (μ, μ) coincides with the positive linear sum of the (factorizable) coherence functions of a number of mode-pure fields (for $\mu \geq 2$) then all the coherence functions of orders (λ, ν) also factorize in this fashion whenever $\max(\lambda, \nu) \geq \mu \geq 2$. The density matrix of such a system is the positive linear combination of density matrices each of which has eigenvectors which correspond to states having components with arbitrary excitations in a distinguished mode k and a component with less than μ photons in all the modes together.

The method of proof breaks down for the case of $\mu = 1$; this case has to be dealt with separately.

In the above discussion we have considered the coherence function of order (μ, μ) . But we know that any principal submatrix of the infinite partitioned matrix Γ is

hermitian and nonnegative. Further, for a density matrix which is a mixture of modepure density matrices we can make a decomposition

$$(8.15) \quad \Gamma(x; y) = \|\Gamma^{(\mu, \nu)}(x; y)\| = \sum_k \gamma_k F_k(y) F_k^\dagger(x)$$

where F is a column matrix $\|F\|$ whose elements are the functions $f_k(y_1, \dots, y_\mu)$ with

$$f_k(y_1, \dots, y_\mu) = c_{\mu k} u_k(y_1) \dots u_k(y_\mu)$$

where $c_{\mu k}$ are constants. If we now construct the auxiliary function (1×1 matrix)

$$E_k(\eta, \xi) = \sum_\mu \int dx_2 \dots \int dx_\mu f_k^*(\xi, x_2, \dots, x_\mu) f(\eta, x_2, \dots, x_\mu)$$

we get

$$(8.16) \quad E_k(\eta, \xi) = u_k(\eta) u_k^*(\xi) \sum_\mu |c_{\mu k}|^2.$$

Conversely we may show that if $E_k(\eta, \xi)$ has only one vector corresponding to a nonzero eigenvalue then $f_k(\xi, x_1, \dots, x_\mu)$ should factorize. To show this we define

$$f_{kr}(x_2, \dots, x_\mu) = \int d\xi u_{kr}^*(\xi) f_k(\xi, x_2, \dots, x_\mu).$$

Then we can show that as long as $E_k(\eta, \xi)$ has only one eigenvector $u_k(\xi)$ having nonzero eigenvalue that all the $f_{kr}(x_2, \dots, x_\mu)$ except the one corresponding to $u_{kr}(\xi) = u_k(\xi)$ vanish; hence we may write

$$f_k(\xi, x_2, \dots, x_\mu) = f_k(x_2, \dots, x_\mu) u_k(\xi)$$

and hence

$$(8.17) \quad f_k(x_1, \dots, x_\mu) = c_{\mu k} u_k(x_1) \dots u_k(x_\mu).$$

Thus the necessary and sufficient condition for an expression for $\|\Gamma(x; y)\|$ to be a positive linear combination of the corresponding set of coherence matrices is that every $E_k(\eta; \xi)$ have only one eigenvector $u_k(\xi)$ having a nonzero eigenvalue.

Provided $\max(\mu, \nu) \geq 2$ we can proceed further to show that

$$(8.18) \quad \left\| \Gamma^{(\mu', \nu')}(x; y) \right\| = \sum_k \gamma_k \left\| c_k^{(\mu', \nu')} u_k^*(x_1) \dots u_k^*(x_{\mu'}) u_k(y_1) \dots u_k(y_{\nu'}) \right\|$$

with

$$c_k^{(\mu', \nu')} = c_{\mu' k} c_{\nu' k}$$

for $\max(\mu', \nu') \geq \min(\mu, \nu)$. To prove this result we can proceed in exactly the same fashion as in the previous case of (μ, μ) order coherence. We have thus proved the following result:

COROLLARY 4.1 The results of theorem (4) continue to be true given that $E_k(\eta, \xi)$ for any principal minor $\|\Gamma^{(\mu, \nu)}\|$ of the matrix of coherence functions has a single eigenvector with nonzero eigenvalues with the understanding that

$$\max(\mu', \nu') \geq \max(\mu, \nu) \geq 2.$$

COROLLARY 4.2 These results could be transcribed into the language of characteristic functionals. Theorem (4) then asserts that if the coherence function of order (μ, μ) is a positive linear sum of the mode-pure coherence functions of order (μ, μ) then, the characteristic functional may be written in the form

$$(8.19) \quad N[\zeta(\cdot)] = \frac{\sum_k \gamma_k N_k[\zeta(\cdot)] + P[\zeta(\cdot)]}{\sum_k \gamma_k + P[0]}$$

where $P[\zeta(\cdot)]$ is a polynomial functional of $\zeta(x)$ of degree less than μ .

9. ENUMERATION OF EXCITED MODES FROM SECOND ORDER COHERENCE FUNCTION

The study of mode-mixed fields of illumination in the previous section concentrated on a special class of fields which could be viewed essentially as positive linear sums of mode-pure fields (except for certain component states with less than μ photons). This required a special property of the coherence functions; we have found that this property could be expressed in terms of certain auxiliary functions $E_k(\eta, \xi)$. For a general optical field, these constraints on the higher coherence function are not satisfied. However when one considers the coherence function of order $(1, 1)$ these questions become degenerate: for this coherence function we *always* have the decomposition

$$(9.1) \quad \Gamma^{(1,1)}(x; y) = \sum_k \gamma_k v_k^*(x) v_k(y)$$

and hence it is already in the form of a positive linear sum of factorized contributions. However a careful examination of the proof of theorem (4) show that it *does not* hold in this case; we cannot conclude that all fields of illumination have positive linear combination of mode-pure density matrices

Second order coherence functions and unexcited modes

We can however see that the decomposition (9.1) can be transcribed into

$$(9.2) \quad \text{tr} \left\{ a_k \rho a_\lambda^\dagger \right\} = \gamma_k \delta_{k\lambda}$$

which leads to the result that if $\gamma_k = 0$ for some member $v_\alpha(x)$ of the complete set of mode functions then

$$(9.3) \quad a_\alpha \rho a_\alpha^\dagger = 0$$

since the trace of this nonnegative matrix cannot vanish unless the matrix itself vanishes. If ψ is any eigenvector of ρ it follows that

$$a_\alpha \psi = 0; \quad \gamma_\alpha = 0.$$

Hence the annihilation operator for those modes for which $\gamma_\alpha = 0$ (which are therefore orthogonal to $\Gamma(x, y)$ considered as a function of y) will annihilate the density matrix

$$a_\alpha \rho = 0; \quad \gamma_\alpha = 0.$$

If we now consider

$$\begin{aligned} V(\xi) a_k \rho &= \sum_\alpha u_\alpha(\xi) a_\alpha a_k \rho \\ &= \sum_\alpha u_\alpha(\xi) a_\alpha a_k \rho; \quad \gamma_\alpha \neq 0 \end{aligned}$$

so that all coherence functions can be written in the form :

$$\begin{aligned} (9.4) \quad \Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu, y_1, \dots, y_\nu) \\ = \sum_{\alpha_1, \dots, \alpha_\mu} \sum_{\beta_1, \dots, \beta_\nu} u_{\alpha_1}^*(x_1) \dots u_{\alpha_\mu}^*(x_\mu) u_{\beta_1}(y_1) \dots u_{\beta_\nu}(y_\nu) \cdot g^{(\mu, \nu)}(\alpha; \beta) \\ \gamma_{\alpha_1}, \dots, \gamma_{\alpha_\mu}, \gamma_{\beta_1}, \dots, \gamma_{\beta_\nu} \neq 0 \end{aligned}$$

where

$$(9.5) \quad g^{(\mu, \nu)}(\alpha; \beta) = \text{tr} \left\{ a_{\alpha_1} \dots a_{\alpha_\mu} \rho a_{\beta_1}^\dagger \dots a_{\beta_\nu}^\dagger \right\}.$$

Thus only those modes can be excited in the higher order coherence functions which are spanned by those modes which are excited in the second order coherence function. In particular if only a finite number of modes are excited in the second order coherence function, then all the coherence functions of higher order can be expanded in terms of suitable products (including cross products) of these mode functions. We have therefore proved the following result :

THEOREM 5. Every second order coherence function can be expanded in the form

$$\Gamma(\xi, \eta) = \sum \gamma_k u_k^*(\xi) u_k(\eta)$$

with the $u_k(\xi)$ constituting a complete set of functions. Given an arbitrary coherence function $\Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu; y_1, \dots, y_\nu)$ we will have

$$(9.6) \quad \int d\xi v(\xi) \Gamma^{(\mu, \nu)}(\xi, x_2, \dots, x_\mu; y_1, \dots, y_\nu) \neq 0$$

$$\int d\eta v^*(\eta) \Gamma^{(\mu, \nu)}(x_1, \dots, x_\mu; \eta, y_2, \dots, y_\nu) \neq 0$$

only if we have

$$(9.7) \quad \int d\xi v(\xi) \Gamma^{(\nu, \nu)}(\xi, \eta) \neq 0$$

This result holds even for the coherence functions $\Gamma^{(\nu, \nu)}(\cdot; y)$ and $\Gamma^{(\nu, \nu)}(x; \cdot)$. It is plausible since the second moment is the analogue of the mean square of a random variable; its vanishing requires that with measure unity the random variable must have value zero and hence all the moments must vanish. Even though we are dealing with quantum mechanical normal ordered expectation values, this result continues to be valid.

Structure of the characteristic functional

We can express the content of the theorem (5) in terms of properties of characteristic functionals. We have then the following statement:

COROLLARY 5.1 If the term bilinear in $\zeta(x)$ and $\zeta^*(x)$ in the expansion of the characteristic functional $N[\zeta(\cdot)]$ depends only a projection $\zeta_1(x)$ of the square integrable function $\zeta(x)$ onto a subspace of such functions, then the functional $N[\zeta(\cdot)]$ is independent of the orthogonal projection and is thus a functional of $\zeta_1(x)$ only.

ACKNOWLEDGEMENT

The author is pleased to thank Professor N. Mukunda for illuminating discussions and Professor F. Rohrlich for his constructive criticism of an earlier version of the manuscript. He is particularly grateful to Professor C. L. Metha for his advice and criticism of this paper and for several valuable discussions.

REFERENCES

1. M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, New York, (1964) Second Edition, Chapter X.
2. G. K. Batchelor, *Homogeneous Turbulence*, Cambridge Univ. Press (1953).
3. P. Roman and E. Wolf, *Nuovo Cimento* 17 (1960) 462-477.
4. D. Gabor, *J. Inst. Elec. Engineers* 93 Part III (1946) 429. E. Wolf, *Proc. Roy. Soc. A230*, (1955) 246.
5. C. L. Mehta, E. Wolf and A. P. Balachandran, *J. Math. Phys.* 7, (1966) 133.
6. Karlin, *Mathematical Methods and Theory in Games, Programming and Economics*, Addison-Wesley Inc., New York, 1 (1959).
7. F. Reisz and B. Sz-Nagy, *Fundamental Analysis*, Frederick Ungar Publishing Co., New York (1955).
8. R. Hanburg Brown and R. Q. Twiss, *Proc. Roy. Soc. A* 242 (1957) 300, A 243 (1957) 291.
L. Mandel, *Progress in Optics*, edited by E. Wolf, 2, John Wiley and sons, New York, (1963).
E. Wolf, *Proceedings of the Symposium on Optical Masers*, Polytechnic Press, Brooklyn, New York and John Wiley and Sons, New York, (1963) 29.
9. T. F. Jordan and N. Mukunda, *J. Math. Physics*, 7 (1966) 849.
10. R. J. Glauber, *Phys. Rev. Letters* 10 (1963) 84, *Phys. Rev.* 130 (1963) 2529.
Proceedings of the Symposium on Optical Masers, Polytechnic Press, Brooklyn, New York and John Wiley and Sons, New York (1963) 45.
11. C. L. Mehta and E. C. G. Sudarshan, Time Evolution of Coherent States, *Physics Letters* (in press).
12. J. R. Klauder, *Annals of Physics* 11 (1960) 123, *J. Math. Phys.* 4 (1963) 1055, 1058, 5 (1963) 177.
J. McKenna and J. R. Klauder, *J. Math. Phys.* 5 (1964) 1. See also references 10.
13. R. J. Glauber, *Phys. Letters* (in press)
C. L. Mehta and E. C. G. Sudarshan, Syracuse University preprint submitted to *Physics Letters*.
14. E. C. G. Sudarshan, Coherence Indices in Quantum Optics, Syracuse University preprint, submitted to *Phys. Rev. Letters*.
C. L. Mehta, Degree of Higher Order Optical Coherence, University of Rochester preprint submitted to *J. Math. Phys.*
15. R. J. Glauber. *Phys. Rev.* 130 (1963) 2529.
L. Mandel and E. Wolf, *Rev. Mod. Phys.* 37 (1965) 231.
16. C. L. Mehta and E. C. G. Sudarshan, *Phys. Rev.* 138 (1965) B274.
17. U. M. Titulaer and R. J. Glauber, *Phys. Rev.* 140, (1965) B676.
18. E. C. G. Sudarshan, *Phys. Rev. Letters* 10 (1963) 277.
19. Y. Kano, *J. Math. Phys.* 6 (1965) 1913.
20. J. E. Moyal, *Proc. Cambridge Phil. Soc.* 45, (1948) 99.
21. R. Bonifacio, L. M. Narducci and E. Montaldi, *Phys. Rev. Letters*, 16 (1966) 1125.
22. N. Aronszajn, *Trans. Am. Math. Soc.* 68 (1950) 377.
23. I. E. Segal, *Can. J. Math.* 13, (1961) 1.
E. Hopf, *J. Rat. Mech. and Anal.* 1 (1952) 87.