Structure and representation of correlation functions and the density matrix for a statistical wave field in optics

E. C. G. Sudarshan

Center for Particle Theory, University of Texas, Austin, Texas 78712

N. Mukunda

Centre for Theoretical Studies, Indian Institute of Science, Bangalore, India 560 012

(Received 20 October 1978; revised manuscript received 13 November 1978)

A systematic structure analysis of the correlation functions of statistical quantum optics is carried out. From a suitably defined auxiliary two-point function we are able to identify the excited modes in the wave field. The relative simplicity of the higher order correlation functions emerge as a byproduct and the conditions under which these are mode pure are derived. These results depend in a crucial manner on the notion of coherence indices and of unimodular coherence indices. A new class of approximate expressions for the density operator of a statistical wave field is worked out based on discrete characteristic sets. These are even more economical than the diagonal coherent state representations. An appreciation of the subtleties of quantum theory obtains. Certain implications for the physics of light beams are cited.

INTRODUCTION

Light is essentially quantum mechanical in nature.[1] A light beam is therefore to be represented by an ensemble of quantum states. For a variety of purposes it is convenient to specify a statistical state by giving the set of correlation functions of all possible orders corresponding to that state. It is therefore of value to study representations of the statistical state that make the calculation of correlation functions as direct as possible. One must of course satisfy oneself that a set of (acceptable) correlation functions defines the state uniquely.

By a systematic analysis of the structure and properties of the correlation functions we can identify the "modes," i.e., the natural one-photon wavefunctions, in terms of which the given correlation function, and so the associated states, assume their simplest form. Such an analysis basically brings out the consequences of the positivity properties of the density matrix and the Bose nature of light.[2] Unlike the case, say, of hydrodynamic turbulence,[3] the correlation functions for light obey uncoupled equations of propagation,[4] and the interaction with matter can be treated perturbatively. This fact justifies the structure analysis of statistical states of the free electromagnetic field and also leads to a remarkable simplification in that we basically need study only systems with one degree of freedom.

Correlation functions for the quantized free electromagnetic field are defined as expectation values of normal ordered operator functions of the field.[5] The diagonal coherent state representation[6] of the density matrix therefore leads to expressions for these functions closely mimicking classical correlation functions defined as averages over classical statistical ensembles.[4] The diagonal representation asserts that any density operator for a quantum system with one degree of freedom may be displayed in the form

$$\rho = \int \phi(z)|z\rangle\langle z|d\mu(z),$$

$$d\mu(z) = (1/\pi)d^2z = (1/\pi)d(\mathrm{Re}z)d(\mathrm{Im}z).$$

Here ϕ is a c-number weight function, the coherent states $|z\rangle$ are eigenstates of the annihilation operator for complex eigenvalues z and the integration is over the entire complex plane.[6] These states taken for all z form at the vector space level an overcomplete family[7] permitting a general vector to be expanded linearly in terms of them in more than one way. The diagonal representation for statistical states exploits this overcompleteness to avoid all nondiagonal outer products of coherent states in the above representation and so achieve a form similar to a classical ensemble. The weight function ϕ is generally a distribution belonging to the space Z_2. Nevertheless with this weight function all "normal ordered" correlation functions can be calculated as if the system were classical with the same weights.[6] It must be stressed that we have here a description of the fully quantum mechanical system and not just of its classical limit or semiclassical approximation. The quantum nature of the system lies in the set of weights ϕ to be admitted, and in their properties. Such a description is valuable in any assessment of the information carried by a light beam.

For such a diagonal representation to exist it is necessary that the set of states used be overcomplete. A set of states constituting a complete, but not overcomplete, basis at the vector space level would represent almost all density operators in a nondiagonal form. Since a system with one degree of freedom already calls for a countably infinite set of basis vectors, we do expect every complete as well as overcomplete set to contain infinitely many elements. The coherent states form a two-parameter continuous infinity of states labelled by the real and imaginary parts of z. The question naturally arises as to whether there are "smaller" overcom-

 0022-2488/79/081801-10$01.00

plete families of states which are yet rich enough to permit a diagonal representation for any density operator in terms of them. More precisely, are there smaller families using which we can approximate any given density operator through expressions of the diagonal form to any desired accuracy?

We find that there are indeed such families which need not even be continuously infinite. There is a great variety of countable overcomplete sets of coherent states using any one of which arbitrarily good diagonal approximations to a given density operator can be obtained.[8] It is curious, however, that if we spread these states "as uniformly as possible" over the complex z-plane, then these states must be more dense than one per unit phase cell in most parts of the phase plane. Indeed, Planck's constant fails to provide a natural size for cells in phase space with respect to which the density of such overcomplete sets may be meaningfully stated. This indicates yet another subtle aspect of quantum theory.

The picture of the general density operator for a light beam has then the following structure: The (electric) field is expanded in terms of a complete orthonormal set of natural modes. Under rather general conditions this is a discrete set. All these natural modes will generally contribute to the two-point correlation function. The density operator can now be associated with a multivariate weight function in the excitations of these modes. If we use all the coherent states associated with the annihilation operator corresponding to each mode, we deal with the diagonal representation in its conventional form, and the weight function is in general a distribution in a discrete set of complex variables one per mode. We may alternatively choose a countable overcomplete set of coherent states to go with each mode, rich enough to allow diagonal-type approximations to any statistical state as far as this mode is concerned; thus any state with respect to this mode can be approximated through ensembles over the chosen discrete set of complex eigenvalues for the mode annihilation operator. And a general density operator for the total field system can be approximated arbitrarily closely by ensembles over the collection of discrete sets of eigenvlaues for the annihilation operators of all modes.

The plan of the paper is as follows. Section 1 gives a brief development of the quantum mechanics of a system with one degree of freedom mostly with a view to establishing notation and deriving certain results in a form to be used later. Sections 2 and 3 deal with the analysis of correlation functions and the unravelling of natural modes; both sections are primarily concerned with the descriptions of those parts of a statistical state that correspond to "large photon numbers." After preliminaries, Sec. 2 analyzes this aspect for a state for which it is assumed that a correlation function of some definite order obeys a *condition of coherence*. Section 3 on the other hand analyzes this aspect for a general state by a slightly different means. Section 4 introduces the Weyl operators and the associated expansions of density operators. The problem of the diagonal representation for operators is then identified with the one representing an arbitrary vector state in terms of linear combinations of subsets of coherent states. The basic tool here is the notion of a second Hilbert space made up of operators on the original one. In Sec. 5 we make use of these results and those given in Sec. 1 to derive the general discrete-discrete approximation to the density operator of a light beam. Concluding remarks and some open problems make up Sec. 6.

1. QUANTUM THEORY OF ONE DEGREE OF FREEDOM

A classical one-dimensional harmonic oscillator with unit mass and frequency ω has canonical variables q, p obeying equations of motion

$$\dot{q} = p, \quad \dot{p} = -\omega^2 q. \tag{1.1}$$

Use of the complex canonical variables

$$a = (\omega q + ip)/\sqrt{2\omega}, \quad a^* = (\omega q - ip)/\sqrt{2\omega}, \tag{1.2}$$

allows the solution of Eq. (1.1) to be completely expressed as

$$a(t) = a(0)\exp(-i\omega t), \quad a^*(t) = a^*(0)\exp(i\omega t). \tag{1.3}$$

The instantaneous state of the oscillator is given by the value of the complex dynamical variable a, and as time advances the representative point in the a plane describes a circle. This plane is a rescaled version of the phase space.

For a quantum system with one degree of freedom we have two unbounded Hermitian operators q, p obeying the commutation relation

$$[q,p] = i. \tag{1.4}$$

(Planck's constant has been set equal to unity, and no special symbols such as carets are used to distinguish operators, as there will be no cause for confusion on this account.) This relation can be transcribed in terms of bounded operators by introducing the Weyl families of unitary operators

$$U(\sigma) = \exp(i\sigma q), \quad V(\tau) = \exp(i\tau p), \qquad -\infty < \sigma,\tau < \infty, \tag{1.5}$$

Then Eq. (1.4) is equivalent to

$$\begin{aligned} &U(\sigma)U(\sigma') = U(\sigma+\sigma'),\\ &V(\tau)V(\tau') = V(\tau+\tau'),\\ &U(\sigma)V(\tau) = V(\tau)U(\sigma)\exp(-i\sigma\tau). \end{aligned} \tag{1.6}$$

Setting the frequency ω equal to unity for simplicity, the annihilation operator a and its Hermitian adjoint $a^\dagger$ are defined, following the classical definition (1.2), as

$$a = (q + ip)/\sqrt{2}, \quad a^\dagger = (q - ip)/\sqrt{2}, \tag{1.7}$$

and then the commutation relation (1.4) appears as

$$\{a, a^\dagger\} = 1. \tag{1.8}$$

Coherent states[9] are eigenvectors of the annihilation operator,

$$a|z\rangle = z|z\rangle, \tag{1.9}$$

with the eigenvalue z being any complex number. These states are normalizable, and when normalized to unity their Schrödinger wavefunctions may be taken to be

$$\langle q'|z\rangle = \pi^{-1/4}\exp\left[-\tfrac{1}{2}(q' - z\sqrt{2})^2 - \tfrac{1}{2}z(z^* - z)\right]. \tag{1.10}$$

No two of these states are mutually orthogonal as one has

$$\langle z|z'\rangle = \exp(-\tfrac{1}{2}|z|^2 - \tfrac{1}{2}|z'|^2 + z^*z'). \tag{1.11}$$

The coherent states taken together for all complex z are complete as they furnish a resolution of the identity in the form[10]

$$\psi = \int |z\rangle\langle z|d\mu(z) \equiv (1/\pi)\int |z\rangle\langle z|d^2z. \tag{1.12}$$

Actually, however, they are overcomplete, as one can easily exhibit linear dependences among them in the form of integral relationships. On the other hand any finite number of distinct coherent states are linearly independent. We come back to the use of Eq. (1.12) in a moment.

We can rewrite the operators (1.5) of the Weyl family in the form

$$U(\sigma) = \exp[i\sigma(a + a^\dagger)/\sqrt{2}],$$
$$V(\tau) = \exp[\tau(a - a^\dagger)/\sqrt{2}]. \tag{1.13}$$

This motivates the introduction of the more general Weyl family of unitary operators

$$W(\alpha) = \exp(\alpha a^\dagger - \alpha^* a)$$
$$= U\left(\frac{\alpha - \alpha^*}{i\sqrt{2}}\right) V\left(-\frac{\alpha + \alpha^*}{\sqrt{2}}\right)$$
$$\times \exp\left(\frac{\alpha^{*2} - \alpha^2}{4}\right), \tag{1.14}$$

where α is any complex number. This family will be put to use in Sec. 4. At this point we note the diagonal coherent state matrix elements of these operators,

$$\langle z|W(\alpha)|z\rangle = \exp(-\tfrac{1}{2}|\alpha|^2 + \alpha z^* - \alpha^* z). \tag{1.15}$$

Because of the overcompleteness of coherent states, one expects to be able to "expand" any state $|\psi\rangle$ in terms of them in more than one way. A particular expansion is supplied by the resolution of the identity, Eq. (1.12); one has for any $|\psi\rangle$,

$$|\psi\rangle = \int \langle z|\psi\rangle|z\rangle\, d\mu(z). \tag{1.16}$$

The particular "wavefunction" occurring in this expansion has certain characteristic features. If for convenience we write

$$\langle z^*|\psi\rangle = \exp(-\tfrac{1}{2}|z|^2) f(z), \tag{1.17}$$

then $f(z)$ is an entire analytic function whose behavior for large $|z|$ is controlled by

$$|f(z)| \leqslant \| |\psi\rangle \| \exp(\tfrac{1}{2}|z|^2). \tag{1.18}$$

Using Eq. (1.10) we can relate $f(z)$ to the Schrödinger wavefunction of $|\psi\rangle$ through

$$f(z) = \pi^{-1/4}\exp(-\tfrac{1}{2}z^2)\int_{-\infty}^{\infty} \psi(q')$$
$$\times \exp(-\tfrac{1}{2}q'^2 + q'z\sqrt{2})\, dq'. \tag{1.19}$$

An alternative expansion possibility arises by considering a suitable subset of coherent states. One choice of subset is given by $|i\sqrt{2}r\rangle$ for all real r. If we tentatively write an expansion

$$|\psi\rangle = \pi^{-1/2}\int_{-\infty}^{\infty} v(r)|i\sqrt{2}r\rangle\, dr, \tag{1.20}$$

a means must be found to evaluate the weight function $v(r)$. One way is to take the scalar product of both sides of Eq. (1.20) with another vector of the subset, $|i\sqrt{2}r'\rangle$. Then use of Eq. (1.11), (1.17), (1.19) gives:

$$\int_{-\infty}^{\infty} v(r)\exp[-(r - r')^2]\, dr$$
$$= \pi^{1/4}\int_{-\infty}^{\infty} \psi(q')\exp(-\tfrac{1}{2}q'^2 + 2ir'q')\, dq', \tag{1.21}$$

so that

$$v(r) = \pi^{-1/4}\int_{-\infty}^{\infty} \psi(q')\exp(\tfrac{1}{2}q'^2)\exp(2irq')\, dq'. \tag{1.22}$$

Since the integrand may grow fast at infinity the weight function $v(r)$ may not be an ordinary function but a distribution (in the family Z_2). Alternately we may take the scalar product of the two sides of Eq. (1.20) with a coherent state $|\sqrt{2}s\rangle$, s real, to obtain

$$\langle\sqrt{2}s|\psi\rangle = \pi^{-1/2}\int_{-\infty}^{\infty} v(r)\exp(-s^2 - r^2 + 2irs)\, dr,$$

so that

$$v(r) = \frac{\exp(r^2)}{\sqrt{\pi}}\int_{-\infty}^{\infty} \langle\sqrt{2}s|\psi\rangle\exp(s^2 - 2irs)\, ds. \tag{1.23}$$

The two possible expansions (1.16, 1.20), with characteristically different properties for their integrands, use, respectively, a two-parameter and a one-parameter continuous infinity of coherent states. Instead of such expansions, if we were satisfied with merely being able to approximate arbitrarily closely to any $|\psi\rangle$ through combinations of coherent states, more economical possibilities in terms of the so-called characteristic sets exist.[11] A set S of points in the complex plane is a characteristic set if we can assert that

$$\langle z|\psi\rangle = 0, \quad z\in S \Rightarrow |\psi\rangle = 0. \tag{1.24}$$

A set S with a finite limit point; the set of all real numbers; the set of all imaginary numbers; any sequence $\{z_n\}$ of distinct nonzero complex numbers for which

$$\sum_{n=1}^{\infty} |z_n|^{-2-\epsilon} = \infty \tag{1.25}$$

for some positive ϵ—all these are examples of characteristic sets. Let us restrict ourselves to discrete sets. In terms of the corresponding coherent states $|z_n|$ we could approximate a given $|\psi\rangle$ to arbitrary accuracy: For each $\eta > 0$ we can find an integer $N(\eta)$ and coefficients $b_n(\eta)$ such that

$$\left\| |\psi\rangle - \sum_{n=1}^{N(\eta)} b_n(\eta)|z_n\rangle \right\| < \eta. \tag{1.26}$$

However, for a general vector $|\psi\rangle$ there is no guarantee that there exists a definite set of coefficients b_n such that the sequence of vectors

$$|\psi_N\rangle = \sum_{n=1}^{N} b_n|z_n\rangle$$

forms a Cauchy sequence converging to $|\psi\rangle$.

We conclude this resumé of the properties of coherent states by quoting one more interesting example of a characteristic set and mentioning a property of such sets in general. The example, due to von Neumann and Perelomov,[12] is the set made up of points in the complex plane of the form

$$z = \sqrt{s}(l + im), \quad l,m = 0, \pm 1, \pm 2, \cdots \quad ; s < \pi. \tag{1.27}$$

Thus by essentially picking one coherent state in a phase volume smaller than a unit cell in phase space we get a set with which any $|\psi\rangle$ can be approximated arbitrarily closely. The property of characteristic sets we have in mind is this: if from any such set S any finite number of points are removed, the remaining points still make up a characteristic set.

2. CORRELATION FUNCTIONS AND NATURAL MODES

The free electromagnetic field may be characterized either by the transverse vector potential $\mathbf{A}(\mathbf{r},t)$ or by the electric field $\mathbf{E}(\mathbf{r},t)$, both of which are transverse and gauge invariant. The interaction Hamiltonian of an electron with the field[13] is expressed in terms of $\mathbf{A}$ directly, so that theoretical expressions relating to experiments based on photoelectric detection naturally involve $\mathbf{A}$ also.[14] We shall thus choose $\mathbf{A}$ as the basic variable for defining correlation functions, though one can always pass to the variable $\mathbf{E}$ by time differentiation. The positive frequency part of $\mathbf{A}(\mathbf{r},t)$, sometimes called the analytic signal,[15,4] consists entirely of annihilation operators and has the time dependence

$$\mathbf{A}^{(+)}(\mathbf{r},t) = e^{-i\hat{\omega}t}\mathbf{A}^{(+)}(\mathbf{r},0), \tag{2.1}$$

where $\hat{\omega}$ is the (positive) frequency operator defined by the wave equation

$$(\nabla^2 + \hat{\omega}^2)\mathbf{A}^{(+)}(\mathbf{r},0) = 0, \quad \hat{\omega} = (-\nabla^2)^{1/2}. \tag{2.2}$$

The transversality of $\mathbf{A}^{(+)}$ is expressed by

$$\nabla\cdot\mathbf{A}^{(+)}(\mathbf{r},t) = 0. \tag{2.3}$$

Therefore, there are only two independent components to $\mathbf{A}^{(+)}$. In momentum space these are the components orthogonal to the momentum direction, and may be chosen to be the two circular polarizations denoted by a two-valued polarization index ϵ. Therefore, we may write for the vector potential,

$$\mathbf{A}^{(+)}(\mathbf{r},t) \to V(\mathbf{r},\epsilon,t) \equiv V(x,t). \tag{2.4}$$

When it is not essential to indicate the polarization index ϵ explicitly, it will be combined with the position vector $\mathbf{r}$ into a single symbol x; and formal integration over x will mean a sum over ϵ plus ordinary integration over space. The time t is *not* combined into x in this way. In fact in all the following analysis we shall be concerned only with conditions at one instant of time, and time variables t will be dropped entirely.

Let some statistical state of the field be given, and let the corresponding density operator be ρ. The general (m,n) order correlation function is defined as the expectation value of the normal-ordered product[5] of m negative frequency (creation) field operators and n positive frequency (annihilation) field operators:

$$\begin{aligned}\Gamma^{(m,n)}(x_1,x_2,\ldots,x_m;y_1,y_2,\ldots,y_n) &= \langle V(x_1)^\dagger\cdots V(x_m)^\dagger V(y_1)\cdots V(y_n)\rangle \\ &= \mathrm{Tr}[V(y_1)\cdots V(y_n)\rho V(x_1)^\dagger\cdots V(x_m)^\dagger].\end{aligned} \tag{2.5}$$

Here, m and n are nonnegative integers, and for given values of these, the correlation function depends symmetrically on the m x's and also on the n y's. For conciseness, we may write $\Gamma^{(m,n)}(x;y)$ for the above correlation function; the superscripts imply that x actually stands for m arguments, each consisting of a position vector and a polarization label, etc. As seen by inspection of Eq. (2.5), the following relation holds,

$$\Gamma^{(n,m)}(y,x) = [\Gamma^{(m,n)}(x;y)]^*. \tag{2.6}$$

Further, for $m = n$, the "matrix" $\Gamma^{(m,m)}(x;y)$ with continuous matrix indices x and y is nonnegative,

$$\int dx_1\cdots dx_m\, dy_1\cdots dy_m f(x_1\cdots x_m)^*\Gamma^{(m,m)}(x;y) f(y_1\cdots y_m) \geqslant 0. \tag{2.7}$$

More generally, we may view the entire collection $\Gamma^{(m,n)}(x;y)$ as constituting a giant matrix Γ in which $\Gamma^{(m,n)}$ stands in the (m,n) position when Γ is partitioned.[16] If similarly f is a giant vector at whose mth position stands a symmetric function $f_m(x_1,\ldots,x_n)$ of m x's, then we may identify Γ as a nonnegative matrix according to

$$f^\dagger\Gamma f \equiv \sum_{m,n=0}^{\infty}\int dx_1\cdots dx_m dy_1\cdots dy_n f_m(x_1\cdots x_m) \times {}^*\Gamma^{(m,n)}(x;y) f_n(y_1\cdots y_n) \geqslant 0. \tag{2.8}$$

We now wish to study in more detail some general properties of these correlation functions, especially for "large" orders.

For this purpose we introduce the family of "optical discriminants" $\Delta^{(m,n)}(x;y)$ according to

$$\Delta^{(m,n)}(x;y) \equiv \Gamma^{(m,m)}\Gamma^{(n,n)}(y;y) - \Gamma^{(m,n)}(x;y)\,\Gamma^{(n,m)}(y;x). \tag{2.9}$$

These are real functions of the indicated variables, and the essential point in the definition is that these are nonnegative quantities. In fact, in terms of the operator combination

$$G^{(m,n)}(x;y) = \Gamma^{(n,n)}(y;y)V(x_1)\cdots V(x_m) - \Gamma^{(n,m)}(y;x)V(y_1)\cdots V(y_n), \tag{2.10}$$

we have

$$\Delta^{(m,n)}(x;y)\Gamma^{(n,n)}(y;y) = \mathrm{Tr}(G^{(m,n)}(x;y)\rho G^{(m,n)}(x;y)^\dagger \geqslant 0, \tag{2.11}$$

so that (provided $\Gamma^{(n,n)}$ does not vanish identically) it follows that

$$\Delta^{(m,n)}(x;y) \geqslant 0. \tag{2.12}$$

Another expression of this result can be given in terms of the "coherence indices." The coherence index of order (m,n) is defined to be

$$S^{(m,n)}(x;y) = \Gamma^{(m,n)}(x;y)/[\Gamma^{(m,m)}(x;x)\Gamma^{(n,n)}(y;y)]^{1/2}. \quad (2.13)$$

Then Eq. (2.12) says the same thing as

$$0 \leqslant |S^{(m,n)}(x;y)| \leqslant 1. \quad (2.14)$$

Let us now have a statistical state ρ for which it is given that for a certain pair of integers (m,n), the inequality in Eq. (2.12) becomes an equality for all choices of x and y. (We assume that $\Gamma^{(n,n)}$ is not identically vanishing, and without loss of generality set $m \geqslant n$.) Equivalently, ρ is such that the coherence index of order (m,n) is unimodular for all arguments:

$$|S^{(m,n)}(x;y)| = 1. \quad (2.15)$$

Two interesting questions arise: (i) what can be said about the other correlation functions $\Gamma^{(m',n')}$; (ii) what is the form of the most general ρ that leads to Eq. (2.15)? We examine these in turn, basing the analysis just on positivity of ρ and the Bose nature of V. The characteristic differences between the cases $m = n$ and $m > n$ will be pointed out at the appropriate places.

As pointed out elsewhere[16] that the coherence indices have a maximum modulus of unity may be seen as a consequence of Schwarz's inequality as applied to the m-fold and n-fold products of the field operator. The coherence index may be viewed as a generalized visibility index. In the analysis of Ref. 16 it was concluded that unimodularity of $A(n,n)$ implies it for all $S(m',n')$ with $\max(m',n') \geqslant n$. We shall see below that an even stronger conclusion can be derived from the unimodularity of $S(m,n)$ with $m \neq n$.

The vanishing of $\Delta^{(m,n)}(x;y)$, combined with Eq. (2.11) and the fact that ρ is nonnegative, leads to the following operator condition on ρ:

$$V(x_1)\cdots V(x_m)\rho = [\Gamma^{(n,m)}(y;x)/\Gamma^{(n,n)}(y;y)]\, V(y_1)\cdots V(y_n)\,\rho. \quad (2.16)$$

This is to be satisfied for all independent choices of the x's and y's. This is easily exploited to give the relations

$$\Gamma^{(m,n)}(x;y) = \Gamma^{(m,n)}(x;y')\Gamma^{(n,n)}(y';y)/\Gamma^{(n,n)}(y';y')$$

$$= \Gamma^{(m,m)}(x;x')\Gamma^{(m,n)}(x';y)/\Gamma^{(m,m)}(x';x'), \quad (2.17)$$

where again all arguments may be chosen independently. Stated in terms of the coherence indices, these relations are

$$S^{(m,n)}(x;y) = S^{(m,n)}(x;y')\, S^{(n,n)}(y';y)$$

$$= S^{(m,m)}(x;x')\, S^{(m,n)}(x';y). \quad (2.18)$$

In boths Eqs. (2.17) and (2.18) we have two independent relations only if $m > n$. These functional equations imply that each of the three coherence indices has a separable dependence on its two sets of arguments:

$$S^{(m,n)}(x;y) = S^{(m)}(x)/S^{(n)}(y),$$
$$S^{(m,m)}(x;x') = S^{(m)}(x)/S^{(m)}(x'),$$
$$S^{(n,n)}(y;y') = S^{(n)}(y)/S^{(n)}(y'), \quad (2.19)$$

We have here a symmetric function $S^{(m)}$ of m arguments and another $S^{(n)}$ of n arguments, the two coinciding for $m = n$. Then Eq. (2.15) shows that $S^{(m)}$ and $S^{(n)}$ may each be taken to be unimodular, which in turn leads to both $S^{(m,m)}$ and $S^{(n,n)}$ being unimodular along with $S^{(m,n)}$.

The result (2.19) can now be fed into the operator condition (2.16) on ρ, and this then takes the form

$$V(x_1)\cdots V(x_m)\rho = [\alpha^{(m)}(x)/\alpha^{(n)}(y)]\, V(y_1)\cdots V(y_n)\,\rho,$$
$$\alpha^{(m)}(x) \equiv [\Gamma^{(m,m)}(x,x)]^{1/2}/S^{(m)}(x),$$
$$\alpha^{(n)}(y) \equiv [\Gamma^{(n,n)}(y,y)]^{1/2}/S^{(n)}(y). \quad (2.20)$$

Once again we note that we have two distinct symmetric functions $\alpha^{(m)}$ and $\alpha^{(n)}$ when $m > n$ and just one when $m = n$, and the operator condition holds for all choices of x's and y's. We now exploit the fact that any two V's commute:

$$V(\xi)V(x_1)\cdots V(x_m)\rho = V(x_1)V(\xi)\cdots V(x_m)\rho,$$

i.e.,

$$V(\xi)[V(y_1)\cdots V(y_n)/\alpha^{(n)}(y)]\rho$$
$$= [\alpha^{(m)}(\xi,x_2,\ldots,x_m)/\alpha^{(m)}(x_1,x_2,\ldots,x_m)]$$
$$\times V(x_1)[V(y_1)\cdots V(y_n)/\alpha^{(n)}(y)]\,\rho. \quad (2.21)$$

Since the right-hand side must be independent of $x_n,\ldots,x_m$, it follows in the first instance that the dependence of $\alpha^{(m)}$ on its first argument must separate from its dependence on the remaining ones; but since $\alpha^{(m)}$ is a symmetric function, it must factorize all the way. Thus we get the result[17]

$$\alpha^{(m)}(x) = C_m\, u(x_1)\cdots u(x_m), \quad (2.22)$$

where we assume $u(x)$ is a normalized "mode" function, and C_m is a constant. By a parallel argument, for $m > n$, we have

$$\alpha^{(n)}(y) = C_n\, v(y_1)\cdots v(y_n), \quad (2.23)$$

but it is easy to show that v must coincide with u. Assuming this done, the operator condition (2.16), or equally well (2.20), on ρ has the form

$$V(x_1)\cdots V(x_m)\rho = [C_m\, u(x_1)\cdots u(x_m)/C_n u(y_1)\cdots u(y_n)]$$
$$\times V(y_1)\cdots V(y_n)\,\rho. \quad (2.24)$$

To fully exploit the fact that this must hold for all x and y, let us introduce the annihilation operator a for the mode u, and its adjoint as

$$a = \int (dx)\, u^*(x)V(x), \qquad a^\dagger = \int dx\, u(x)V(x)^\dagger. \quad (2.25)$$

Then Eq. (2.24) is equivalent to three conditions on ρ:

$$V(x_1)\cdots V(x_m)\rho = u(x_1)\cdots u(x_m)\, a^m\rho, \quad (2.26a)$$
$$V(y_1)\cdots V(y_n)\rho = u(y_1)\cdots u(y_n)\, a^n\rho, \quad (2.26b)$$
$$a^m\rho = (C_m/C_n)a^n\rho. \quad (2.26c)$$

For $m = n$, the first two conditions coincide while the third is vacuous. We shall see later that for $m > n$ (2.26c) is a very strong condition on ρ.

We have demonstrated thus far that Eqs. (2.26) are necessary consequences of Eq. (2.15); if the latter is valid, there

is some mode function $u(x)$ with associated operator a, and constants C_m, C_n, such that Eqs. (2.26) are obeyed by ρ. The converse is also true: From Eqs. (2.26) we can obtain Eq. (2.15). We easily get the results, given (2.26):

$$\Gamma^{(m,n)}(x;y) = u(x_1)^*\cdots u(y_n)\ \mathrm{Tr}(a^n\rho a^{\dagger m}),$$
$$\Gamma^{(m,m)}(x;x') = u(x_1)^*\cdots u(x'_m)\ \mathrm{Tr}(a'^m\rho a^{\dagger m}),$$
$$\Gamma^{(n,n)}(y;y') = u(y_1)^*\cdots u(y'_n)\ \mathrm{Tr}(a^n\rho a^{\dagger n}); \tag{2.27}$$

and from here, remembering Eq. (2.26c), the unimodularity of $S^{(m,n)}$ follows. Moreover, the factorizability of all correlation function $\Gamma^{(m',n')}$ for $m',n' \geqslant n$ is an easy consequence of Eqs. (2.26) [16]:

$$\Gamma^{(m',n')}(x;y) = g^{(m',n')}\ u(x_1)^*\cdots u(y_{n'}),$$
$$g^{(m',n')} = \mathrm{Tr}(a^{n'}\rho a^{\dagger m'}), \qquad m',n' \geqslant n. \tag{2.28}$$

In particular, if the optical discriminant $\Delta^{(1,1)}$ vanishes so that the two-point correlation function $\Gamma^{(1,1)}(x;y)$ factorizes, then *all* correlation functions, *including* the ones $\Gamma^{(m,0)}$ and $\Gamma^{(0,m)}$, will factorize in terms of just one mode function $u(x)$, and the statistical state may be said to be "mode pure." This is not necessarily a coherent state, but it will be so provided

$$\mathrm{Tr}(a^n\rho a^{\dagger m}) = z^n z^{*m} \tag{2.29}$$

for some complex z and all $m,n \geqslant 0$.

The second part of the analysis concerns the most general form ρ can have if Eq. (2.15) is to be satisfied. We know that we must find the most general solution to Eqs. (2.26); of these, Eq. (2.26a) is a consequence of Eq. (2.26b) if $m > n$, so we need solve just Eqs. (2.26b) and (2.26c), which we write again for definiteness:

$$V(y_1)\cdots V(y_n)\,\rho = u(y_1)\cdots u(y_n)\,a^n\rho, \tag{2.30a}$$
$$a^m\rho = (C_m/C_n)\,a^n\rho. \tag{2.30b}$$

The case $m = n$ [when (2.30b) is empty] is taken up first, the case $m > n$ later. To begin with, we recall the general form of ρ, given that it is Hermitian, nonnegative, and of unit trace, its eigenvalues ρ_j form a discrete set summing up to unity. If the corresponding eigenvectors are $|\psi_j\rangle$, we have:

$$\rho = \sum_j \rho_j |\psi_j\rangle\langle\psi_j|\,, \quad \rho_j > 0, \quad \sum_j \rho_j = 1. \tag{2.31}$$

There are, naturally, no terms here corresponding to zero being a possible eigenvalue of ρ; the $|\psi_j\rangle$ will be an orthonormal set which is in general not a complete one but can always be extended to a complete orthonormal set. Now it can be seen that even if the vectors $|\psi_j\rangle$ are not pairwise orthogonal, the expression in Eq. (2.31) yields an acceptable density matrix provided only that the numbers ρ_j remain positive, sum up to unity, and each $|\psi_j\rangle$ is of norm unity. In this more general situation, we do not interpret the ρ_j and $|\psi_j\rangle$ as eigenvalues and eigenvectors of ρ, but interpret ρ as being a convex combination of the pure state density operators $|\psi_j\rangle\langle\psi_j|$. It follows that we need find the most general solutions to Eqs. (2.30) assuming ρ to be as given in Eq. (2.31) but need not insist that the $|\psi_j\rangle$ be pairwise orthogonal. But the fact that the constants ρ_j in Eqs. (2.30) are positive is enough to show that the operator conditions developed above for ρ must in fact be satisfied by each $|\psi_j\rangle$; it is not necessary, in order to make this statement, that (2.30) be an expansion of ρ in terms of its eigenvalues and eigenvectors. Specifically, each of Eqs. (2.16), (2.20), (2.24), (2.26), and (2.30) must remain valid if $|\psi_j\rangle$ replaces ρ in them.

For the case $m = n$, we choose any normalized mode function u and ask for the most general vector $|\psi_j\rangle$ obeying

$$V(x_1)\cdots V(x_m)|\psi\rangle = u(x_1)\cdots u(x_m)\,a^m|\psi\rangle. \tag{2.32}$$

If $u(x)$ is chosen as the first member of a complete orthonormal set $(u,v_1,v_2,\cdots)$ and the field $V(x)$ is expanded as

$$V(x) = u(x)\,a + \sum_\alpha v_\alpha(x)\,b_\alpha, \tag{2.33}$$

then the content of Eq. (2.32) can be stated in words thus: $|\psi\rangle$ must be annihilated by every product of m annihilation operators provided at least one of them is one of the b_α. $|\psi\rangle$ must be of the form

$$|\psi\rangle = \{F(a^\dagger) + P^{(m-1)}[V^\dagger(x)]\}|0\rangle, \tag{2.34}$$

where F is an arbitrary function of its argument and $P^{(m-1)}$ is a polynomial functional of the creation field $V^\dagger$ of degree not more than $(m-1)$. We may take any number of such (normalized) vectors $|\psi_j\rangle$, choose any positive constants ρ_j summing to unity, and put them into Eq. (2.31) to get the most general statistical state ρ for which the (m,m) optical discriminant vanishes identically. The quantities to be chosen freely are the mode function u, the functions F_j, the functionals $P_j^{(m-1)}$, and the ρ_j.

For the case $m = n + N$, $N \geqslant 1$, we choose again a mode $u(x)$ and a complex number z,

$$C_m/C_n = z^N. \tag{2.35}$$

We then look for the general solution $|\psi\rangle$ to

$$V(x_1)\cdots V(x_n)|\psi\rangle = u(x_1)\cdots u(x_n)\,a^n|\psi\rangle, \tag{2.36a}$$
$$a^N(a^n|\psi\rangle) = z^N(a^n|\psi\rangle). \tag{2.36b}$$

As in Eq. (2.34), Eq. (2.36a) is solved by

$$|\psi\rangle = \{F(a^\dagger) + P^{(n-1)}[V(x)^\dagger]\}|0\rangle, \tag{2.37}$$

but now the function F is severely restricted by the remaining condition (2.36b). In fact, apart from the freedom to add a polynomial of degree $(n-1)$ in $a^\dagger$ (which could be absorbed in $P^{(n-1)}$), we find

$$F(a^\dagger) = \sum_{r=0}^{N-1} \beta_r \exp(e^{2\pi i r/N} z a^\dagger), \tag{2.38}$$

where β_r are arbitrary constants. Thus the acceptable states $|\psi\rangle$ are largely determined in terms of coherent states as defined in the last section, with respect to the mode u:

$$|\psi\rangle = \sum_{r=0}^{N-1} \beta'_r |z e^{2\pi i r/N}\rangle_u + P^{(n-1)}[V(x)]^\dagger|0\rangle. \tag{2.39}$$

Once again taking several such (normalized) vectors $|\psi_j\rangle$ and choosing ρ_j, having first picked a $u(x)$ and a z, we get via Eq. (2.31) the most general statistical state with unimodular coherence index of order (m,n).

3. ENUMERATION OF EXCITED MODES

In the previous section we discovered the general form

of a statistical state for which it was assumed that some definite optical discriminant $\Delta^{(m,n)}$ vanished identically. The parts of ρ corresponding to "large" (greater than or equal to n) total photon numbers was to a considerable extent determined, and was seen to depend on just one single mode function $u(x)$. Now we outline a method of directly analyzing these components of ρ without assuming anything about the optical discriminants. As is to be expected, the large photon number description of ρ will involve more than just one mode function in general.

The total photon number operator, N, is given in terms of the potential $\mathbf{A}$ and field $\mathbf{E}$ as

$$N = -2i\int d^3r \mathbf{A}^{(+)}(\mathbf{r})^\dagger \cdot \mathbf{E}^{(+)}(\mathbf{r}). \tag{3.1}$$

Since $\mathbf{E}^{(+)}$ is $i\omega$ times $\mathbf{A}^{(+)}$, where ω is the frequency operator, we can rewrite N in the form

$$N = \int dx\, dy\, V(x)^\dagger H(x,y)\, V(y), \tag{3.2}$$

where we have denoted the Fourier transform of 2ω by $H(x,y)$, and the integrations include polarization sums. Let us now start with the (n,n) order correlation function $\Gamma^{(n,n)}(x;y)$,

$$\Gamma^{(n,n)}(x;y) = \mathrm{Tr}[V(y_1)\cdots V(y_n)\rho V(x_1)^\dagger \cdots V(x_n)^\dagger]. \tag{3.3}$$

Since the trace is unchanged by cyclic permutation of its arguments, and since any two V's ($V^\dagger$'s) commute, we can use Eq. (3.2) to "contract" one V and one $V^\dagger$ and produce an N at the proper place within the trace:

$$\begin{aligned}&\int H(x_1,y_1)\,\Gamma^{(n,n)}(x;y)\,dx_1 dy_1\\ &= \mathrm{Tr}[\rho V(x_2)^\dagger \cdots V(x_n)^\dagger N V(y_2)\cdots V(y_n)]\\ &= \mathrm{Tr}[V(y_2)\cdots V(y_n)(N-n+1)\rho V(x_2)^\dagger \cdots V(x_n)^\dagger].\end{aligned} \tag{3.4}$$

Carrying out this operation $(n-2)$ times more, and using cyclic invariance, we find:

$$\begin{aligned}&\int H(x_1,y_1)\cdots H(x_{n-1},y_{n-1})\,\Gamma^{(n,n)}(x;y)\,dz_1 dy_1 \cdots dx_{n-1}\,dy_{n-1}\\ &\quad = \mathrm{Tr}[V(y_n)(N-1)(N-2)\cdots(N-n+1)\rho V(x_n)^\dagger]\\ &\quad = \mathrm{Tr}[V(y_n)\sigma V(x_n)^\dagger],\\ &\sigma = \sqrt{(N-1)(N-2)\cdots(N-n+1)}\\ &\qquad \times \rho \sqrt{(N-1)(N-2)\cdots(N-n+1)}.\end{aligned} \tag{3.5}$$

The operator σ, like ρ, is Hermitian and nonnegative; all states in ρ with at most $(n-1)$ photons drop out in σ, while states in ρ with n or more photons survive with positive numerical factors. Thus by this process we obtain for $\Gamma^{(n,n)}$ an auxiliary two-point function,

$$\Phi(x,y) = \mathrm{Tr}[V(y)\,\sigma V(x)^\dagger], \tag{3.6}$$

which retains the positivity and hermiticity property but involves only n or more photon contributions to ρ. Let us then make an eigenvector decomposition of Φ:

$$\Phi(x,y) = \sum_\nu \lambda_\nu u_\nu(x)^* u_\nu(y), \quad \lambda_\nu > 0. \tag{3.7}$$

Only the strictly positive eigenvalues λ_ν of Φ and the corresponding eigenfunctions appear here; the functions u_ν form an orthonormal set which if not complete can be extended to a complete orthonormal set. Let u_ν, $\nu = 1,2,\cdots\infty$, be the latter complete set, and let the subset of ν values covered by the sum in Eq. (3.7) be denoted by κ. Make an expansion of V in the basis u_ν,

$$V(x) = \sum_{\nu=1}^{\infty} u_\nu(x) a_\nu. \tag{3.8}$$

Then Eq. (3.7) may be written

$$\sum_{\mu,\nu=1}^{\infty} u_\mu(x)^* u_\nu(y)\mathrm{Tr}(a_\nu \sigma a_\mu^\dagger) = \sum_{\nu\in\kappa} \lambda_\nu u_\nu(x)^* u_\nu(y), \tag{3.9}$$

that is,

$$\mathrm{Tr}(a_\nu \sigma a_\mu^\dagger) = \begin{cases} 0 & \text{if } \mu \neq \nu \text{ or } \nu \notin \kappa, \\ \lambda_\nu & \text{if } \mu = \nu \in \kappa. \end{cases} \tag{3.10}$$

Because of the basic nonnegativity of σ we may conclude from here that

$$a_\mu \sigma = 0 \quad \text{if } \mu \notin \kappa. \tag{3.11}$$

Moreover, since the most general structure for σ is of the form

$$\sigma = \sum_j \sigma_j |\phi_j\rangle\langle\phi_j|, \quad \sigma_j > 0, \tag{3.12}$$

with the (normalized but not necessarily mutually orthogonal) states $|\phi_j\rangle$ having at least n photons, the condition (3.11) passes over to each $|\phi_j\rangle$,

$$a_\mu |\phi_j\rangle = 0 \quad \text{if } \mu \notin \kappa. \tag{3.13}$$

For any given n, this analysis tells us that only the eigenmodes corresponding to nonzero eigenvalues of the auxiliary two-point function Φ can be excited in the contributions to ρ having at least n photons.[16] If u_ν, $\nu\in\kappa$, is less than a complete set, this is nontrivial information concerning ρ. The general solution to Eq. (3.13) is

$$|\phi\rangle = F^{(>n)}(a_\nu^\dagger, \nu\in\kappa)|0\rangle, \tag{3.14}$$

where $F^{(>n)}$ is a possibly infinite polynomial in the indicated operators with the lowest degree terms being at least of order n. Taking several ρ solutions of this form we can get the general structure of σ and then of ρ remembering Eq. (3.5):

$$\begin{aligned}&\rho = \sum_j \rho_j |\psi_j\rangle\langle\psi_j|, \quad \rho_j > 0, \quad \sum_j \rho_j = 1,,\\ &|\psi_j\rangle = \{F_j(a_\nu^\dagger, \nu\in\kappa) + P_j^{(n-1)}[V(x)^\dagger]\}|0\rangle.\end{aligned} \tag{3.15}$$

We assume the $|\psi_j\rangle$ are unit vectors. $P_j^{(n-1)}$ are polynomial functionals of degree no more than $(n-1)$ in $V(x)^\dagger$; and the F_j are arbitrary functions of the indicated operators. We note that this result is the generalization of the results of Section 2, with the single mode u being enlarged to the set u_ν, $\nu\in\kappa$. The correlation function $\Gamma^{(n,n)}$ in this case involves the distinguished modes u_ν, $\nu\in\kappa$:

$$\begin{aligned}&\Gamma^{(n,n)}(x;y) = \sum_{\mu_i,\nu\in\kappa} u_{\mu_1}(x_1)^* \cdots u_{\nu_n}(y_n)\, g^{(n,n)}(\mu;\nu),\\ &g^{(n,n)}(\mu;\nu) = \mathrm{Tr}(a_{\nu_1}\cdots a_{\nu_n}\rho a_{\mu_1}^\dagger \cdots a_{\mu_n}^\dagger).\end{aligned} \tag{3.16}$$

But it is immediately obvious that a similar restriction to the

set κ occurs in all $\Gamma^{(m',n')}$ for $m',n' \geqslant n$. It is as if we may ignore all modes except those present in Eq. (3.7) for computing the "sufficiently high" order correlation functions.

Additional information on the correlation functions beyond (n,n) allows repetition of this analysis. We now supply the auxiliary $\Phi(x,y)$ with an index n. Then the $\geqslant n$ photon parts of the statistical state ρ can involve only the eigenmodes of $\Phi_n(x,y)$ (corresponding to nonzero eigenvalues!). If we have a sequence of auxiliary functions $\Phi_{n_1}, \Phi_{n_2}, \cdots$ with $n_1 < n_2 < \cdots$, then we have a corresponding sequence of sets of eigenmodes $[u_\nu(x), \nu \in k_1]$, $[u_\nu(x), \nu \in \kappa_2], \cdots$ with the inclusion relations $\kappa_1 \supseteq \kappa_2 \supseteq \cdots$. The higher the order of the correlation function considered, the fewer the contributing modes!

4. SECOND HILBERT SPACE AND COHERENT STATES

Given a Hilbert space $\mathscr{H}$, we can induce a vector space structure with an inner product among the linear operators $A,B,\cdots$ on $\mathscr{H}$ by defining

$$(A,B) = \mathrm{Tr}(A^\dagger B). \tag{4.1}$$

So the operators of Hilbert–Schmidt class with finite values of $\mathrm{Tr}(A^\dagger A)$ constitute a second Hilbert space $\mathscr{K}$ based on $\mathscr{H}$. In particular we may choose $\mathscr{H}$ as the representation space of the (unbounded) operators $a,a^\dagger$ or q,p associated with a quantum system with one degree of freedom.

Let A be a general operator on $\mathscr{H}$. We define a set of four "superoperators" which are operators on $\mathscr{K}$ and whose effects on A considered as a vector in $\mathscr{K}$ are given by:

$$\mathscr{A}_1 A = (aA - Aa^\dagger)/\sqrt{2}, \quad \mathscr{A}_1^\dagger A = (a^\dagger A - Aa)/\sqrt{2},$$
$$\mathscr{A}_2 A = -i(aA + Aa^\dagger)/\sqrt{2}, \quad \mathscr{A}_2^\dagger = i(a^\dagger A + Aa)/\sqrt{2}. \tag{4.2}$$

These hermiticity relations are in accordance with the definition (4.1) of the scalar product. By virtue of their definitions it follows that

$$[\mathscr{A}_1, \mathscr{A}_1^\dagger] = [\mathscr{A}_2, \mathscr{A}_2^\dagger] = 1,$$
$$[\mathscr{A}_1, \mathscr{A}_2] = [\mathscr{A}_1, \mathscr{A}_2^\dagger] = 0. \tag{4.3}$$

Hence these operators correspond to a quantum system with *two* degrees of freedom.[8]

The simultaneous coherent states corresponding to it, and $\mathscr{A}_2$ are supplied by the outer products of the coherent states in $\mathscr{H}$:

$$\mathscr{A}_1 |z_1\rangle\langle z_2| = \frac{z_1 - z_2^*}{\sqrt{2}} |z_1\rangle\langle z_2|,$$
$$\mathscr{A}_2 |z_1\rangle\langle z_2| = -i \frac{z_1 + z_2^*}{\sqrt{2}} |z_1\rangle\langle z_2|. \tag{4.4}$$

In particular,

$$\mathscr{A}_1 |z\rangle\langle z| = i\sqrt{2}\, \mathrm{Im}\, z |z\rangle\langle z|,$$
$$\mathscr{A}_2 |z\rangle\langle z| = -i\sqrt{2}\, \mathrm{Re}\, z |z\rangle\langle z|. \tag{4.5}$$

Hence from among all the coherent states $|z_1\rangle\langle z_2|$ in $\mathscr{K}$, those with $z_1 = z_2 = z$ correspond to pure imaginary eigenvalues for both $\mathscr{A}_1$ and $\mathscr{A}_2$. By a straightforward generalization of Eq. (1.20) to two degrees of freedom, we may now expand a density operator ρ on $\mathscr{H}$ considered as a vector in $\mathscr{K}$ in terms of the imaginary eigenvalue coherent states,

$$\rho = (1/\pi) \int \phi(z) |z\rangle\langle z| d^2 z. \tag{4.6}$$

$\phi(z)$ is the (distribution) weight function analogous to $v(r)$ in Eq. (1.20). To determine $\phi(z)$ by an equation similar to (1.22) [or (1.23)], we must get the description of $\mathscr{K}$ by Schrödinger wavefunctions like $\psi(q')$. We define

$$Q_1 = (\mathscr{A}_1 + \mathscr{A}_1^\dagger)/\sqrt{2}, \quad P_1 = -i(\mathscr{A}_1 - \mathscr{A}_2^\dagger)/\sqrt{2},$$
$$Q_2 = (\mathscr{A}_2 + \mathscr{A}_2^\dagger)/\sqrt{2}, \quad P_2 = -i(\mathscr{A}_2 - \mathscr{A}_2^\dagger)/\sqrt{2}. \tag{4.7}$$

On a general $A \in \mathscr{K}$ these act as

$$Q_1 A = \frac{1}{\sqrt{2}}[q,A], \quad P_1 A = \frac{1}{\sqrt{2}}\{p,A\},$$
$$Q_2 A = \frac{1}{\sqrt{2}}[p,A], \quad P_2 A = \frac{-1}{\sqrt{2}}\{q,A\}, \tag{4.8}$$

and the only nonvanishing commutators among them are

$$[Q_1,P_1] = [Q_2,P_2] = i. \tag{4.9}$$

The (generalized) "eigenstates" of Q_1 and Q_2 turn out to be the Weyl family of unitary operators $W(\alpha)$ defined in Eq. (1.14). If α is written as $r + is$, we find:

$$Q_1 W(r+is) = rW(r+is), \quad Q_2 W(r+is) = sW(r+is),$$
$$P_1 W(r+is) = \frac{i\partial}{\partial r} D(r+is),$$
$$P_2 W(r+is) = \frac{i\partial}{\partial s} W(r+is),$$
$$(W(r'+is'), W(r+is)) = \pi\delta(r'-r)\delta(s'-s). \tag{4.10}$$

These operators have general coherent state matrix elements which may be viewed as a generalization of Eq. (1.10),

$$\langle z'|W(\alpha)|z\rangle = [|z'\rangle\langle z|, W(\alpha)] = \exp(-\tfrac{1}{2}|\alpha|^2 - \tfrac{1}{2}|z|^2$$
$$-\tfrac{1}{2}|z'|^2 + \alpha z'^* - \alpha^* z + zz'^*). \tag{4.11}$$

With the operators $W(\alpha)$ forming an "ideal" basis for $\mathscr{K}$, any $A \in \mathscr{K}$ and in particular a density operator ρ has an expansion

$$\rho = \int t(\alpha)\, W(\alpha) d^2\alpha \tag{4.12}$$

with the "Weyl weight" playing the role of $\psi(q')$ in Sec. 1. The distribution $\phi(z)$ obeys the integral equations[18]

$$\int \phi(z) \exp(-|z-z'|^2) d^2 z$$
$$= \pi \int t(\alpha) \exp(-\tfrac{1}{2}|\alpha|^2 + \alpha z'^* - \alpha^* z')\, d^2\alpha, \tag{4.13}$$

which generalizes Eq. (1.21) and arises by taking the diagonal coherent state matrix element in Eqs. (4.6) and (4.12) and using Eqs. (1.11) and (1.15). The solution

$$\phi(z) = \int t(\alpha) \exp(\tfrac{1}{2}|\alpha|^2) \exp(\alpha z^* - \alpha^* z)\, d^2\alpha \tag{4.14}$$

is the analog of Eq. (1.22). We can also make use of the properties of coherent states to derive the known result[18]

$$\alpha(z) = \frac{\exp(|z|^2)}{\pi} \int \langle -z'|\rho|z'\rangle \exp(|z'|^2 + zz'^* - z^*z')\, d^2z', \tag{4.15}$$

which is in the pattern of Eq. (1.23).

5. DISCRETE DIAGONAL REPRESENTATIONS

The preceding section has shown that the problem of the diagonal coherent state representation (4.6) for a density operator is in essence the same as the problem of expanding a vector $|\psi\rangle$ in terms of a subset of "pure imaginary" coherent states as attempted in Eq. (1.20). But for vectors we know that even more economical subsets exist through which arbitrarily accurate expansions can be made. We can thus combine the properties of these characteristic sets recounted in Sec. 1 with the approach of Sec. 4 to get new diagonal approximations for density operators.

We begin with the one-degree-of-freedom case. Since by Eq. (4.5) "diagonal" outer products of coherent states correspond to pure imaginary eigenvalues for $\mathscr{A}_1$ and $\mathscr{A}_2$, we proceed as follows. We pick two discrete characteristic sets $\{i\sqrt{2}y_n\}$ and $\{-i\sqrt{2}\,x_m\}$, both consisting of points on the imaginary axis, so that x_m and y_n are real. Then the double sequence of points $\{i\sqrt{2}\,y_n, -i\sqrt{2}\,x_m\}$ is a characteristic set[11] in the product of the complex plane by itself. Consequently the simultaneous coherent states of $\mathscr{A}_1$ and $\mathscr{A}_2$ with respective eigenvalues $i\sqrt{2}\,y_n$, $-i\sqrt{2}\,x_m$ give a set of elements in $\mathscr{H}$ through linear combinations of which one can approximate any member of $\mathscr{H}$ arbitrarily closely. But by Eq. (4.5) these elements of $\mathscr{H}$ are just the elements

$$|z_{mn}\rangle\langle z_{mn}|, \quad z_{mn} = x_m + iy_n. \tag{5.1}$$

Consequently we can approximate any density operator arbitrarily closely (in Hilbert–Schmidt norm!) by a discrete sum

$$\sum_{m,n} \phi_{mn} |z_{mn}\rangle\langle z_{mn}|. \tag{5.2}$$

The number of terms needed and the coefficients ϕ_{mn} to be used both depend on the desired accuracy of approximation.[19] In the case with one degree of freedom and one operator a, the operator (5.2) describes an ensemble over the preassigned set of values $\{z_{mn}\}$ for a, with the real coefficient ϕ_{mn} being the quasiprobability associated with the realization z_{mn} for a. Provided the trace property is maintained in these approximations to ρ, the ϕ_{mn} must add up to unity but in general the nonnegativity of ϕ_{mn} cannot be guaranteed; thus Eq. (5.2) is *not* in general, a convex combination of the pure state density operators $|z_{mn}\rangle\langle z_{mn}|$.

Since a set of points with a finite limit point does form a characteristic set,[11] we can have approximations of the above type with all but a finite number of points z_{mn} inside one phase cell! In certain respects this would mimic an amplitude-stabilized (mode-pure) classical light beam, but *every* ρ could be reached this way. The quantum aspects are hidden in the rapid variations of ϕ_{mn} as z_{mn} varies inside the chosen phase cell and also as the degree of approximation is improved.

At the other extreme, we can try to choose the set $\{z_{mn}\}$ as uniformly spread out as possible over the complex plane: the basic limitation now comes from the condition (1.25) for a characteristic set if finite limit points are to be avoided. This condition is to be applied both to $\{x_m\}$ and $\{y_n\}$. Let us then demand that

$$\sum_{n=1}^{\infty} |x_n|^{-2-\epsilon_1} = \infty, \quad \sum_{n=1}^{\infty} |y_n|^{-2-\epsilon_2} = \infty, \tag{5.3}$$

for some positive ϵ_1 and ϵ_2. It is clear that x_n and y_n cannot depend linearly on n, since these series would then converge for any $\epsilon_1, \epsilon_2 \geqslant 0$. Among fractional power dependences, possible candidates are, for example,

$$y_n, x_n \sim n^{\alpha}, \quad \alpha < \tfrac{1}{2}. \tag{5.4}$$

This means that the number of points of the set $\{z_{mn}\}$ contained in a square of side L is approximately $L^{2/\alpha}$ which is, for large L, larger than L^2. In this sense, one needs much more than one point for phase cell to have a set of pure state density operators $|z_{mn}\rangle\langle z_{mn}|$ built from coherent states, with combinations of which any ρ can be approximated.

Formally, representations of the type (5.2) can be extended from one degree of freedom to the entire field. The operator $V(x)$ is expanded in some complete orthonormal set $u_\alpha(x)$,

$$V(x) = \sum_{\alpha} u_\alpha(x) a_\alpha. \tag{5.5}$$

For each mode α, we choose some mesh of points $z^{(\alpha)}_{mn}$ in the complex plane in the manner described for one degree of freedom. Then a density operator ρ for the whole field can be approximated through expressions of the form

$$\sum_{|m,n|} \phi(\{m_\alpha, n_\alpha\}) |\{z^{(\alpha)}_{m_\alpha n_\alpha}\}\rangle\langle\{z^{(\alpha)}_{m_\alpha n_\alpha}\}| \tag{5.6}$$

with

$$a_\beta |z^{(\alpha)}_{m_\alpha n_\alpha}\rangle = \{z^{(\beta)}_{m_\beta n_\beta}\} |\{z^{(\alpha)}_{m_\alpha n_\alpha}\}\rangle. \tag{5.7}$$

We interpret (5.6) as an ensemble over the preassigned set of realizations of $V(x)$ given by

$$V(x;\{m,n\}) = \sum_{\alpha} \{z^{(\alpha)}_{m_\alpha n_\alpha}\}\, u_\alpha(x). \tag{5.8}$$

This is of course a c-number function of x. As in the one degree of freedom case, the quasiprobability $\phi(\{m,n\})$ that $V(x;\{m,n\})$ is realized is real but may be negative. Even the term quasiprobability is used only figuratively since two coherent states are never orthogonal!

6. CONCLUDING REMARKS

In this paper we have studied the properties of statistical states describing general light beams from two points of view. On the one hand we have shown how from the analysis of coherence functions one can systematically look for simplicity in the large photon number sectors of a given state.

On the other hand, we have given new understanding of the diagonal coherent state representation, and in the process discovered new representations embodying very subtle features of quantum theory.

The structure of a state for which the coherence index of order (m,n) is unimodular is essentially determined, for photon numbers greater than or equal to n, by a single mode function and a sequence of constants. The striking differences between the cases $m = n$ and $m > n$ are worth pointing out. In the former, for instance, even if one knew that m photon states were present in ρ, nothing definite could be said about $m + 1$ and higher photon states; they may or may not be present in ρ. If $m > n$, on the other hand, the presence of m photon states in ρ guarantees the presence of states with arbitrarily large photon numbers, all in the mode $u(x)$ of course. This is because the relevant parts of ρ are determined in terms of coherent states which are superpositions of states will all possible numbers of photons.

Turning to the theory of discrete diagonal coherent state approximations to ρ, we would like to make two comments. The first is to clarify the situation concerning the use of a characteristic set of coherent states distributed "as uniformly as possible" over the phase plane. At the level of making approximations to *vectors* through linear combinations of coherent states, the example due to von Neumann shows that by taking one coherent state per unit phase cell we get a characteristic set. In other words, the set of points

$$z_{l,m} = \sqrt{\pi}\,(l + im), \quad l,m = 0, \pm 1, \pm 2, \cdots$$

is a characteristic set in the complex plane, and any $|\psi\rangle$ can be approximated through combinations of coherent states $|z_{l,m}\rangle$. But the problem of the discrete diagonal approximations to operators through forming linear combinations of

$$|z_{mn}\rangle\langle z_{mn}|$$

is *different*. Here, the real and imaginary parts, x_m and y_n, of Z_{mn} must be such that the set (x_m, y_n), essentially, must be characteristic in the product of the complex plane by itself. Equivalently, $\{x_m\}$ and $\{y_n\}$ must each be a real characteristic set, and this precludes $\{z_{mn}\}$ being distributed uniformly over the complex plane!

The second comment concerns the fact that one must be content with arbitrarily close approximations to ρ but may "never quite get there." To us this seems to point to the very thin line dividing "definition" and "existence." One can well imagine being completely innocent of the theory of distributions as yet conceiving of the possibility of the diagonal representation in its conventional form, Eq. (4.6). Faced with equations such as (4.13), (4.14), and (4.15) to find the weight function ϕ, one would be forced to say that in general no ϕ exists obeying these equations, but that with "good" functions ϕ_α in the formula (4.6) one can produce density operators ρ_α that are "arbitrarily close" to any given ρ! Indeed, distributions may usefully be thought of in this way. It seems that the situation is qualitatively quite similar with the discrete representations based on characteristic sets.

An interesting practical problem is to find ways of computing the coefficients ϕ_{mn} in the discrete diagonal approximation (5.2) to a given ρ for a desired accuracy. One may try to get ϕ_{mn} from a knowledge of the Weyl weight $t(\alpha)$ of ρ, but the technique of Fourier transformation seems not useful in this context. This seems to be a genuinely difficult problem; we have only succeeded in establishing the *existence* of these representations.

ACKNOWLEDGMENT

This work was supported in part by the Department of Energy in contract with the University of Texas at Austin under contract No. EY-76-S-05-3992.

[1]Max Planck, Verh. D. Phys. Ges. **2**, 237 (1900).
[2]S.N. Bose, Z. Physik **26**, 178 (1924).
[3]G.K. Batchelor, *The Theory of Homogeneous Turbulence* (Cambridge U.P., London and New York, 1963).
[4]M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1975).
[5]R.J. Glauber, Phys. Rev. Lett. **10**, 84 (1963); Phys. Rev. **130**, 2529 (1963).
[6]E.C.G. Sudarshan, Phys. Rev. Lett. **10**, 277 (1963).
[7]J.R. Klauder, J. Math. Phys. **4**, 1055, 1058 (1963); **5**, 177 (1963).
[8]N. Mukunda and E.C.G. Sudarshan, "New Light on the Optical Equivalence Theorem and a New Type of Discrete Diagonal Coherent State Representation," Pramāṇa **10**, 227–38 (1978).
[9]E. Schrödinger, Naturwissenschaften **14**, 644 (1927).
[10]I.E. Segal, Illinois J. Math. **6**, 500 (1962); S.S. Schweber, J. Math. Phys. **3**, 831 (1962).
[11]V. Bargman, Commun. Pure Appl. Math. **14**, 187 (1961).
[12]J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton U.P., Princeton, New Jersey, 1955); J.R. Klauder and E.C.G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968); A.M. Perelomov, Theor. Math. Phys. (Russian) **6**, 213 (1971). Perelomov shows that the example given by von Neumann $z = \sqrt{\pi}\,(l + im)$, $l,m = 0, \pm 1, \pm 2, \cdots$ is overcomplete but only by *one* state and is therefore a characteristic set in the sense of (1.24) but *not* in the sense that it remains a characteristic set if any finite number of points are removed from it. We thank John R. Klauder for information on Perelomov's work.
[13]P.A.M. Dirac, Proc. R. Soc. London **114**, 710 (1927).
[14]L. Mandel, E.C.G. Sudarshan, and E. Wolf, Proc. Phys. Soc. London **84**, 435 (1964).
[15]D. Gabor, J. Inst. Elec. Eng. **93**, 429 (1946); E. Wolf, Proc. R. Soc. A **230**, 246 (1955).
[16]E.C.G. Sudarshan, J. Math. Phys. Sci. **3**, 121 (1969); see also J.R. Klauder and E.C.G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968), pp. 166–7.
[17]U.M. Titulaer and R.J. Glauber, Phys. Rev. **140**, B676 (1965).
[18]C.L. Mehta and E.C.G. Sudarshan, Phys. Rev. **138**, B274 (1965).
[19]R.E. Paley and N. Wiener, *Fourier Transforms in the Complex Domain* (Am. Math. Soc., Providence, 1934).
[20]E.P. Wigner, Phys. Rev. **40**, 749 (1932); J.E. Moyal, Proc. Cambridge Philos. Soc. **45**, 99 (1948); E.C.G. Sudarshan, Phys. Rev. Lett. **10**, 277 (1963).