Coherent-State Representations for the Photon Density Operator

KEVIN E. CAHILL*
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts
(Received 18 August 1964; revised manuscript received 12 February 1965)

The "diagonal" P representation of the photon density operator in terms of the coherent states of the radiation field is studied. It is shown that the class of weight functions $P(\alpha)$ required to represent all density operators by means of the P representation is much larger than the class of all probability distributions and that therefore the analogy linking the use of the P representation with certain classical stochastic techniques is useful only for a limited class of density operators. The class of density operators for which the P representation is appropriate is shown to be limited in the sense that there are density operators that require weight functions so singular as to lie outside the class of all tempered distributions. Conditions on the weight function $P(\alpha)$ under which the expectation values of the normal ordered operators $(a^\dagger)^n a^m$ assume a particularly simple integral form are presented. It is shown that if $P(\alpha)$ is a tempered distribution, then these expectation values are given by a simple limiting process. A limited correspondence with classical optics is suggested by a formula for these expectation values. Conditions that a density operator must satisfy if its weight function is to be non-negative are presented and are shown to exclude all pure states except the coherent states themselves. The problem of representing an arbitrary operator in terms of the coherent states is studied. New measures of the completeness of the coherent states are established.

I. INTRODUCTION

BECAUSE of the zero rest mass of the photon, the wave properties of light are more obvious than its corpuscular aspects and were, in fact, exhibited more than a century earlier. The wave theory of Young and Maxwell has provided a sound conceptual framework for optics, within which all the classic experiments have been successfully understood. Recently, however, the introduction of the laser and of new techniques for counting photons has vastly increased the variety and precision of optical experiments. It has become important in these new areas of optics to make use of the greater insight and detail afforded by the quantum theory.

In this context, Glauber[1–5] has begun the development of a fully quantum dynamical theory of the statistical properties of light beams. A fundamental method in his formulation is the use of the "coherent" states as a basis for describing the radiation field. These states $|\alpha\rangle$ are eigenstates of the photon annihilation operator a, i.e., we have $a|\alpha\rangle=\alpha|\alpha\rangle$, and their eigenvalues α cover the entire complex plane. In the classical limit, $|\alpha|\gg1$, these state vectors correspond to states in which the field vectors have a precisely defined set of Fourier coefficients. The coherent states form a complete set, as is evident from the fact that the unit operator may be expressed as

$$1=\frac{1}{\pi}\int|\alpha\rangle\langle\alpha|d^2\alpha\,, \qquad (1.1)$$

where $d^2\alpha\equiv d(\mathrm{Re}\alpha)d(\mathrm{Im}\alpha)$ is a real element of area.

Interest has been focused lately upon the manner in which a density operator describing the statistical properties of an arbitrary optical system may be represented in terms of the coherent states. It is always possible to represent the density operator ρ as an integral over all outer products of coherent states, since we may use Eq. (1.1) to write

$$\rho=\frac{1}{\pi^2}\int|\alpha\rangle\langle\alpha|\rho|\beta\rangle\langle\beta|d^2\alpha d^2\beta\,. \qquad (1.2)$$

This representation has been studied by Glauber, who has shown that when the function $\langle\alpha|\rho|\beta\rangle$ possesses a certain integral representation, the expression (1.2) may be reduced to a simpler form which he called the P representation. In this form, the density operator is expressed as an integral over all outer products of identical coherent states:

$$\rho=\int|\alpha\rangle P(\alpha)\langle\alpha|d^2\alpha\,. \qquad (1.3)$$

For a wide variety of fields, the use of the P representation can simplify the computation of an important class of expectation values. The expectation values of the normally ordered products $(a^\dagger)^n a^m$ of the fundamental field operators are diagonal sums of the form $\mathrm{Tr}[\rho(a^\dagger)^n a^m]$. When the P representation is used for ρ and when the orders of summation and integration may be interchanged, we obtain from Eq. (1.3) the trace relations

$$\mathrm{Tr}[\rho(a^\dagger)^n a^m]=\int(\alpha^*)^n\alpha^m P(\alpha)d^2\alpha\,. \qquad (1.4)$$

These relations express the quantum-mechanical expectation values in a form resembling classical ensemble

---

* National Science Foundation Predoctoral Fellow, 1963–65.
[1] R. J. Glauber, Phys. Rev. Letters **10**, 84 (1963).
[2] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).
[3] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).
[4] R. J. Glauber, *Quantum Electronics, Proceedings of the Third International Congress, Paris, 1963* (Columbia University Press, New York, 1964), Vol. I, p. 111.
[5] R. J. Glauber, *Quantum Optics and Electronics, Lecture Notes of the 1964 Session of the Summer School for Theoretical Physics, Les Houches, France* (Gordon and Breach Science Publishers, Inc., New York, 1965).

averages and thereby provide an instructive and useful analogy with classical optics. In this analogy, the weight function $P(\alpha)$ (or $P(\{\alpha\})$ for a multimode system) corresponds to a classical probability distribution for the Fourier coefficient(s) of the classical fields.[6] The weight function $P(\alpha)$ is real and normalized to unity, as can be seen from Eqs. (1.3) and (1.4), by using the facts that $\rho^\dagger=\rho$ and $\mathrm{Tr}(\rho)=1$. The density operators belonging to some physical states may be represented by using well-behaved, positive-definite weight functions $P(\alpha)$, which go smoothly over into the corresponding classical probability distributions as these states approach the classical limit.

This correspondence is not of universal validity, however. The weight functions $P(\alpha)$ may be extremely singular and may assume negative values; in some cases they share none of the mathematical properties of probability distributions (except reality and normalization). In Sec. VI, we shall present the very restrictive conditions that a density operator must satisfy if its weight function is to be non-negative. We shall show that these conditions exclude all pure states, except the coherent states themselves. The density operators that may be represented by weight functions having only delta function singularities are also rather special, as will be proven in Sec. IV.[7]

The class of density operators for which the P representation is appropriate has not yet been fully characterized. We shall not attempt this characterization in the present paper but shall only remark that not all density operators can be said to possess this representation in any useful form. For there are density operators that require weight functions so intractably singular as to lie outside the class of all tempered distributions.[8,9] In fact, as will be shown in Sec. IV, a density operator must satisfy a set of rather restrictive conditions if its weight function is to be a tempered distribution. In a recent paper, Sudarshan[10] has proposed an explicit construction for any density operator of a weight functional that formally satisfies Eqs. (1.3) and (1.4). This weight functional, however, is in general so singular that it not only fails to be a tempered distribution, it lies outside the much wider class of all distributions, as we shall demonstrate in Sec. II.

The class of tempered distributions is the widest class of functionals that share with probability distributions a certain continuity property. As we shall show in Sec. II, the expectation values $\langle F_n\rangle$ of a sequence of functions $F_n(\alpha)$ calculated from a probability distribution $\mathcal{P}(\alpha)$ according to the rule

$$\langle F_n\rangle=\int F_n(\alpha)\mathcal{P}(\alpha)d^2\alpha \tag{1.5}$$

will converge to the limit $\langle F\rangle$ if the sequence of functions $F_n(\alpha)$ converges uniformly to the function $F(\alpha)$. Thus the expectation value $\langle F\rangle$ may be said to be a continuous functional of the function $F(\alpha)$. Expectation values calculated from tempered distributions exhibit a similar, though somewhat weaker, continuity. Members of the wider class of all distributions, in general, possess this continuity only when operating on a very special class of functions $F(\alpha)$. The general weight functional of Sudarshan, being in general neither a distribution nor a tempered distribution, does not possess this continuity property characteristic of probability distributions.

The radical extent to which the weight functions $P(\alpha)$ differ mathematically from probability distributions is a natural reflection of the fact that α is not a quantum-mechanical observable, as is evident from the relation $[a,a^\dagger]=1$. For this reason, one may only speak of the probability $W(A)$ for finding α within a range of values A large compared to the uncertainty in the measurement of α. In Sec. II, we shall consider the relationship between the probability distribution $W(A)$ and the weight function $P(\alpha)$.

It will be made evident in Sec. III that the trace relations, Eq. (1.4), impose a restriction on the asymptotic behavior as $|\alpha|\to\infty$ of $P(\alpha)$. We shall prove that the trace relations hold whenever the functions involved, $P(\alpha)(\alpha^*)^n\alpha^m$, are absolutely integrable or, more generally, whenever $P(\alpha)$ is a tempered distribution that decreases in magnitude rapidly enough for large $|\alpha|$. We shall also show that if $P(\alpha)$ is a tempered distribution then the following limit formula is valid:

$$\mathrm{Tr}[\rho(a^\dagger)^n a^m]=\lim_{\epsilon\to0+}\int e^{-\epsilon|\alpha|^2}P(\alpha)(\alpha^*)^n\alpha^m d^2\alpha. \tag{1.6}$$

In Sec. V, we shall study two equations that are generalizations of Eqs. (1.3) and (1.4):

$$\Omega=\int|\alpha\rangle\Omega(\alpha)\langle\alpha|\,d^2\alpha \tag{1.7}$$

and

$$\mathrm{Tr}[\Omega(a^\dagger)^n a^m]=\int\omega(\alpha)(\alpha^*)^n\alpha^m d^2\alpha, \tag{1.8}$$

where Ω is an arbitrary operator. We shall see that Eq.

[6] The classical stochastic techniques are described in J. Lawson and G. E. Uhlenbeck, *Threshold Noise Signals* (McGraw-Hill Book Company, Inc., New York, 1950), pp. 33-56. See also, M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Inc., London, 1959), Chap. X.

[7] Limitations on the class of density operators for which the P representation is appropriate were first discussed by R. J. Glauber in Ref. 3.

[8] Distributions and tempered distributions form classes of singular functions that include the delta function and its derivatives. Definitions are presented in the Appendix.

[9] This result was first suggested by D. Kastler and R. J. Glauber (private communication).

[10] E. C. G. Sudarshan, Phys. Rev. Letters **10**, 277 (1963), and *Proceedings of the Symposium on Optical Masers* (Polytechnic Institute of Brooklyn, 1963), p. 45. The form of the weight functional for the case of many modes is incorrectly written; a corrected version appears in Ref. 14.

(1.8), when specialized to density operators, provides a correspondence with classical optics that is universal and somewhat closer than that afforded by the P representation.

In Sec. VII, we shall exploit the analytic properties of the coherent states to obtain new measures of their completeness.

II. P REPRESENTATION AND CLASSICAL OPTICS

For a given mode of excitation of the electromagnetic field and for each complex number α, the coherent state $|\alpha\rangle$ is given by

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{(n!)^{1/2}}|n\rangle\,, \tag{2.1}$$

where $|n\rangle$ is the occupation number state having n photons in the given mode. The state $|\alpha\rangle$ is the eigenstate of the annihilation operator a with eigenvalue α, i. e., $a|\alpha\rangle=\alpha|\alpha\rangle$, as is evident from the relation $a|n\rangle=n^{1/2}|n-1\rangle$. These states are normalized but not orthogonal,

$$\langle\beta|\alpha\rangle=\exp[\beta^*\alpha-\tfrac{1}{2}(|\alpha|^2+|\beta|^2)]\,, \tag{2.2}$$

and form a complete set of states for the given mode, as is shown by Eq. (1.1).

For a multimode system, the coherent state $|\{\alpha\}\rangle$ is specified by a sequence of complex numbers α_i, one for each mode i, and is defined by $|\{\alpha\}\rangle=\prod_i|\alpha_i\rangle_i$. These multimode coherent states are eigenstates of each of the annihilation operators: $a_j|\{\alpha\}\rangle=\alpha_j|\{\alpha\}\rangle$, for all j. In terms of these states, Eqs. (1.3) and (1.4), which specify the P representation for a single mode, have the following natural extensions:

$$\rho=\int|\{\alpha\}\rangle P(\{\alpha\})\langle\{\alpha\}|\textstyle\prod_i d^2\alpha_i \tag{2.3}$$

and

$$\mathrm{Tr}[\rho(a_i^\dagger)^n a_j^m]=\int(\alpha_i^*)^n\alpha_j^m P(\{\alpha\})\textstyle\prod_k d^2\alpha_k\,. \tag{2.4}$$

The positive frequency part of the electric field operator $E_\mu^{(+)}(x)$ at the space-time point x may be expressed as a linear combination of the annihilation operators

$$E_\mu^{(+)}(x)=i\textstyle\sum_k(\frac{1}{2}\hbar\omega_k)^{1/2}a_k u_{k,\mu}(x)\,, \tag{2.5}$$

where the $u_{k,\mu}(x)$ are a complete orthonormal set of vector mode functions. The coherent state $|\{\alpha\}\rangle$ is therefore an eigenstate of $E_\mu^{(+)}(x)$, and we may write

$$E_\mu^{(+)}(x)|\{\alpha\}\rangle=\mathcal{E}_\mu(x,\{\alpha\})|\{\alpha\}\rangle\,, \tag{2.6}$$

where the function $\mathcal{E}_\mu(x,\{\alpha\})$ is obtained from the operator $E_\mu^{(+)}(x)$ by replacing the annihilation operator a_k by its eigenvalue α_k for each mode k.

The usefulness of the P representation, as well as the correspondence between it and classical optics, may be illustrated by considering the infinite hierarchy of correlation functions $G^{(n)}$, of which the first is defined[2] by

$$G_{\mu\nu}^{(1)}(x_1,x_2)=\mathrm{Tr}[\rho E_\mu^{(-)}(x_1)E_\nu^{(+)}(x_2)]\,. \tag{2.7}$$

If the P representation is used for ρ, Eq. (2.4) allows us to simplify Eq. (2.7) to

$$G_{\mu\nu}^{(1)}(x_1,x_2)=\int\mathcal{E}_\mu^*(x_1,\{\alpha\})\mathcal{E}_\nu(x_2,\{\alpha\})P(\{\alpha\})\textstyle\prod_i d^2\alpha_i\,. \tag{2.8}$$

An entirely similar simplification occurs for the higher order correlation functions, which, like $G^{(1)}$, are also expectation values of normally ordered products of $E_\mu^{(-)}$ and $E_\nu^{(+)}$.

In the classical version of Eq. (2.8), the weight function $P(\{\alpha\})$ is replaced by a probability distribution for the Fourier coefficients, $\{\alpha\}$, of the electric field vector $\mathcal{E}_\mu(x,\{\alpha\})$.[6] This is a significant restriction since the class of all weight functions $P(\{\alpha\})$ is much larger than the class of all probability distributions, as is shown in Secs. IV and VI. This wider range of mathematical possibilities is characteristic of the greater physical generality of the quantum description.

It is perhaps worth noting how the greater generality of the quantum description arises from the uncertainty principle. The classical description assumes that the electric field is precisely measurable at all points in space time and therefore that the Fourier coefficients $\{\alpha\}$ of $\mathcal{E}_\mu(x,\{\alpha\})$ can also be measured exactly. From this it follows that statistical phenomena arising from $\mathcal{E}_\mu(x,\{\alpha\})$ may be described in terms of probability distributions for these Fourier coefficients. In the quantum theory, however, the commutator $[E_\mu^{(+)}(x),E_\nu^{(-)}(y)]$ does not, in general, vanish; a measurement of the electric field at one point disturbs measurement at other points. The Fourier coefficients $\{\alpha\}$ are therefore not precisely measurable, and there is no physically meaningful probability distribution for them in the quantum theory. Consequently, the quantum theory is free of the restrictive assumption that all optical phenomena may be adequately described in terms of probability distributions for these Fourier coefficients.

For simplicity, most of the subsequent discussion will be carried out for systems having excitations in only one mode. The generalization of our results to the case of finitely many modes is straightforward. The case of infinitely many modes introduces convergence problems characteristic of integrations over infinitely many variables. These problems also occur in the classical theory, and we shall not consider them in this paper.

In our dimensionless units the uncertainty in the measurement of the parameter α is of the order of unity: $(\Delta\,\mathrm{Re}\alpha)(\Delta\,\mathrm{Im}\alpha)\approx 1$. It is therefore possible to estimate the probability $W(A)$ that a field has a value of α lying within a square A in the complex α plane whose area $\|A\|$ is large compared to unity. To do this, we shall exhibit a Hermitian operator $M(A)$ that, in an approximate[11] sense, is a projection operator for α lying within A, i.e., $M(A)|\alpha\rangle\approx|\alpha\rangle$, if α is in A; and $M(A)|\alpha\rangle\approx 0$ if α is not in A. The Hermitian operator

$$M(A)=\frac{1}{\pi}\int_A |\alpha\rangle\langle\alpha|d^2\alpha\,, \tag{2.9}$$

where the integral extends over the square A, has this property. For, from Eq. (2.2), we have

$$\langle\beta|M(A)|\beta\rangle=\frac{1}{\pi}\int_A e^{-|\alpha-\beta|^2}d^2\alpha\,, \tag{2.10}$$

and since

$$\pi=\int d^2\alpha\,\exp(-|\alpha-\beta|^2)\,,$$

then

$$\begin{aligned}\langle\beta|M(A)|\beta\rangle&\approx 1 \qquad \beta \text{ in } A\,,\\ &\approx 0 \qquad \beta \text{ not in } A\,,\end{aligned} \tag{2.11}$$

provided $\|A\|\gg 1$.

The approximate probability $W(A)$ is therefore given by the relation $W(A)=\mathrm{Tr}[\rho M(A)]$. When A is the entire complex plane, $M(A)$ is the unit operator, so that $W(A)=\mathrm{Tr}(\rho)=1$, which indicates that $W(A)$ is correctly normalized. Using the P representation for ρ and the trace relations, Eq. (1.4), we obtain

$$W(A)=\frac{1}{\pi}\int_A d^2\alpha\int d^2\beta\,P(\beta)e^{-|\alpha-\beta|^2}. \tag{2.12}$$

When the weight function $P(\alpha)$ has the mathematical properties of a probability distribution there are closer relationships between $P(\alpha)$ and $W(A)$. If $P(\alpha)$ is well behaved, we have, from Eqs. (2.10)–(2.11),

$$W(A)\approx\int_A P(\alpha)d^2\alpha. \tag{2.13}$$

If $P(\alpha)$ is non-negative and slowly varying throughout A, then we may write

$$W(A)\approx P(\alpha)\|A\|\,, \tag{2.14}$$

where α is in A.

As we have mentioned earlier, in order to describe some states of the radiation field by means of the P representation it is necessary to use extremely singular weight functions. These singular weight functions may be classified and contrasted with the probability distributions of the classical theory by studying the continuity properties of their expectation values,

$$\langle F\rangle=\int F(\alpha)P(\alpha)d^2\alpha. \tag{2.15}$$

[11] The approximation is necessarily poor for α lying near the boundary of the square A.

The relevant question is, if $F_n(\alpha)$ is a sequence of functions converging to a function $F(\alpha)$, under what conditions will the expectation values $\langle F_n\rangle$ converge to the limit $\langle F\rangle$?

If the weight function $P(\alpha)$ has all the mathematical properties of a probability distribution, we can easily show that the sequence $\langle F_n\rangle$ converges to $\langle F\rangle$, provided the functions $F_n(\alpha)$ and $F(\alpha)$ are integrable and the convergence of the sequence $F_n(\alpha)$ is uniform. Because the convergence is uniform, we have that $|F_n(\alpha)-F(\alpha)|<\epsilon$, for all α if $n>N(\epsilon)$. Since $P(\alpha)$ is assumed to be non-negative and normalized, it follows that

$$\begin{aligned}|\langle F\rangle-\langle F_n\rangle|&\leq\int P(\alpha)|F(\alpha)-F_n(\alpha)|d^2\alpha\\ &\leq\epsilon\int P(\alpha)d^2\alpha=\epsilon.\end{aligned}$$

Thus the sequence $\langle F_n\rangle$ converges to $\langle F\rangle$. In this sense, weight functions having the character of a probability distribution generate an expectation value functional $\langle F\rangle$ that is continuous with respect to the function $F(\alpha)$.

Tempered distributions[8] are defined by the requirement that they generate an expectation value functional that has a restricted form of the continuity characteristic of probability distributions. The conditions under which the sequence $\langle F_n\rangle$ converges to $\langle F\rangle$ are more restrictive in two ways. First, the functions $F_n(\alpha)$ must be infinitely differentiable and must, together with their derivatives of all orders, go to zero at infinity faster than any power of $1/|\alpha|$. Second, the convergence of the sequence of functions $F_n(\alpha)$ to the function $F(\alpha)$ must be stronger, and the derivatives of the $F_n(\alpha)$ must converge to the derivatives of $F(\alpha)$.

The class of all distributions is much wider than the class of all tempered distributions, but its greater generality does not seem relevant to the P representation. This is because distributions are required to make sense only when operating on infinitely differentiable functions $F(\alpha)$ that vanish outside some bounded region.

Some index of the restrictiveness of the classical assumption that weight functions may be replaced by probability distributions is provided by Sudarshan's[10] general weight functional $P_S(\{\alpha\})$, which, for one mode, reduces to

$$\begin{aligned}P_S(\alpha)=\sum_{n,m=0}^{\infty}&\frac{\langle n|\rho|m\rangle(n!m!)^{1/2}}{2\pi r(n+m)!}\delta^{(n+m)}(r)\\ &\times\exp[r^2+i(m-n)(\theta+\pi)]\,,\end{aligned} \tag{2.16}$$

where $\delta^{(n)}(x)\equiv(d/dx)^n\delta(x)$ and $\alpha=re^{i\theta}$. When only finitely many matrix elements, $\langle n|\rho|m\rangle$, are nonzero, $P_S(\alpha)$ is a tempered distribution. For such density operators, it does not seem possible to avoid the sort of singularities exhibited by $P_S(\alpha)$.

When infinitely many matrix elements $\langle n|\rho|m\rangle$ are nonzero, $P_S(\alpha)$ contains infinitely many derivatives of the delta function, all evaluated at the origin. However, we shall prove in the Appendix that $P_S(\alpha)$ can have only a finite number of derivatives of the delta function if it is to be a distribution. Consequently, Sudarshan's weight functional is not a distribution (and therefore not a tempered distribution), except in those special cases mentioned above.

Many of the density operators for which Sudarshan's weight functional is not a distribution possess other weight functions that are tempered distributions. As we shall show in Sec. IV, however, a density operator must satisfy a set of restrictive conditions if its weight function is to be a tempered distribution.

The radical extent to which some weight functions differ from probability distributions does not appear to have been generally recognized. In a recent paper, Mandel[12] has used the P representation to compute a probability distribution $p[V(x,\alpha)]$ for a field amplitude $V(x,\alpha)$, depending linearly upon α, according to the relation

$$p[V(x,\beta)]=\int P(\alpha)\delta[V(x,\alpha)-V(x,\beta)]d^2\alpha. \quad (2.17)$$

While this relation may be valid for those well-behaved weight functions to which Eq. (2.13) may be applied, it is not a general result.[13]

In another paper by Mandel,[14] the direct use of $P_S(\{\alpha\})$ as a probability distribution has led to the conclusion that for a stationary density operator, i.e., one for which $[\rho,H]=0$, "the expectation value of any product operator containing unpaired creation and annihilation operators $a_k^\dagger$ or a_k as factors vanishes." A more accurate result is that for a stationary density operator, expectation values of the form

$$\mathrm{Tr}[\rho\prod_k(a_k^\dagger)^{n_k}(a_k)^{m_k}] \quad (2.18)$$

vanish unless

$$\sum_k(n_k-m_k)E_k=0\,, \quad (2.19)$$

where E_k is the energy of a photon in mode k. To prove this, we note that quantities like (2.18) must be time-independent if ρ is stationary. From this and the fact that $a_k(t)=a_k(0)\exp(-iE_kt)$, we see that (2.18) must vanish unless (2.19) holds true.

It is easy to construct density operators for which (2.18) is nonzero when the degeneracy condition (2.19) is satisfied—in fact, the ordinary interference phenomena exhibited by stationary fields (e.g., in Young's double-slit experiment) are described only by such density operators. To illustrate this last assertion, let us consider the electric field intensity $\langle I(x)\rangle$ for an ensemble of plane-polarized photons. It has been shown[2] that $\langle I(x)\rangle\propto\mathrm{Tr}[\rho E^{(-)}(x)E^{(+)}(x)]$, where the polarization index has been suppressed. Using the plane-wave expansion $E^{(+)}(x)\propto\sum_l a_l\exp(ik_l\cdot x)$, we have

$$\langle I(x)\rangle\propto\sum_l\mathrm{Tr}[\rho a_l^\dagger a_l] +\sum_{j\neq l}2\cos[i(k_j-k_l)\cdot x]\,\mathrm{Tr}[\rho a_j^\dagger a_l].$$

The second summation is time-independent, if ρ is stationary, since, in view of (2.19), those terms for which $E_j\neq E_l$ vanish. The remaining terms of the second summation exhibit a three-dimensional interference pattern, and if they also vanish, $\langle I(x)\rangle$ is independent of x.

[12] L. Mandel, Phys. Letters **7**, 117 (1963).

[13] In fact, when $P_s(\alpha)$ is used for $P(\alpha)$, it follows directly from Eqs. (2.16)–(2.17) that $p[V(x,\beta)]=0$ for $V(x,\beta)\neq V(x,0)$.

[14] L. Mandel, Phys. Letters **10**, 166 (1964).

III. TRACE RELATIONS

In this section, we shall present conditions under which a weight function $P(\alpha)$ satisfies the trace relations, Eq. (1.4). These relations are essential both to the analogy with classical optics and to the usefulness of the P representation as a computational device.

An examination of Eq. (1.4) suggests that if the expectation value $\mathrm{Tr}[\rho(a^\dagger)^na^n]$ is finite, the weight function $P(\alpha)$ must, as $|\alpha|\to\infty$, tend to zero faster than $|\alpha|^{-(2n+1)}$. If these traces are finite for all n, it would seem that $P(\alpha)$ must go to zero at infinity faster than any inverse power of $|\alpha|$. For a wide class of weight functions, these asymptotic conditions are necessary for the trace relations to hold. They are also sufficient if $P(\alpha)$ is well behaved, as the following result will show.

Let us assume that the P representation, Eq. (1.3), exists for the density operator ρ and that $P(\alpha)$ is a well-behaved, i.e., measurable, weight function. In this case, we shall now show that the trace relation

$$\mathrm{Tr}[\rho(a^\dagger)^na^m]=\int(\alpha^*)^n\alpha^mP(\alpha)d^2\alpha \quad (3.1)$$

follows from Eq. (1.3), provided the Lebesgue integral $\int|P(\alpha)|\,|\alpha|^{n+m}d^2\alpha$ exists and is finite.

By the definition of the trace we have

$$\mathrm{Tr}[\rho(a^\dagger)^na^m]=\sum_{p=0}^{\infty}\langle p|a^m\rho(a^\dagger)^n|p\rangle$$

$$\sum_{p=0}^{\infty}\langle p+m|\rho|p+n\rangle\frac{[(n+p)!(m+p)!]^{1/2}}{p!},$$

which, using Eqs. (1.3) and (2.1), becomes

$$\mathrm{Tr}[\rho(a^\dagger)^na^m]=\sum_{p=0}^{\infty}\int\frac{|\alpha|^{2p}}{p!}P(\alpha)e^{-|\alpha|^2}(\alpha^*)^n\alpha^md^2\alpha. \quad (3.2)$$

Now the sequence of functions

$$S_i(\alpha)=(\alpha^*)^n\alpha^me^{-|\alpha|^2}\sum_{p=0}^{i}\frac{|\alpha|^{2p}}{p!} \quad (3.3)$$

converges to $S(\alpha)=(\alpha^*)^n\alpha^m$, and we have the inequality $|S_i(\alpha)| \le |S(\alpha)|$ for all α and all i. Then, by the Lebesgue dominated convergence theorem,[15] the interchange

$$\lim_{i\to\infty} \int S_i(\alpha)P(\alpha)d^2\alpha = \int S(\alpha)P(\alpha)d^2\alpha \qquad (3.4)$$

is valid, provided the integral $\int |S(\alpha)P(\alpha)|d^2\alpha$ exists and is finite. Equations (3.2) and (3.4) directly imply Eq. (3.1) and our proof is complete. This proof is easily generalized to cover the case of finitely many modes.

We now consider the case in which $P(\alpha)$ is not well behaved. We shall show that Eq. (1.3) implies Eq. (3.1), for all n and m, provided that for some arbitrarily small positive constants ϵ and δ the function

$$T(\alpha) = P(\alpha)\exp(\epsilon|\alpha|^\delta) \qquad (3.5)$$

is a tempered distribution. Our proof will utilize the continuity property of the tempered distribution $T(\alpha)$. We let $f_i(\alpha) = S_i(\alpha)\exp(-\epsilon|\alpha|^\delta)$, where $S_i(\alpha)$ is given by Eq. (3.3). It is easy to see that $f_i(\alpha)$ converges to $f(\alpha) = S(\alpha)\exp(-\epsilon|\alpha|^\delta)$ and that the convergence is uniform on every bounded set. Since $|f_i(\alpha)| \le |f(\alpha)| = |\alpha|^{n+m}\exp(-\epsilon|\alpha|^\delta)$, it is clear that we may find constants $C(r,s,u,v)$ such that, for all i, r, s, u, and v,

$$\left| x^r y^s \frac{\partial^{u+v}}{\partial x^u \partial y^v} f_i(\alpha) \right| \le C(r,s,u,v), \qquad (3.6)$$

where $\alpha = x+iy$. Then, by the continuity of $T(\alpha)$,[16] we have

$$\lim_{i\to\infty} \int T(\alpha)f_i(\alpha)d^2\alpha = \int T(\alpha)f(\alpha)d^2\alpha = \int S(\alpha)P(\alpha)d^2\alpha, \qquad (3.7)$$

which, together with Eq. (3.2), implies Eq. (3.1). Since n and m are arbitrary, our proof is complete. For the case of finitely many modes, one must require that condition (3.5) hold for each mode.

We shall now show that if $P(\alpha)$ is a tempered distribution, the formula

$$\mathrm{Tr}[\rho(a^\dagger)^n a^m] = \lim_{\epsilon\to 0+} \int e^{-\epsilon|\alpha|^2}(\alpha^*)^n\alpha^m P(\alpha)d^2\alpha \qquad (3.8)$$

is valid, provided $|\mathrm{Tr}[\rho(a^\dagger)^n a^m]| < \infty$. For each pair of non-negative integers, n and m, we define the function

$$Q(s,n,m) = \int e^{-s|\alpha|^2}(\alpha^*)^n\alpha^m P(\alpha)d^2\alpha, \qquad (3.9)$$

where s is a complex variable. As we show in the Appendix, the function $Q(s,n,m)$ is analytic throughout the right half-plane, $\mathrm{Re}s>0$, and in this region the kth derivative is given by

$$Q^{(k)}(s,n,m) = \int (-|\alpha|^2)^k e^{-s|\alpha|^2}(\alpha^*)^n\alpha^m P(\alpha)d^2\alpha. \qquad (3.10)$$

But, by Eqs. (1.3) and (2.1), we have

$$Q^{(k)}(1,n,m) = (-1)^k k!\langle k|a^m\rho(a^\dagger)^n|k\rangle. \qquad (3.11)$$

Then, since $Q(s,n,m)$ is analytic for $\mathrm{Re}s>0$, we have for $|s-1|<1$ that

$$Q(s,n,m) = \sum_{k=0}^{\infty}(1-s)^k\langle k|a^m\rho(a^\dagger)^n|k\rangle. \qquad (3.12)$$

If the trace $\mathrm{Tr}[\rho(a^\dagger)^n a^m]$ is finite, then the power series for $Q(s,n,m)$ converges also at $s=0$, and, by Stoltz's theorem,[17] we may conclude that

$$\mathrm{Tr}[\rho(a^\dagger)^n a^m] = \lim_{\epsilon\to 0+} Q(\epsilon,n,m),$$

which is just Eq. (3.8). This result and its proof are easily generalized to cover the case of finitely many modes.

IV. SINGULAR NATURE OF SOME WEIGHT FUNCTIONS

In this section, we shall show that the weight functions corresponding to some density operators are extremely singular. As we have mentioned earlier, this is an evident limitation both upon the computational usefulness of the P representation and upon the analogy between it and classical optics.

The assumption that $P(\alpha)$ is a tempered distribution leads to certain restrictive conditions on the density operator ρ, as we shall now show. If $P(\alpha)$ is a tempered distribution, then the functions $Q(s,n,m)$, defined by Eq. (3.9), are analytic for $\mathrm{Re}s>0$, and Eqs. (3.10)–(3.12) hold true. Now the power series for $Q(s,n,m)$, Eq. (3.12), which converges for $|s-1|<1$, may be rewritten as

$$Q(s,n,m) = \sum_{k=0}^{\infty}\frac{(1-s)^k}{k!}[(k+m)!(k+n)!]^{1/2} \times\langle k+m|\rho|k+n\rangle. \qquad (4.1)$$

Since $Q(s,n,m)$ is analytic for $\mathrm{Re}s>0$, this power series must have an analytic continuation $Q(s,n,m)$ throughout the right half-plane, $\mathrm{Re}s>0$. This is true for all non-negative integers n and m.

These are conditions on the matrix elements $\langle n|\rho|m\rangle$ that by no means all density operators satisfy. For example, if $\langle 2n|\rho|2n\rangle = (1-x)x^n$ and $\langle 2n+1|\rho|2n+1\rangle = 0$,

[15] W. Rudin, *Principles of Mathematical Analysis* (McGraw-Hill Book Company, Inc., New York, 1953), p. 209.

[16] This follows directly from the definition of a tempered distribution, which is given in the Appendix.

[17] E. Hille, *Analytic Function Theory* (Blaisdell Publishing Company, New York, 1963), Vol. I, p. 123.

where $1>x>0$, then $Q(s,0,0)=(1-x)[1-(1-s)^2x]^{-1}$, which has a pole at $s=1+x^{-1/2}$.

If $P(\alpha)$ is a tempered distribution, the analytic continuation of the power series of Eq. (4.1) must also satisfy the asymptotic condition

$$\lim_{|s|\to\infty} Q(|s|,n,m)|s|^{-N}=0\,,\quad \text{for some } N(n,m)\,. \tag{4.2}$$

This is true for all non-negative integers, n and m. As we shall show in the Appendix, this follows directly from Eq. (3.9) and the continuity of the tempered distribution $P(\alpha)$.

An example of a density operator that violates these conditions is one for which $\langle 2n|\rho|2n\rangle=(n!)^{-1}x^ne^{-x}$ and $\langle 2n+1|\rho|2n+1\rangle=0$, where $x>0$. For in this case $Q(s,0,0)=\exp[sx(s-2)]$.

Since some density operators do not possess weight functions that are tempered distributions, it is reasonable to consider whether these density operators may be represented in some more general way, e.g., in terms of a limiting process. We shall now mention a difficulty connected with representing density operators in terms of a sequence of tempered distributions $P_n(\alpha)$, in the form[18]

$$\rho=\lim_{n\to\infty}\int|\alpha\rangle P_n(\alpha)\langle\alpha|d^2\alpha\,. \tag{4.3}$$

It has been shown[19] that a sequence of tempered distributions $P_n(\alpha)$ is convergent if and only if, for every test function $F(\alpha)$, the sequence of expectation values $\langle F\rangle_{P_n}$ is convergent. Consequently, if ρ is a density operator that cannot be represented by a tempered distribution but for which the limiting process (4.3) is valid, then there will be test functions $F(\alpha)$ such that the limit

$$\lim_{n\to\infty}\int P_n(\alpha)F(\alpha)d^2\alpha$$

will not exist.

If a weight function $P(\alpha)$ is to bear any close mathematical resemblance to a probability distribution, it can have, at worst, delta-function singularities. We shall now show, under quite general assumptions, that if $P(\alpha)$ has only delta-function singularities, then the analytic continuation of the power series, $Q(s,0,0)$, of Eq. (4.1) must obey a strong asymptotic condition.

Performing the angular integration in Eq. (3.9) and letting $x=|\alpha|^2$ and $P'(x)=\pi\int_0^{2\pi}P(\alpha)d\theta$, we obtain

$$Q(s,0,0)=\int_0^\infty e^{-sx}P'(x)dx\,. \tag{4.4}$$

We shall not assume that $P(\alpha)$ is a tempered distribution but shall rather assume that, for some $t<1$, the function

$$\beta(y)=\int_0^y e^{-tx}P'(x)dx$$

is of bounded variation on every finite portion of the interval $[0,\infty\rangle$ and that $\beta(y)$ converges as $y\to\infty$. Under these assumptions, it may be shown[20] that $Q(s,0,0)$ is analytic for $\mathrm{Re}\,s>t$ and that, as before, $Q^{(n)}(1,0,0)=(-1)^n\langle n|\rho|n\rangle n!$. Consequently, the power series of Eq. (4.1) may be analytically continued to the function $Q(s,0,0)$ for $\mathrm{Re}\,s>t$. Since $\beta(y)$ is of bounded variation, it follows that the limit

$$\lim_{|s|\to\infty} Q(|s|,0,0) \tag{4.5}$$

exists and is finite.

V. ALTERNATIVE COHERENT STATE REPRESENTATION

We shall now show that, for an arbitrary operator Ω, Eqs. (1.7) and (1.8) always have solutions $\Omega(\alpha)$ and $\omega(\alpha)$ that possess formal Fourier expansions in terms of relatively well-behaved functions. We let $\alpha=re^{i\theta}$ and Fourier-analyze $\Omega(\alpha)$ and $\omega(\alpha)$:

$$\begin{aligned}\Omega(\alpha)&=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}e^{in\theta}\Omega_n(r)\,,\\ \omega(\alpha)&=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}e^{in\theta}\omega_n(r)\,.\end{aligned} \tag{5.1}$$

Substituting into Eqs. (1.7)–(1.8), we have, for $m\geqslant 0$ and $n+m\geqslant 0$,

$$[m!(n+m)!]^{1/2}\langle m|\Omega|m+n\rangle=\int_0^\infty r^{2m+n+1}\Omega_n(r)e^{-r^2}dr \tag{5.2}$$

and

$$\mathrm{Tr}[\Omega(a^\dagger)^{n+m}a^m]=\int_0^\infty r^{2m+n+1}\omega_n(r)dr\,. \tag{5.3}$$

It is important to note that for all pairs of integers n and m, with fixed n, the left-hand sides of Eqs. (5.2) and (5.3) depend only upon $\Omega_n(r)$ and $\omega_n(r)$, respectively. For each integer n, we form the sequences

$$\mathrm{M}(n,m)=[m!(n+m)!]^{1/2}\langle m|\Omega|m+n\rangle$$

and

$$\mu(n,m)=\mathrm{Tr}[\Omega(a^\dagger)^{n+m}a^m]\,,$$

where each sequence starts at the first non-negative integer m such that $n+m\geqslant 0$.

Now, according to the Stieltjes-Boas theorem,[21] there

---

[18] Limiting procedures of this type were introduced by J. R. Klauder, J. McKenna, and D. Currie (to be published) and by C. L. Mehta and E. C. G. Sudarshan, Phys. Rev. **138**, B274 (1965); both procedures are subject to the limitation mentioned above.

[19] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (W. A. Benjamin, Inc., New York, 1964), p. 34.

[20] D. V. Widder, *The Laplace Transform* (Princeton University Press, Princeton, New Jersey, 1946), p. 57.

exist functions $B_n(r)$ and $\beta_n(r)$ of bounded variation such that

$$M(n,m)=\int_0^\infty r^{2m+n+1}dB_n(r)$$

and

$$\mu(n,m)=\int_0^\infty r^{2m+n+1}d\beta_n(r)\,.$$

This is true for all n, as long as all the elements $M(n,m)$ and $\mu(n,m)$ are finite as we have implicitly assumed. Setting $\Omega_n(r)=e^{r^2}(dB_n(r)/dr)$ and $\omega_n(r)=d\beta_n(r)/dr$, we we see that $\Omega_n(r)$ and $\omega_n(r)$ contain, at most, a countable number of singularities of the form $\delta(r-r')$, since $B_n(r)$ and $\beta_n(r)$ are of bounded variation. An additional consequence of the Stieltjes-Boas theorem is that the integrals $\int_0^\infty|\Omega_n(r)|e^{-r^2}dr$ and $\int_0^\infty|\omega_n(r)|dr$ are finite for all n. The presence of the damping factor e^{-r^2} in the integral of $|\Omega_n(r)|$ reflects the mathematical difficulties associated with the universal use of the P representation.

The Fourier series expansions for $\Omega(\alpha)$ and $\omega(\alpha)$, Eq. (5.1), are purely formal, and we shall not investigate the question of their convergence. Instead, we shall exhibit the extent to which these series, when cut off at the points $n=\pm N$, still respect Eqs. (1.7)–(1.8). For each $N\geqslant 0$, we define the cutoff series $\Omega(\alpha,N)$ and $\omega(\alpha,N)$ by Eq. (5.1) with the proviso that the summation extend only from $n=-N$ to $n=+N$, i.e.,

$$\Omega(\alpha,N)=\frac{1}{2\pi}\sum_{n=-N}^{N}e^{in\theta}\Omega_n(r)\,, \tag{5.4}$$

with a similar equation for $\omega(\alpha,N)$.

Now, from Eq. (5.2), it follows that the operator Ω_N defined by

$$\Omega_N=\int|\alpha\rangle\Omega(\alpha,N)\langle\alpha|d^2\alpha\,, \tag{5.5}$$

differs from Ω only in that for $|n-m|>N$, $\langle n|\Omega_N|m\rangle=0$. Consequently, if the trace $\mathrm{Tr}[\Omega^\dagger\Omega]$ is finite, it is easy to check that as $N\to\infty$, $\mathrm{Tr}[(\Omega-\Omega_N)^\dagger(\Omega-\Omega_N)]\to 0$, and therefore that the convergence of Ω_N to Ω is strong.

From Eq. (5.3), it follows that if $N\geqslant|n-m|$, then the exact value of the trace $\mathrm{Tr}[\Omega(a^\dagger)^n a^m]$ is given by

$$\mathrm{Tr}[\Omega(a^\dagger)^n a^m]=\int\omega(\alpha,N)(\alpha^*)^n\alpha^m d^2\alpha\,, \tag{5.6}$$

without any limiting process.

These results are easily generalized to the case of finitely many modes.

Let us now consider the analog of Eq. (5.6) for the case in which Ω is a density operator describing a field of finitely many modes. The relations $\Omega^\dagger=\Omega$ and $\mathrm{Tr}(\Omega)=1$ imply that the "trace function" $\rho_N(\{\alpha\})\equiv\omega(\{\alpha\},N)$ is real and normalized to unity for all N. Since the correlation function $G^{(n)}$ contains n factors of $E^{(-)}$ and n factors of $E^{(+)}$, it is clear that $\rho_N(\{\alpha\})$ may be used to represent those correlation functions $G^{(n)}$ for which $n\leq N$. Thus, for example, Eq. (2.8) may be replaced by

$$G_{\mu\nu}^{(1)}(x_1,x_2)=\int\mathcal{E}_\mu^*(x_1,\{\alpha\})\mathcal{E}_\nu(x_2,\{\alpha\})\rho_N(\{\alpha\})\textstyle\prod_i d^2\alpha_i\,, \tag{5.7}$$

where $N\geqslant 1$. Since the trace functions $\rho_N(\{\alpha\})$ are tempered distributions with no singularities worse than $\delta(r-r')$, this representation for the correlation functions would seem to suggest a correspondence with classical optics that is universal and perhaps somewhat closer than that afforded by the P representation.

This correspondence is also limited, however, since, as we shall show in Sec. VI, $\rho_N(\{\alpha\})$ is not, in general, positive definite. In addition, as the results of Sec. II indicate, $\rho_N(\{\alpha\})$ must not be thought of as a probability distribution for the Fourier coefficients $\{\alpha\}$, except perhaps in an approximate sense for states near the classical limit.

VI. POSITIVE DEFINITENESS

In this section, we shall present restrictions on the density operator ρ, if its weight function $P(\alpha)$ or its trace function $\rho_N(\alpha)$ is to be non-negative. We shall show that the requirement $P(\alpha)\geqslant 0$ excludes all pure states, except the coherent states themselves. We recall that Eq. (2.14), which expresses a close relationship between the probability function $W(A)$ and the weight function $P(\alpha)$, is valid only when $P(\alpha)\geqslant 0$.

Let $A(\alpha)$ be any non-negative polynomial in the complex variables α and α^*, i.e.,

$$A(\alpha)=\sum_{r,s=0}^{N}a(r,s)\alpha^r(\alpha^*)^s\geqslant 0\,. \tag{6.1}$$

Then, if $P(\alpha)\geqslant 0$, we have

$$\sum_{r,s=0}^{N}a(r,s)\langle s|\rho|r\rangle(r!s!)^{1/2}=\int A(\alpha)P(\alpha)e^{-|\alpha|^2}d^2\alpha\geqslant 0; \tag{6.2}$$

and, if $\rho(\alpha)\geqslant 0$, we have

$$\sum_{r,s=0}^{N}a(r,s)\,\mathrm{Tr}[\rho(a^\dagger)^s a^r]=\int A(\alpha)\rho(\alpha)d^2\alpha\geqslant 0\,. \tag{6.3}$$

These conditions have an obvious generalization for the case of finitely many modes.

We shall illustrate the restrictiveness of Eqs. (6.2)–(6.3) by showing that the pure state density operator $\rho=|1\rangle\langle 1|$ cannot have either a non-negative weight function or a non-negative trace function. Letting $A(\alpha)=(1-|\alpha|^2)^2$, we have

$$\langle 0|\rho|0\rangle-2\langle 1|\rho|1\rangle+2\langle 2|\rho|2\rangle=-2$$

and

$$\mathrm{Tr}(\rho)-2\,\mathrm{Tr}(\rho a^\dagger a)+\mathrm{Tr}[\rho(a^\dagger)^2a^2]=-1\,,$$

which violate Eqs. (6.2)–(6.3).

We shall now show that if ρ represents a pure state and $P(\alpha)\geqslant 0$, then $\rho=|\alpha'\rangle\langle\alpha'|$ for some α'. If ρ represents any pure state, then we have $\rho^2=\rho$ and

$$\mathrm{Tr}(\rho^2)=1=\int P(\alpha)d^2\alpha\,. \tag{6.4}$$

But we may also write

$$\mathrm{Tr}(\rho^2)=\mathrm{Tr}\left[\int|\alpha\rangle\langle\alpha|\beta\rangle\langle\beta|P(\alpha)P(\beta)d^2\alpha d^2\beta\right]$$

$$=\int e^{-|\alpha-\beta|^2}P(\alpha)P(\beta)d^2\alpha d^2\beta\,. \tag{6.5}$$

Now if $P(\alpha)\geqslant 0$, Eqs. (6.4) and (6.5) lead to the contradiction $\mathrm{Tr}(\rho^2)<1$, unless $P(\alpha)=\delta(\alpha-\alpha')\equiv\delta[\mathrm{Re}(\alpha-\alpha')]\times\delta[\mathrm{Im}(\alpha-\alpha')]$ for some α', in which case $\rho=|\alpha'\rangle\langle\alpha'|$. Thus the only pure states representable by non-negative weight functions are the coherent states themselves.

VII. SUPERCOMPLETENESS OF THE COHERENT STATES

We shall show that the coherent states $|\alpha\rangle$ are "supercomplete," in the sense that if α_n is any convergent sequence of complex numbers, then the corresponding coherent states $|\alpha_n\rangle$ are themselves complete.

Let $\langle f|$ be any normalized state vector that is orthogonal to all the $|\alpha_n\rangle$, i.e.,

$$\langle f|\alpha_n\rangle=0;\quad n=0,1,2,\cdots\infty\,. \tag{7.1}$$

We shall show that $\langle f|=0$.

Since the states $|n\rangle$ are complete, the expansion $\langle f|=\sum_n\langle f|n\rangle\langle n|$ is valid, and we have $\sum_n|\langle f|n\rangle|^2=1$. Then the function

$$f(\alpha)=e^{\frac{1}{2}|\alpha|^2}\langle f|\alpha\rangle=\sum_{n=0}^{\infty}\langle f|n\rangle\frac{\alpha^n}{(n!)^{1/2}} \tag{7.2}$$

is entire. But by Eq. (7.1), we find that $f(\alpha_n)=0$ on the convergent sequence of complex numbers, α_n, and hence that $f(\alpha)$ vanishes identically. Thus $\langle f|n\rangle=0$, for all n, and $\langle f|=0$. Since only the null vector is orthogonal to the states $|\alpha_n\rangle$ for all n, the states $|\alpha_n\rangle$ are complete, i.e., they span the Hilbert space of the occupation number states $|n\rangle$.

As a consequence of this result, any vector $\langle g|$ in the Hilbert space spanned by the states $|n\rangle$ is uniquely determined by the scalar products $\langle g|\alpha_n\rangle$, provided the sequence α_n is convergent. Similarly, any Hilbert space operator Ω is uniquely determined by the matrix elements $\langle\alpha_n|\Omega|\alpha_m\rangle$.

We shall now show that if Ω is an operator such that $\langle\alpha|\Omega|\beta\rangle$ is finite for all α and β, then Ω is uniquely determined by the diagonal elements, $\langle x_n+iy_m|\Omega|x_n+iy_m\rangle$, where x_n and y_m are any two convergent sequences of real numbers.

To prove this, we shall show that if

$$\langle x_n+iy_m|\Omega|x_n+iy_m\rangle=0\,,$$

then under these conditions $\Omega=0$. The function $f(x,y)=\langle x+iy|\Omega|x+iy\rangle$ may be written as an everywhere convergent power series in the complex variables α, α^* or in the real variables x, y, since

$$\langle\alpha|\Omega|\alpha\rangle=e^{-|\alpha|^2}\sum_{n,m=0}^{\infty}\langle n|\Omega|m\rangle\frac{(\alpha^*)^n\alpha^m}{(n!m!)^{1/2}}\,. \tag{7.3}$$

Therefore, as a function of the two complex variables z, w, the function $f(z,w)$ is entire.[22] Consequently, if $f(z,w)=0$ for $z=x_n$ and $w=y_m$ for each n and m, where x_n and y_m are two convergent sequences of real numbers, $f(z,w)$ vanishes identically. Hence, by Eq. (7.3), $\Omega=0$.

A special case of this theorem is that the matrix elements $\langle\alpha|\Omega|\alpha\rangle$, where α ranges over any region of the complex α plane having nonzero area, uniquely determine the operator Ω. The weaker result that Ω is determined by the elements $\langle\alpha|\Omega|\alpha\rangle$, for all α, was first suggested by Jordan.[23]

ACKNOWLEDGMENTS

The author is grateful to Professor R. J. Glauber for an introduction to this subject, provided by his course on Quantum Optics, and for many valuable conversations. He is also indebted to W. A. Drumin for a discussion of measure theory.

APPENDIX

In this Appendix we shall define the terms distribution and tempered distribution and shall then use these defintions to prove three assertions that were made in Secs. II, III, and IV.

Let S denote the set of all infinitely differentiable functions $\phi(x)$ that, together with their derivatives of all orders, go to zero, as $|x|\to\infty$, faster than any power of $1/|x|$. The members of S are the "test functions" on which tempered distributions are defined to operate. A sequence of test functions $\phi_n(x)$ is said to *converge in S* to the function $\phi(x)$, if in every bounded region the derivatives of all orders of the $\phi_n(x)$ converge uniformly to the corresponding derivatives of $\phi(x)$ and if there are

[21] D. V. Widder, Ref. 20, p. 139.

[22] S. Bochner and W. T. Martin, *Several Complex Variables* (Princeton University Press, Princeton, New Jersey, 1948), p. 34.

[23] An incomplete proof of this result was first given by T. F. Jordan, Phys. Rev. Letters, **12**, 607 (1964).

constants C_{kq} such that, for all x, n, k, and q,

$$|x^k(d^q/dx^q)\phi_n(x)|\leq C_{kq}. \tag{A1}$$

The set K of test functions for distributions is the set of all infinitely differentiable functions $\phi(x)$ that vanish outside some bounded region, which may be different for each of the $\phi(x)$. A sequence of test functions $\phi_n(x)$ is said to *converge in* K to the function $\phi(x)$, if the functions $\phi_n(x)$ all vanish outside some fixed bounded region, the same for all of them, and if their derivatives of all orders converge uniformly to the corresponding derivatives of $\phi(x)$. A sequence that converges in S (in K) to the null function $o(x)=0$ is said to *converge to zero in* S (*in* K).

Distributions (*tempered distributions*) are continuous linear functionals on K (on S). More explicitly, a functional $f(x)$ that associates with each $\phi(x)$ in K (in S) a number $\langle\phi\rangle_f$, given symbolically by $\langle\phi\rangle_f=\int\phi(x)f(x)dx$, is said to be a distribution (tempered distribution), if the following conditions are satisfied: (1) If $\phi(x)$ and $\psi(x)$ are in K (in S) and α and β are complex numbers, then $\langle\alpha\phi+\beta\psi\rangle_f=\alpha\langle\phi\rangle_f+\beta\langle\psi\rangle_f$. (2) If $\phi_n(x)$ is a sequence of test functions that converges in K (in S) to the function $\phi(x)$, then $\lim_{n\to\infty}\langle\phi_n\rangle_f=\langle\phi\rangle_f$. We note that these definitions imply that every tempered distribution is a distribution.

We shall now prove that if $P(x)$ is a tempered distribution, then the function

$$Q(s,n)=\int_{-\infty}^{\infty}P(x)x^ne^{-sx^2}dx \tag{A2}$$

is analytic for $\mathrm{Re}s>0$ for each integer $n\geqslant 0$ and that its kth derivative is given by

$$Q^{(k)}(s,n)=(-1)^k\int_{-\infty}^{\infty}P(x)x^{n+2k}e^{-sx^2}dx. \tag{A3}$$

Now it is clear that $\phi_{n,s}(x)\equiv x^ne^{-sx^2}$ is in S for all $n\geqslant 0$, so that Eqs. (A2) and (A3) are well defined. We must show that for each fixed s, $\mathrm{Re}s>0$, and each sequence $s'\to s$,

$$\lim_{s'\to s}\left[\frac{Q^{(k)}(s',n)-Q^{(k)}(s,n)}{s'-s}-Q^{(k+1)}(s,n)\right]=0, \tag{A4}$$

for all n and k, where $Q^{(k)}(s,n)$ is given by Eq. (A3). Since $P(x)$ is a tempered distribution, it is sufficient to show that for all integers $n\geqslant 0$ and for each sequence $s'\to s$, the sequence of test functions, $\psi_{n,s'}(x)\equiv(s'-s)^{-1}\times[\phi_{n,s'}(x)-\phi_{n,s}(x)]+\phi_{n+2,s}(x)$, converges to zero in S. Now it is easy to verify that

$$\psi_{n,s'}(x)=(s'-s)x^{n+4}e^{-sx^2}r_{s'}(x),$$

where

$$r_{s'}(x)=\sum_{m=2}^{\infty}(s-s')^{m-2}\frac{x^{2(m-2)}}{(n!)}. \tag{A5}$$

Because of the factor e^{-sx^2}, the functions $\psi_{n,s'}(x)$ satisfy Eq. (A1), since for $|s-s'|<\epsilon$ the constants C_{kq} may be chosen independent of s'. Now, in any bounded region $|x|<R$, the functions $r_{s'}(x)$ converge uniformly to $\frac{1}{4}$ as $s'\to s$, and all their derivatives converge uniformly to zero. For $|x|<R$ and fixed s, $\mathrm{Re}s>0$, the function $x^{n+4}e^{-sx^2}$ and all its derivatives are bounded. Hence $\psi_{n,s'}(x)$ and all its derivatives approach zero uniformly for $|x|<R$ as $s'\to s$. Consequently, for each n, the functions $\psi_{n,s'}(x)$ converge to zero in S for each sequence $s'\to s$, $\mathrm{Re}s>0$, Eq. (A4) is established, and our proof is complete.

Equations (3.9)–(3.10) are the two-variable generalizations of this result. The arithmetic is somewhat more involved, but the argument is identical to that given here.

We shall now show that if $P(x)$ is a tempered distribution, then for all n the functions $Q(s,n)$, given by Eq. (A2), must satisfy the asymptotic condition

$$\lim_{|s|\to\infty}Q(|s|,n)|s|^{-N}=0,\quad\text{for some } N(n). \tag{A6}$$

We note that the functions $\theta_{n,s}(x)\equiv x^ne^{-|s|x^2}$ and all their derivatives converge uniformly to zero, as $|s|\to\infty$, in the region $|x|>r$, for any $r>0$. Thus it is easy to verify that constants $C(k,q,n)$ may be found such that

$$|x^k(d^q/dx^q)\theta_{n,s}(x)|\leq|s|^qC(k,q,n),$$

for $|s|>R$. But if $Q(s,n)$ violates condition (A6), then we may find a sequence s_q such that $|Q(|s_q|,n)|\geqslant q|s_q|^qC(k,q,n)$. Thus the sequence

$$\epsilon_{q,n}(x)\equiv|Q(|s_q|,n)|^{-1}\theta_{n,s_q}(x)$$

converges to zero in S, as $q\to\infty$. Hence the sequence $\delta_{q,n}\equiv\int P(x)\epsilon_{q,n}(x)dx$ converges to zero, as $q\to\infty$. But this is a contradiction, since by Eq. (A2), $|\delta_{q,n}|=1$, and our proof is complete.

The proof of Eq. (4.2), which is the two-variable generalization of this result, follows this argument very closely.

We shall now show that Sudarshan's functional, Eq. (2.16), is not a distribution unless $\langle n|\rho|m\rangle=0$ for $n>N$ and $m>N$, where N is a finite integer. The relevant structure of $P_S(\alpha)$ is exhibited by

$$P_S'(x)=\sum_{n=0}^{\infty}c_n\delta^{(n)}(x), \tag{A7}$$

where $\delta^{(n)}(x)\equiv(d^n/dx^n)\delta(x)$ and where an infinite number of the c_n are different from zero whenever the matrix $\langle n|\rho|m\rangle$ has an infinite number of nonzero entries.

Let $U(x)$ be an infinitely differentiable (Urysohn) function such that $U(x)=1$, if $|x|\leq\frac{1}{3}$; $U(x)=0$, if $|x|\geqslant\frac{2}{3}$; and $|U(x)|\leq 1$, for all x.[24] We define $\phi_n(x)=0$, if $c_n=0$; $\phi_n(x)=U(x)x^n(n!)^{-1}$, if $|c_n|\geqslant 1$; and $\phi_n(x)$

[24] The existence of such functions is proven in I. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Academic Press Inc., New York, 1964), Vol. I, p. 142.

$=c_n^{-1}U(c_n^{-1}x)x^n(n!)^{-1}$, if $0<|c_n|<1$. The sequence $\phi_n(x)$ has been explicitly constructed so that it converges to zero in K, as $n\to\infty$, as may easily be verified. Then, if $P_S'(x)$ is a distribution, we have

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}P_S'(x)\phi_n(x)dx=0. \qquad \text{(A8)}$$

However, from Eq. (A7), we have as well

$$\left|\int_{-\infty}^{\infty}P_S'(x)\phi_n(x)dx\right|\geqslant 1, \qquad \text{(A9)}$$

unless $c_n=0$. Thus there is a contradiction, unless $c_n=0$ for $n>N$, and our proof is complete.

PHYSICAL REVIEW VOLUME 138, NUMBER 6B 21 JUNE 1965

Classical Relativistic Mechanics of Interacting Point Particles

H. VAN DAM* AND E. P. WIGNER†
Summer School of the University of Wisconsin, Madison, Wisconsin
(Received 15 February 1965)

The possibility of formulating a classical relativistically invariant mechanics of an arbitrary number of interacting point particles is demonstrated. This theory is similar to Newtonian mechanics inasmuch as the interaction between any pair of particles contains an arbitrary function of their distance. The conservation laws for energy and linear and angular momenta are established in the sense that the sum of these quantities for the particles entering a collision is equal to the corresponding sum for the collision products. The possibility of such a mechanics contrasts with the impossibility, demonstrated recently, of establishing a relativistic mechanics within the framework of the canonical representations of the Lorentz group.

1. INTRODUCTION

THERE appears to be a body of opinion according to which an interaction between point particles must necessarily be via a field if the theory is to be invariant under Lorentz transformations.[1] This field is thought to be essential and, in collisions between the particles, the field can carry away energy, linear momentum, and angular momentum. Also there exists a wide spread opinion according to which the interaction between point particles must be via "signals," the velocity of which does not exceed that of light, if the theory is to be Lorentz-invariant. The present considerations show that these opinions must be revised. The equations which we shall consider postulate that a particle "interacts" at a definite space-time point with those points of the other orbits which are spacelike with respect to the point under consideration (see Fig. 1). The interaction is analogous to Newton's "*actio in distantia*" and reduces to the gravitational interaction in the nonrelativistic limit and if the distances between the particles (in the usual sense) are always "large."

The possibility of such a mechanics contrasts with the impossibility, demonstrated recently,[2] of establishing a relativistic mechanics within the framework of the canonical representations of the Lorentz group.

2. KINEMATICS

It is our aim to describe equations of motion for interacting point particles which satisfy the following postulates:

1. The equations are Lorentz-invariant in the sense that the orbits transform as expected under proper inhomogeneous Lorentz transformations.
2. The rest mass of each of the particles is constant.
3. The total energy, total linear momentum, and total angular momentum are conserved in the asymptotic sense, i.e., they have the same value when the particles have separated after a collision as they had before the particles came close together. The law for the motion of the center of mass should be valid in the same sense.

We shall also make the assumption that there are $6n$ degrees of freedom, where n is the number of particles. This means that $6n$ independent position and velocity components completely characterize the orbits.

In order to give the equations of motion in a manifestly Lorentz-invariant form, it is convenient to introduce a proper time τ_i for each of the mass points and give the orbits parametrically, in terms of these proper

* University of North Carolina, Chapel Hill, North Carolina.
† Oak Ridge National Laboratory, Oak Ridge, Tennessee.

[1] L. D. Landau and E. Lifshitz, *The Classical Theory of Fields*, (Addison-Wesley Publishing Company, Inc., Cambridge, Massachusetts, 1951), p. 41.

[2] D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. **35**, 350 (1963); also D. G. Currie, J. Math. Phys. **4**, 1470 (1963) and J. T. Cannon and T. F. Jordan, *ibid.* **5**, 299 (1964); also H. Ekstein, Université d'Aix-Marseille, 1964 (unpublished). These "no interaction" theorems apply to Dirac's Hamiltonian framework, given in Rev. Mod. Phys. **21**, 392 (1949). In this framework, Lorentz transformations are canonical transformations. The "no interaction" theorem can be circumvented if one is willing to give up the existence of world lines, as was done by L. H. Thomas, Rev. Mod. Phys. **17**, 182 (1945); B. Bakamjian and L. H. Thomas, Phys. Rev. **92**, 1300 (1953); L. L. Foldy, *ibid.* **122**, 275 (1961).