case for each and every k, the kth component u_k vanishes in the sense just described. Our requirement is then that if u is such a state vector, then Su shall likewise be such a state vector.

Mathematically this may be expressed as follows: Let M_b and M_f denote the manifolds in H_b and H_f consisting of those wave functions vanishing on the given backward ray. The S operator must then be such as to leave invariant the submanifold of K of the form $(0\oplus M_b\oplus M_b\otimes' M_b\oplus\cdots)\otimes(0\oplus M_f\oplus M_f\otimes_1 M_f\oplus\cdots)$, where $\otimes_1$ denotes the antisymmetric tensor product. Equivalently, if P denotes the operation of projection of K onto this subspace, $SP=PSP$. Each backward light ray gives a projection P for which this equation must hold, but for a covariant interaction it suffices to take any one such ray.

In the present paper we are concerned not with the entire S operator, but only with its restriction to the subspace of K in which exactly one boson and one fermion are present, i.e., with the operator A_1SA_1, where A_1 denotes the projection of K onto the subspace $H_b\otimes H_f$, or with the corresponding operator in the case of an arbitrary scatterer. The methods, however, apply in principle to the complete S operator.

PHYSICAL REVIEW VOLUME 109, NUMBER 6 MARCH 15, 1958

Formulation of Quantum Mechanics Based on the Quasi-Probability Distribution Induced on Phase Space*

GEORGE A. BAKER, JR.†
University of California, Berkeley, California
(Received March 25, 1957)

We postulate a formulation of quantum mechanics which is based solely on a quasi-probability function on the classical phase space. We then show that this formulation is equivalent to the standard formulation, and that the quasi-probability function is exactly analogous to the density matrix of Dirac and von Neumann. We investigate the theory of measurement in this formulation and derive the following remarkable results. As is well known, the correspondence between classical functions of both the position and conjugate momentum and quantum mechanical operators is ambiguous because of noncommutativity. We show that the solution of this correspondence problem is completely equivalent to the solution of the eigenvalue problem. This result enables us to give a constructive method to compute eigenvalues and eigenfunctions.

I. INTRODUCTION AND SUMMARY

IT is well known that, as a general rule, for macroscopic phenomena, classical mechanics furnishes quite a good description of nature. If we have a mechanical system, it is described classically by a Hamiltonian function $H(q_k,p_k,t)$. Classical mechanics asserts that if we measure the system, we will find it with unit probability at a point, $(q_k(t),p_k(t))$, in phase space which moves in accordance with Hamilton's canonical equations,

$$\dot{q}_k=\{q_k,H\},\quad \dot{p}_k=\{p_k,H\},$$

where $\{A,B\}$ is the classical Poisson bracket.[1]

We find experimentally, however, that it is not possible to make the measurements necessary to establish the classical trajectory. The fundamental limitation is expressed by Heisenberg's uncertainty principle which states that it is impossible to ascertain the position of a system in phase space more accurately than to say that it is in a volume of the order of h^n, where n is the number of degrees of freedom and h is Planck's constant. The uncertainty principle shows us the need for a different representation than the classical, moving phase-point.

For the case of quantum-mechanical systems in which all observables may be expressed as functions of the coordinates and their canonical momenta (q_k,p_k), we may represent the system by a quasi-probability (not everywhere necessarily non-negative) distribution in phase space, instead of the more usual Heisenberg or Schrödinger representations. We shall see that the impossibility of simultaneously measuring complementary quantities (such as q and p) will be closely related to the occurrence of "negative probability." We show that the quasi-probability distributional representation is equivalent to the standard formulation. In our formulation, we replace the classical condition of a point representation with a corresponding quantum condition, and with the aid of the correspondence principle, are able to derive the dynamical law.

By introducing the appropriate orthonormal set, we are able to show that the quasi-probability function which we use is isomorphic to the statistical operator of von Neumann.[2]

* Submitted in partial fulfillment of the requirement for the Ph.D. degree, University of California, Berkeley, California.

† Now at Los Alamos Scientific Laboratory, Los Alamos, New Mexico.

[1] H. Goldstein, *Classical Mechanics* (Addison-Wesley Publishing Company, Inc., Cambridge, 1953).

[2] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, translated by R. T. Beyer (Princeton University Press, Princeton, 1955).

As a result of our study of the quantum theory of measurement, we are able to develop a method for constructing the solution to any quantum mechanical eigenfunction problem. The problem of the correspondence between phase space functions and the powers of a given physical quantity is shown to be equivalent to the solution of the eigenfunction problem, and we give an explicit rule to determine this correspondence.

II. QUASI-PROBABILITY DISTRIBUTIONAL FORMULATION OF QUANTUM MECHANICS

This formulation of quantum mechanics is based on the following postulate:

Postulate Q.—There exists a quasi-probability distribution function $f(q_k,p_k,t)$ of the conjugate coordinates (q_k,p_k) and the time, t, satisfying the conditions

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f dq_1\cdots dq_n dp_1\cdots dp_n=1, \qquad \text{(normalization)} \quad (1)$$

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} |f|^2 dq_1\cdots dq_n dp_1\cdots dp_n \text{ exists}, \qquad \text{(boundedness)} \quad (2)$$

$$f=h^n(f,f), \quad \text{(quantum)}, \qquad (3)$$

$$\frac{\partial f}{\partial t}=\frac{-1}{\hbar}[f,H], \quad \text{(dynamical)}, \qquad (4)$$

where $H(q_k,p_k,t)$ is the classical Hamiltonian function, which completely defines the quantum mechanical state of the system.

We have used the definitions

$$(A,B)=\left(\frac{2}{h}\right)^{2n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\cos\left\{\frac{2}{\hbar}\sum_{j=1}^{n}\det\begin{vmatrix}1 & q_j & p_j\\ 1 & \tau_j & \sigma_j\\ 1 & \xi_j & \eta_j\end{vmatrix}\right\}$$
$$\times A(\tau_k,\sigma_k)B(\xi_k,\eta_k)d\xi_1\cdots d\xi_n d\eta_1\cdots$$
$$\times d\eta_n d\tau_1\cdots d\tau_n d\sigma_1\cdots d\sigma_n,$$

and

$$[A,B]=2\left(\frac{2}{h}\right)^{2n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\sin\left\{\frac{2}{\hbar}\sum_{j=1}^{n}\det\begin{vmatrix}1 & q_j & p_j\\ 1 & \tau_j & \sigma_j\\ 1 & \xi_j & \eta_j\end{vmatrix}\right\}$$
$$\times A(\tau_k,\sigma_k)B(\xi_k,\eta_k)d\xi_1\cdots d\xi_n d\eta_1\cdots$$
$$\times d\eta_n d\tau_1\cdots d\tau_n d\sigma_1\cdots d\sigma_n.$$

We remark that one can show for properly restricted A and B, by applying a suitable form of Riemann's theorem on trigonometric integrals, and an integration by parts in the second case, that, in the limit as h goes to zero,

$$(A,B)\rightarrow A(q_k,p_k)B(q_k,p_k),$$

$$\frac{1}{\hbar}[A,B]\rightarrow\{A,B\}=\sum_{j=1}^{n}\left(\frac{\partial A}{\partial q_j}\frac{\partial B}{\partial p_j}-\frac{\partial A}{\partial p_j}\frac{\partial B}{\partial q_j}\right).$$

The relation for the sine bracket converts condition (4) into Liouville's theorem and hence in the classical limit f changes in time like a classical statistical mechanical distribution would. The relation for the cosine bracket, together with condition (3), implies that f tends to a distribution on a set of measure zero in the classical limit. Thus, in the classical limit, this formulation reduces to a phase point executing a classical trajectory.

It is now our purpose to show how the quasi-probability distributional formulation is related to the density matrix formulation of von Neumann and Dirac. To do so, we first show that the distribution function may be written in the form given by Wigner.[3] We then show, by introducing an appropriate orthonormal set, the one-to-one correspondence between the quasi-probability distributional formulation and the density matrix formulation.

It may be useful in following the derivations given herein to think of the quasi-probability distribution function as a particular representation of the more familiar density matrix, and the sine and cosine brackets as the commutator and one-half the anticommutator brackets, respectively. We show that there is an isomorphism between the density matrix formulation and the quasi-probability distributional formulation.

We now show that we may write

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left\{-\frac{2i}{\hbar}\sum_{k=1}^{n}s_k p_k\right\}f(q_k,p_k)dp_1\cdots dp_n$$
$$=g^*(q_k+s_k)g(q_k-s_k), \qquad (3')$$

where g depends on the state of the system. It follows from the definition that $[A,B]=-[B,A]$. Therefore, $[f,f]=0$. So, by condition (3) of postulate Q,

$$f=h^n\left((f,f)+\frac{i}{2}[f,f]\right),$$

or

$$f=\frac{2^{2n}}{h^n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left\{\frac{2i}{\hbar}\sum_{j=1}^{n}[p_j(\tau_j-\xi_j)-\sigma_j(q_j-\xi_j)\right.$$
$$\left.-\eta_j(\tau_j-q_j)]\right\}f(\tau_k,\sigma_k)f(\xi_k,\eta_k)d\xi_1\cdots$$
$$\times d\xi_n d\eta_1\cdots d\eta_n d\tau_1\cdots d\tau_n d\sigma_1\cdots d\sigma_n.$$

[3] E. Wigner, Phys. Rev. 40, 749 (1932).

Let us make a change of variables of integration:

$$\tau_j-\xi_j=y_j,\quad \tau_j+\xi_j=w_j+q_j,\quad \text{Jacobian}=(\tfrac{1}{2})^n.$$

Then

$$f=\left(\frac{2}{n}\right)^n\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left\{\frac{2i}{\hbar}\sum_{j=1}^{n}[p_jy_j-\tfrac{1}{2}\sigma_j(q_j+y_j-w_j)+\tfrac{1}{2}\eta_j(q_j-y_j-w_j)]\right\}f((w_k+y_k+q_k)/2,\sigma_k)$$
$$\times f((w_k+q_k-y_k)/2,\eta_k)dw_1\cdots dw_ndy_1\cdots\times dy_nd\eta_1\cdots d\eta_nd\sigma_1\cdots d\sigma_n.$$

If we take the Fourier transform of the above relation with respect to (p_k),then, defining the auxiliary function

$$G(q_k+s_k,\,q_k-s_k)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left\{-\frac{2i}{\hbar}\sum_{j=1}^{n}p_js_j\right\}\times f(q_k,p_k)dp_1\cdots dp_n,$$

we obtain, by Fourier's integral theorem,[4]

$$G(q_k+s_k,\,q_k-s_k)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}G(q_k+s_k,\,w_k)G(w_k,\,q_k-s_k)dw_1\cdots dw_n.$$

If we think of $G(q_k+s_k,\,w_k)$ as the kernel of a homogeneous, linear integral equation, we see that it has at least one solution, i.e., $G(w_k,\,q_k-s_k)$ and its eigenvalue is unity. By a slight modification of the arguments of Courant and Hilbert,[5] we know

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}|G(q_k,w_k)|^2dq_1\cdots dq_ndw_1\cdots dw_n\geq\sum_{i=1}^{\infty}\frac{1}{|\lambda_i|^2},$$

where the λ_i are the eigenvalues. But, by the relation we derived above, the integral becomes

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}G(q_k,q_k)dq_1\cdots dq_n,$$

as $G(x,y)=G^*(y,x)$, which is, by definition, equal to

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(q_k,p_k)dq_1\cdots dq_ndp_1\cdots dp_n=1,$$

by the normalization of f. Hence

$$1\geq 1+\sum_{i=2}^{\infty}\frac{1}{|\lambda_i|^2}.$$

[4] E. C. Titchmarch, *Introduction to the Theory of Fourier Integrals* (Clarendon Press, Oxford, 1937), Chap. III.

[5] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953), Chap. III, Sec. 4.

Therefore, there is only one eigenvalue, 1, and by the above-mentioned arguments of Courant and Hilbert, we see that $G(x,y)$ is a degenerate kernel, and so must be of the form

$$G(x,y)=g^*(x)g(y),$$

which is (3′).

If we take the inverse Fourier transform of (3′) on (s_k) and identify g with the wave function, ψ, we obtain the Wigner form for f. Hence

$$f(q_k,p_k)=\left(\frac{2}{h}\right)^n\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left\{\frac{2i}{\hbar}\sum_{k=1}^{n}p_ky_k\right\}\times\psi^*(q_k+y_k)\psi(q_k-y_k)dy_1\cdots dy_n.\quad (3'')$$

It is this form which Wigner[3] chose "from all possible expressions, because it seems to be the simplest," although he knew only that it gave the correct marginal distributions. Moyal[6] has shown that it also gives the correct joint distribution if we make the "Weyl correspondence"[7] (see also, Sec. III below) between operators and phase-space functions. Moyal investigates the quasi-probability distribution function from the point of view of modern statistical theory and the theory of general stochastic processes. Groenwold[8] and Takabayasi[9] have also investigated this form and some equivalent forms of the quasi-probability distribution function.

We remark that, if we integrate first on p and then on q that the normalization of f insures that ψ must be square-integrable, and hence belong to a Hilbert space.

III. RELATION BETWEEN THE QUASI-PROBABILITY DISTRIBUTION AND THE STATISTICAL OPERATOR OF VON NEUMANN[2]

Following von Neumann, we introduce an ensemble of systems each of which is in a "pure state," and each state has a certain frequency of occurrence in the ensemble. The quasi-probability distribution function for the ensemble need not satisfy condition (3) of postulate Q, but rather it is a sum of functions which do. Hence f for the ensemble will be

$$f=\textstyle\sum_\rho w_\rho f_\rho(q_k,p_k).$$

Let us introduce a complete orthonormal set of wave functions $\{\psi_j(q_k)\}$. From the form (3′) of f, we know that to each f_ρ, there corresponds a ψ_ρ which we may expand as

$$\psi_\rho=\textstyle\sum_j a_{\rho j}\psi_j.$$

It then follows at once that

$$f=\sum_{\rho,i,j}w_\rho a_{\rho i}^*a_{\rho j}f_{ij},$$

[6] J. E. Moyal, Proc. Cambridge Phil. Soc. **45**, 99 (1949).

[7] H. Weyl, *The Theory of Groups and Quantum Mechanics*, translated from the German by H. P. Robertson (Dover Publications, New York, 1931), p. 274.

[8] H. J. Groenwold, Physica **12**, 405 (1946).

[9] T. Takabayasi, Progr. Theoret. Phys. Japan **11**, 341 (1954).

where we define

$$f_{ij}(q_k,p_k)=\left(\frac{2}{h}\right)^n\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left\{\frac{2i}{\hbar}\sum_{k=1}^{n}y_kp_k\right\}\times\psi_i^*(q_k+y_k)\psi_j(q_k-y_k)dy_1\cdots dy_n.$$

The f_{ij} have certain orthogonality properties which we shall now note. These properties have been, in essence, derived by Moyal[6] for one degree of freedom, but their proof for n degrees is the same. They are as follows (variables of integration suppressed):

(i) $\int f_{ij}^*f_{km}=0$ if and only if

$$\int\psi_i^*\psi_k=0,\quad\text{or}\quad\int\psi_j^*\psi_m=0.$$

(ii) $\int|f_{ij}|^2=h^{-n}$.

(iii) The ψ_i are an orthonormal set if and only if the $h^{n/2}f_{ii}$ are

(iv) $\int f_{ij}=\delta_{ij}$, if the set $\{\psi_i\}$ is orthonormal.

(v) If and only if the set $\{\psi_i\}$ is a complete orthonormal set,

$$\sum_{i,j}f_{ij}(q_k,p_k)f_{ij}^*(q_k',p_k')=h^{-n}\prod_{k=1}^{n}\delta(q_k-q_k')\delta(p_k-p_k').$$

(vi) If $\{\psi_i\}$ is a complete orthonormal set, then $h^{n/2}f_{ij}$ is a complete orthonormal set in the Hilbert space of phase-space functions. This is to say that, not only do the f_{ij} form a basis for the quasi-probability distribution functions, but they also span the entire Hilbert space (L_2) of functions on phase space.

If we now compute the matrix

$$\left[h^n\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f_{ij}^*fdq_1\cdots dq_ndq_1\cdots dp_n\right],$$

we obtain

$$[\textstyle\sum_p w_pa_{pi}^*,a_{pj}],$$

which is just the matrix for von Neumann's statistical operator $[U_{ij}]$. The matrix corresponding to a quantity $R(q_k,p_k)$ is seen to be

$$[R_{jm}]=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}R(q_k,p_k)\times f_{j,m}^*(q_k,p_k)dq_1\cdots dq_ndp_1\cdots dp_n,$$

as the expected value of $R(q_k,p_k)$ is given correctly by von Neumann's rule:

$$\langle R\rangle=\mathrm{Trace}(RU)$$

for all U. For

$$\begin{aligned}\mathrm{Tr}(RU)&=\sum_{j,m}R_{jm}U_{mj}\\&=\sum_{j,m}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}R(q_k,p_k)f_{j,m}^*(q_k,p_k)\\&\qquad\times[\sum_p w_pa_{pm}^*a_{pj}]dq_1\cdots dq_ndp_1\cdots dp_n,\end{aligned}$$

and as $f_{jm}^*=f_{mj}$, this becomes

$$\begin{aligned}\mathrm{Tr}(RU)&=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}R(q_k,p_k)\\&\qquad\times[\sum_{p,j,m}w_pa_{pm}^*a_{pj}f_{mj}(q_k,p_k)]\\&\qquad\times dq_1\cdots dq_ndp_1\cdots dp_n\\&=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}R(q_k,p_k)f(q_k,p_k)dq_1\cdots\\&\qquad\times dq_ndp_1\cdots dp_n=\langle R(q_k,p_k)\rangle.\end{aligned}$$

These results indicate that the quasi-probability distribution is directly analogous to von Neumann's statistical operator. Where he uses infinite matrices as the basis of his theory, we use functions of the real variables (q_k,p_k). It is worth noting that, using the above method to define a matrix for a function, the matrix for the cosine bracket, (A,B), is one-half the anti-commutator of the matrix for A and the matrix for B. Also the matrix for the sine bracket, $[A,B]$, is simply the commutator divided by i of the matrix for A and the matrix for B. These results serve to establish an isomorphism between the space of functions of real variables and the space of infinite matrices. They may be verified by a straightforward formal calculation, which starts from the following rule for the result of $R(q_k,p_k)$ acting on ψ. This rule follows at once from our definition of the matrix elements R_{jm}. It is

$$\mathcal{R}\psi(q_1,\cdots q_n)=h^{-n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left\{\frac{i}{\hbar}\sum_{k=1}^{n}\eta_k(q_k-\xi_k)\right\}\times R((q_k+\xi_k)/2,\,\eta_k)\psi(\xi_k)d\xi_1\cdots d\xi_nd\eta_1\cdots d\eta_n.$$

We note that this rule may also be derived from the correspondence suggested by Weyl[7] by some fairly straightforward manipulations involving the use of Fourier's integral theorem. Let $\mathcal{P}$ be the operator corresponding to p and $\mathcal{Q}$ be the operator corresponding to q. Let them satisfy the commutation relation

$$\mathcal{P}\mathcal{Q}-\mathcal{Q}\mathcal{P}=(h/i)\mathcal{E},$$

where $\mathcal{E}$ is the identity operator. If

$$R(\xi,p)=\frac{1}{h}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left\{\frac{i}{\hbar}(\sigma\xi+\tau p)\right\}\zeta(\sigma,\tau)d\sigma d\tau,$$

then, according to Weyl, the correct operator is obtained by replacing ξ by $\mathcal{Q}$ and p by $\mathcal{P}$. In this derivation, use is made of an identity of Kermack and McCrea[10]:

$$\exp\left\{\frac{i}{\hbar}(\sigma\mathcal{P}+\tau\mathcal{Q})\right\}=\exp\left(\frac{i\sigma\tau}{2\hbar}\right)\exp\left(\frac{i\tau\mathcal{Q}}{\hbar}\right)\exp\left(\frac{i\sigma\mathcal{P}}{\hbar}\right).$$

Our quantum condition, $(f,f)=h^n f$, becomes then, in matrix language,

$$UU=U,$$

which is just von Neumann's characterization of a "pure state." The physical interpretation in the two cases is similar. In matrix language, it characterizes a projection operator onto some state, while our condition may be thought of as characterizing sort of a smeared-out projection operator for a region of phase space. It represents a modification of the classical delta function which projects onto a phase-point.

IV. QUANTUM DYNAMICS AND THE CORRESPONDENCE PRINCIPLE

We show in this section that the dynamical equation of quantum mechanics can be derived from the quantum condition, with the aid of the Bohr correspondence principle. For this demonstration, it is convenient to define a dot product as

$$A\cdot B=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}A(q_k,p_k)B(q_k,p_k)\times dq_1\cdots dq_n dp_1\cdots dp_n.$$

It is easy to verify from the definitions that

$$[A,B]\cdot C=A\cdot[B,C],$$

and to verify, by formal integration by parts and Fourier's integral theorem, that

$$[A,B]=\hbar\{A,B\},$$

if A is a polynomial, at most quadratic, where $\{A,B\}$ is the classical Poisson bracket.

The large-scale experimental validity of classical mechanics tells us that quantum theory must, in some sense, correspond closely to classical mechanics. We have altered the classical concept of a moving point in phase space to that of a quasi-probability distribution which changes in time. This distribution (see Sec. II) is imagined to be concentrated about the classical point, so that a crude measurement will be unable to differentiate between the two theories. To insure this correspondence, we shall use the statement which actually seems to be given by experiments—on the average, Hamilton's canonical equations hold. It can be shown, say by using the Wigner form (3″) of f and some of the properties given in the next section, and making an infinitesimal change ψ, that the most general infinitesimal change δf which preserves the normalization and quantum conditions is given by

$$\delta f=[f,\delta g],$$

where δg is arbitrary. Since by "the average of $\dot{q}_k$" we mean the time rate of change of the expected value of q_k, we have

$$\text{Average }(\dot{q}_k)=\frac{d}{dt}(q_k\cdot f)=q_k\cdot\frac{\partial f}{\partial t}.$$

Also

$$\delta f\equiv\delta t(\partial f/\partial t).$$

We must have, by the correspondence principle,

$$\begin{aligned}\delta t\left(q_k\cdot\frac{\partial f}{\partial t}\right)&=q_k\cdot[f,\delta g]\\&\equiv-[q_k,\delta g]\cdot f\\&\equiv-\hbar\{q_k,\delta g\}\cdot f\\&=\delta t\{q_k,H\}\cdot f.\end{aligned}$$

Thus we see, as the above equation must hold for all q_k and p_k, and for any possible f, we must (outside an arbitrary, additive constant, V_0) choose for δg

$$\delta g=-H\delta t/\hbar.$$

Thus we obtain the dynamical equation

$$\frac{\partial f}{\partial t}=-\frac{1}{\hbar}[f,H],$$

which is given by condition (4) of postulate Q. It should be noted that this equation is the direct analog of Liouville's theorem of classical statistical mechanics.[1]

We see, therefore, that in this formulation, the change in the formal structure from classical to quantum mechanics consists in replacing the equation $f=(0^+)f^2$ by $f=h^n(f,f)$. (See Sec. II for limiting behavior of the cosine bracket as $h\to 0$.) The quasi-probability distributional formulation has the advantage that it does not depend on the two superfluous constants, the arbitrary phase factor and the additive constant in the classical potential energy which appears in the standard Schrödinger formulation. This lack of dependence on arbitrary, unobservable constants is not only an advantage, *per se*, but should be a great convenience in the treatment of the asymptotic behavior in scattering problems. Furthermore, our formulation provides a sort of intuitive picture of what the system is doing in phase space.

[10] W. O. Kermack and W. H. McCrea, Proc. Edinburg Math. Soc. 2, 224 (1931).

V. FORMAL PROPERTIES

(I) One property of the quasi-probability distribution which is easy to demonstrate is that it is uniformly bounded (see also, Takabayasi[9]). In terms of the wave function $\psi(q_k)$, we have

$$f(q_k,p_k)=(2/h)^n\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left[\frac{2i}{\hbar}\sum_{k=1}^{n}y_kp_k\right]\times\psi^*(q_k+y_k)\psi(q_k-y_k)dy_1\cdots dy_n.$$

By the Schwartz inequality,[4] we have

$$|f(q_k,p_k)|^2\leq(2/h)^{2n}\left\{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\left|\exp\left[\frac{2i}{\hbar}\sum_{k=1}^{n}y_kp_k\right]\times\psi^*(q_k+y_k)\right|^2dy_1\cdots dy_n\right\}\times\left\{\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}|\psi(q_k-y_k)|^2dy_1\cdots dy_n\right\},$$

which, as $\int\psi\psi^*=1$, implies

$$|f(q_k,p_k)|\leq(2/h)^n.$$

(II) A second property is the following one. Let us define

$$f_{\mathrm{I}}=\frac{1}{N}\sum_{k=1}^{N}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left[\frac{2i}{\hbar}\sum_{j=1}^{n}p_jy_j\right]\times\phi_k^*(q_j+y_j)\phi_k(q_j-y_j)dy_1\cdots dy_n,$$

where the ϕ_k's are orthonormal.

Let $[a_{ik}]$ be a unitary transformation and let us also define

$$\chi_i(q_j)=\sum_{k=1}^{n}a_{ik}\phi_k(q_j),$$

and

$$f_{\mathrm{II}}=\frac{1}{N}\sum_{i=1}^{n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left[\frac{2i}{\hbar}\sum_{j=1}^{n}p_jy_j\right]\times\chi_i^*(q_j+y_j)\chi_i(q_j-y_j)dy_1\cdots dy_n,$$

then $f_{\mathrm{I}}=f_{\mathrm{II}}$. This means that if f represents an ensemble composed of equal numbers of systems in N orthogonal states, then we get the same f no matter in which way we make up the orthogonal states. To see this, we expand f_{II} as

$$f_{\mathrm{II}}=\frac{1}{N}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\exp\left[\frac{2i}{\hbar}\sum_{j=1}^{n}p_jy_j\right]\times\sum_{k,m}^{N}\sum_{i=1}^{N}a_{ik}^*\phi_k^*(q_j+y_j)a_{im}\phi_m(q_j-y_j)dy_1\cdots dy_n.$$

Now, as $[a_{ik}]$ is unitary,

$$\sum_{i=1}^{N}a_{ik}^*a_{im}=\delta_{km}.$$

Thus, by summing over m, it reduces to the definition of f_{I}.

(III) The third group of properties listed below follow by straightforward, but somewhat tedious, formal calculation directly from the definitions. They are, however, obvious from the analogy to the density matrix formulation with the dot product playing the role of the trace.

$$[A,B]\cdot f=[f,A]\cdot B=[B,f]\cdot A=A\cdot[B,f],$$
$$[A,B]=-[B,A],$$
$$(A,B)=(B,A),$$
$$(A,B)\cdot f=(f,A)\cdot B,\ \text{etc.},$$
$$[A,[B,C]]+[C,[A,B]]+[B,[C,A]]=0,$$
$$[A,(B,C)]=([A,B],C)+([A,C],B).$$

If f_{ii} and f_{jj} are orthogonal to each other, then

$$(f_{ii},f_{jj})=0,\quad [f_{ii},f_{jj}]=0,\quad\text{and, of course,}\quad [f_{ii},f_{ii}]=0.$$

VI. MEASUREMENT

We are now in a position to discuss the effect of measurement on a quantum-mechanical system. In the standard Schrödinger representation, the measurement of a quantity, $R(q_k,p_k)$, leaves the system in a state described by a ψ which satisfies the eigenvalue equation,

$$\mathcal{R}\psi=\lambda\psi,$$

where $\mathcal{R}$ is the operator corresponding to $R(q_k,p_k)$. We know that this equation is equivalent[11] to the extremal condition

$$\delta(\langle R\rangle)=0,$$

or

$$\delta(R\cdot f)=R\cdot\delta f=R\cdot[f,\delta g]=[R,f]\cdot\delta g=0,$$

where δg is an arbitrary variation. Because δg is arbitrary, we must have

$$[R,f]=0.$$

This condition generates a sequence of quasi-probability distribution functions, $f_{\lambda\lambda}$, indexed by λ, where it is understood that several distinct $f_{\lambda\lambda}$ may be given the same name by this naming process, and

$$\lambda=R\cdot f_{\lambda\lambda}.$$

We shall say that the $\{f_{\lambda\lambda}\}$ form a "complete" set if

$$1=h^n\textstyle\sum_\lambda f_{\lambda\lambda}\cdot f$$

(conservation of probability) for all quasi-probability distribution functions f.

[11] H. Jeffreys and B. S. Jeffreys, *Methods of Mathematical Physics* (Cambridge University Press, New York, 1950), Sec. 10.14.

The case of the degenerate $f_{\lambda\lambda}$ (more than one f with the same value of λ) can be clarified as follows. We know from the standard quantum theory that the ψ_λ corresponding to different λ are orthogonal and hence (Sec. III) the $f_{\lambda\lambda}$ are. Further the ψ_λ corresponding to the same λ can be made orthogonal by the Schmidt process. By property II of Sec. V, it does not matter in which way it is done, since $\sum_\lambda f_{\lambda\lambda}$ involves equal weights to each $f_{\lambda\lambda}$. Thus we must understand by the above "completeness" condition that all the $f_{\lambda\lambda}$ are to be orthogonal to each other, pairwise. We may now formulate the following measurement postulate.

Postulate M.—If we have an ensemble represented by a normalized, weighted sum $\mathfrak{F}$ of quasi-probability distribution functions, then the measurement of a dynamical quantity, $R(q_k,p_k)$, decomposes the ensemble into a set of subensembles indexed by the measured value of $R(q_k,p_k)$. Each subensemble is represented by a quasi-probability distribution function $f_{\lambda\lambda}$, which satisfies the condition $[R,f_{\lambda\lambda}]=0$, and in each subensemble $R(q_k,p_k)$ takes on precisely its measured value, λ. In order for a measurement to be possible, all the conditions of this postulate must be enforceable for all possible $\mathfrak{F}$.

Now by the results of Sec. III, we know that we can expand any quasi-probability distribution function, and hence any weighted, normalized sum of them in terms of a complete orthonormal set $(h^{n/2}f_{ij})$. Now if we assume $R(q_k,p_k)$ measurable, the condition $[R,f_{\lambda\lambda}]=0$ must form a "complete" set, or we would not be able to decompose the whole ensemble. Each $f_{\lambda\lambda}$ implies a corresponding ψ_λ, and hence we can construct a complete orthonormal system, $(h^{n/2}f_{\lambda\nu})$, by the method of Sec. III. We note that this orthonormal system has the property that the $f_{\lambda\lambda}$ are quasi-probability distribution functions, while the $f_{\lambda\nu}$, $\lambda\neq\nu$ are not. Let us expand $\mathfrak{F}$ in terms of it. By Sec. III, it becomes

$$\mathfrak{F}=\sum_{\rho,\lambda,\nu} w_\rho a_{\rho\lambda}{}^* a_{\rho\nu} f_{\lambda\nu}(q_k,p_k).$$

If we make a measurement, by postulate M, the $f_{\lambda\nu}$, $\lambda\neq\nu$, are destroyed. (This results in no loss of normalization as $\int f_{\lambda\nu}=\delta_{\lambda\nu}$ by Sec. III, iv.) Hence a measurement of $R(q_k,p_k)$ transforms $\mathfrak{F}$ into

$$\mathfrak{F}'=\sum_{\rho,\lambda} w_\rho a_{\rho\lambda}{}^* a_{\rho\lambda} f_{\lambda\lambda}(q_k,p_k).$$

We may now compute the distribution of measured values of $R(q_k,p_k)$ by means of the orthogonality relations as

$$F(R)-F(0)=\sum_{0\leq\lambda\leq R}{}' f_{\lambda\lambda}\cdot\mathfrak{F},$$

where $F(R)$ is the cumulative distribution of R. By $\sum'$, we mean that if there is a contribution at either end point, we take only half of it. This is done to adapt the function F to Fourier analysis.

However, we can proceed otherwise to obtain the cumulative distribution (and it is a true cumulative distribution for $\sum w_\rho |a_{\rho\lambda}|^2\geq 0$) and obtain an important result thereby. We first obtain the standard statistical characteristic function

$$C(S)=\sum_{\nu=0}^{\infty}(-iS/\hbar)^\nu\left(\frac{\mu_\nu(\mathfrak{F})}{\nu!}\right),$$

where μ_ν is the νth moment of R, given $\mathfrak{F}$, computed from the above cumulative distribution. It can be shown that there exist functions $R^{(\nu)}(q_k,p_k)$ (if $|\mu_\nu|<\infty$) such that

$$\mu_\nu(\mathfrak{F})=R^{(\nu)}\cdot\mathfrak{F}$$

for all $\mathfrak{F}$. According to Kendall,[12] the cumulative distribution is then

$$F(R)-F(0)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{[1-\exp(iRS/\hbar)]}{iS}C(S)dS.$$

Substituting for $C(S)$ and equating these two expressions, we see, when the appropriate interchange of limit processes is permissible, that we must have, as $\mathfrak{F}$ is arbitrary by the relations of Sec. III (vi),

$$\sum_{0\leq\lambda\leq R}{}' f_{\lambda\lambda}(q_k,p_k)=\frac{1}{2\pi h^n}\int_{-\infty}^{+\infty}\frac{[1-\exp(iRS/\hbar)]}{iS}\times\left(\sum_{\nu=0}^{\infty}\frac{(-iS)^\nu}{\hbar}\frac{R^{(\nu)}(q_k,p_k)}{\nu!}\right)dS.$$

Thus we see that the $f_{\lambda\lambda}$ must be constructed from the $R^{(\nu)}(q_k,p_k)$. Conversely, we must have

$$R^{(\nu)}(q_k,p_k)=h^n\sum_{\text{all }\lambda}\lambda^\nu f_{\lambda\lambda}(q_k,p_k)$$

$$=\int_{\lambda=-\infty}^{+\infty}\lambda^\nu dF_\lambda(q_k,p_k),\quad\text{(Stieltjes integral)},$$

where we define

$$F_\lambda(q_k,p_k)=h^n\sum_{0\leq\mu\leq\lambda}{}' f_{\mu\mu}(q_k,p_k).$$

It can be shown by use of the relations of Sec. V, property III, that the $R^{(\nu)}(q_k,p_k)$ satisfy the equation

$$R^{(\nu)}\cdot\mathfrak{F}=(R,R^{(\nu-1)})\cdot\mathfrak{F}$$

for all $\mathfrak{F}$, as we would expect from the analog pointed out in Sec. III. As Moyal[6] has shown, $R^{(0)}=1$, so that we may use the above relation to construct successively the $R^{(\nu)}$.

This result gives an explicit method of solving the eigenfunction problem for the measurement of R. We use the above equation to compute the $R^{(\nu)}$ and then

[12] M. G. Kendall, *The Advanced Theory of Statistics* (Charles Griffen and Company, Ltd., London, 1947), Chap. 4.

use them to compute the $f_{\lambda\lambda}$. We see that the problem of which quantity corresponds to the νth power of an observed quantity is equivalent to the eigenfunction problem.

VII. SIMULTANEOUS MEASUREMENT

Two quantities R and S are clearly simultaneously measurable if and only if postulate M can be imposed for both at once. This means that $\mathfrak{F}$ must be decomposable into a set of subensembles represented by quasi-probability distribution functions $f_{\rho\rho,\,\sigma\sigma}$ indexed by

$$\rho=R\cdot f_{\rho\rho,\,\sigma\sigma},\quad \sigma=S\cdot f_{\rho\rho,\,\sigma\sigma},$$

where $[R,f_{\rho\rho,\,\sigma\sigma}]=[S,f_{\rho\rho,\,\sigma\sigma}]=0$, and R and S take on the precise values ρ and σ, respectively. We must also have

$$1=h^n\sum_{\rho,\sigma} f_{\rho\rho,\,\sigma\sigma}\cdot\mathfrak{F}$$

for all $\mathfrak{F}$. We now have, as before, for the cumulative joint distribution

$$F(R,S)-F(0,0)=h^n \sum_{\substack{\rho,\,\sigma\\ 0\leq\rho\leq R\\ 0\leq\sigma\leq S}}{}' f_{\rho\rho,\,\sigma\sigma}\cdot\mathfrak{F}.$$

An argument analogous to that given above (Sec. VI) shows the quantity $(R^{(\nu)}S^{(\mu)})$ corresponding to the (ν,μ)th moment of the above distribution is

$$(R^{(\nu)}S^{(\mu)})=h^n\sum_{\text{all }(\rho,\sigma)}\rho^\nu\sigma^\mu f_{\rho\rho,\,\sigma\sigma}(q_k,p_k).$$

We compute symbolically the cosine bracket

$$(R^{(\nu)},S^{(\mu)})=h^{2n}\Big(\sum_{\text{all }(\rho,\sigma)}\rho^\nu f_{\rho\rho,\,\sigma\sigma},\sum_{\text{all }(\rho,\sigma)}\sigma^\mu f_{\rho\rho,\,\sigma\sigma}\Big)$$
$$=(R^{(\nu)}S^{(\mu)}),$$

where use has been made of the relations of Sec. V. By virtue of their nature as weighted sums of the same quasi-probability distribution functions, we see that

$$[R,S]\cdot\mathfrak{F}=0$$

for all $\mathfrak{F}$. That is to say, if two quantities are simultaneously measurable, their operators commute, a well-known result of the standard formulations.

Let us define an Nth order cosine bracket as

$$(T_1,T_2,\cdots,T_N)=\frac{1}{N!}\sum_{\text{all permutations}}\times\{T_1,[T_2,(\cdots,T_N)\cdots]\}.$$

This is totally symmetric in the T_k. We see at once, from the work of this and the previous section, that the joint distribution of N simultaneously measurable quantities $T_1,\cdots,T_N$ must be

$$F(T_1,\cdots,T_N)-F(0,\cdots,0)=[F_{T1}(q_k,p_k),\cdots,F_{TN}(q_k,p_k)]\cdot\mathfrak{F},$$

where

$$F_{Tj}(q_k,p_k)=\sum_{0\leq\tau_j\leq T_j}{}' f_{\tau j}(q_k,p_k),$$

with

$$[T_j,f_{\tau j}]=0,\quad \tau_j=T_j\cdot f_{\tau j},$$

and the condition $[T_j,T_k]\cdot\mathfrak{F}=0$ must hold for all j, k, and $\mathfrak{F}$. Then the expected value of any function

$$G(T_1,\cdots,T_N)$$

is given by

$$\langle G\rangle=\underbrace{\int\cdots\int}_{\substack{\text{entire range}\\ \text{of the }T_k}} G(T_1\cdots,T_N)dF(T_1,\cdots,T_N).$$

As we can form $F(T_1,\cdots,T_N)-F(0,\cdots,0)$ in an unambiguous manner according to our above definition for any (T_k), whether they are simultaneously measurable or not, we might wonder what its significance is, if any, for nonsimultaneously measurable quantities. Now for this case, von Neumann[3] (Chap. IV, Sec. 2) has shown that F cannot be a true cumulative distribution function for all possible states of the system as this would lead to dispersion-free ensembles, which are impossible. We have exhibited an F which is a true distribution, if the (T_k) are simultaneously measurable. We see that the only way it can satisfy von Neumann's theorem in the case of nonsimultaneously measurable variables is that it must imply "negative probabilities." Thus we arrive at the important physically meaningful conclusion that the F defined above is a true distribution function if and only if the (T_k) are simultaneously measurable. This is to say, when quantum mechanics predicts an impossible result like a "negative probability," then the interpretation is that there is no physically realizable experiment to measure the joint distribution. It is worth noting that in the case $T_1=q$ and $T_2=p$, that

$$\frac{d^2[F(q,p)]}{dqdp}=\frac{1}{h^{n/2}}\operatorname{Re}\{\psi(q)\phi^*(p)\exp(-ipq/\hbar)\},$$

which is not the quasi-probability distribution function. Nor could it be expected to be, because of the basic impossibility of establishing an isomorphism between a commutative and a noncommutative linear algebra. As we have seen, it is necessary, to satisfy the measurement postulate, to have the operator of the "square" of a quantity be the square of the operator; thus, if the operators do not commute, we are forced into trying to establish the above-mentioned impossible correspondence, in order to try to make a definition which correctly gives the distribution for the simultaneously measurable variables also give the quasi-probability distribution for the conjugate variables p and q.

We emphasize that these results are in accord with the fact that a dynamical quantity $R(q_k,p_k)$ which is a function of noncommuting variables is a separate and distinct entity which should be denoted by a separate symbol, $\mathbf{R}$. $R(q_k,p_k)$ has the property that $\langle\mathbf{R}\rangle=\langle R(q_k,p_k)\rangle$ for any distribution; however, we do not expect

$$\langle\mathbf{R}^2\rangle=\langle R^2(q_k,p_k)\rangle,$$

but instead

$$\langle\mathbf{R}^2\rangle=\langle R^{(2)}(q_k,p_k)\rangle.$$

In this formulation, we can correctly find the expected value of $\mathbf{R}$ by using $R(q_k,p_k)$, but it is not possible, in general, to study a function $G(R)$ in terms of $G[R(q_k,p_k)]$. As we have seen above, the solution of this correspondence problem in general is equivalent to the solution of the corresponding eigenvalue problem.

APPENDIX. EXAMPLE OF THE QUASI-PROBABILITY DISTRIBUTION: THE HARMONIC OSCILLATOR

It is a matter of straightforward calculation[8,9] to show that for the one-dimensional harmonic oscillator, the energy eigen-quasi-probability-distribution-functions are:

$$f_n(H,\theta)dHd\theta=[(-1)^n/(2\pi n!)]L_n(4H/h\nu)\times\exp(-2H/h\nu)d(2H/h\nu)d\theta,$$

where $L_n(x)$ are the Laguerre polynomials,[13] and we have made the *algebraic* change to the variables

$$\theta=\tan^{-1}[p/(2\pi m\nu q)],\quad H=(p^2/2m)+2\pi^2m\nu^2q^2.$$

The dynamical equation satisfied by f, in this example, is the same as the classical equation. It is

$$\frac{\partial f}{\partial t}=-(p/m)\frac{\partial f}{\partial q}+4\pi^2m\nu^2q\frac{\partial f}{\partial p}.$$

[13] See, for example, P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), p. 784.

It is also of interest to compute the generating function,

$$G(s)=\sum_{\nu=0}^{\infty}(-is/h)^\nu H^{(\nu)}(q,p)/(\nu!).$$

By the relations we have obtained, this is also equal to

$$G(s)=h\sum_{\text{all }\lambda}\exp(-is\lambda/\hbar)f_{\lambda\lambda}(q,p),$$

which we may compute by means of the formula for the generating function for the Laguerre polynomials.[13] Thus

$$\begin{aligned}G(s)=h\sum_{n=0}^{\infty}&\exp[-\tfrac{1}{2}is(2n+1)h\nu/\hbar](2/h\nu)\\&\times(-1)^n(n!)^{-1}L_n(4H/h\nu)\exp(-2H/h\nu)\\=&\exp[-(i/\hbar)(2H/\omega)\tan(\tfrac{1}{2}s\omega)]/\cos(\tfrac{1}{2}s\omega),\end{aligned}$$

where $\omega=2\pi\nu$.

We now obtain the various $H^{(\nu)}$ from $G(s)$ by the relation

$$H^{(\nu)}=\left(-\frac{\hbar}{i}\frac{\partial}{\partial s}\right)^\nu G(s)\bigg|_{s=0},$$

and the eigenfunctions by the relation

$$f_n=(2\pi h)^{-1}\int_{-\infty}^{+\infty}\exp(i\omega sn)[1-\exp(i\omega s)](is)^{-1}G(s)ds.$$

We obtain by differentiation

$$H^{(0)}=1,\quad H^{(1)}=H,\quad H^{(2)}=H^2-(\tfrac{1}{2}h\nu)^2,$$
$$H^{(3)}=H^3-5(\tfrac{1}{2}h\nu)^2H,\quad\text{etc.,}$$

which agree with what we obtain by the direct application of the recursion relation.