Appendix III. Quantum Mechanics as a Statistical Theory by J.E. Moyal

Originally published in the Proceedings of the Cambridge Philosophical Society, Vol. 45, 1949.

NB: page numbers refer to original publication.

QUANTUM MECHANICS AS A STATISTICAL THEORY

By J. E. MOYAL

Communicated by M. S. Bartlett

Received 12 November 1947

1. Introduction

Statistical concepts play an ambiguous role in quantum theory. The critique of acts of observation, leading to Heisenberg's 'principle of uncertainty' and to the necessity for considering dynamical parameters as statistical variates, not only for large aggregates, as in classical kinetic theory, but also for isolated atomic systems, is quite fundamental in justifying the basic principles of quantum theory; yet paradoxically, the expression of the latter in terms of operations in an abstract space of 'state' vectors is essentially independent of any statistical ideas. These are only introduced as a *post hoc* interpretation, the accepted one being that the probability of a state is equal to the square of the modulus of the vector representing it; other and less satisfactory statistical interpretations have also been suggested (cf. Dirac (1)).

One is led to wonder whether this formalism does not disguise what is an essentially statistical theory, and whether a reformulation of the principles of quantum mechanics in purely statistical terms would not be worth while in affording us a deeper insight into the meaning of the theory. From this point of view, the fundamental entities would be the statistical variates representing the dynamical parameters of each system; the operators, matrices and wave functions of quantum theory would no longer be considered as having an intrinsic meaning, but would appear rather as aids to the calculation of statistical averages and distributions. Yet there are serious difficulties in effecting such a reformulation. Classical statistical mechanics is a 'crypto-deterministic' theory, where each element of the probability distribution of the dynamical variables specifying a given system evolves with time according to deterministic laws of motion; the whole uncertainty is contained in the form of the initial distributions. A theory based on such concepts could not give a satisfactory account of such non-deterministic effects as radioactive decay or spontaneous emission (cf. Whittaker (2)). Classical statistical mechanics is, however, only a special case in the general theory of dynamical statistical (stochastic) processes. In the general case, there is the possibility of 'diffusion' of the probability 'fluid', so that the transformation with time of the probability distribution need not be deterministic in the classical sense. In this paper, we shall attempt to interpret quantum mechanics as a form of such a general statistical dynamics.

I. QUANTUM KINEMATICS

2. The existence of phase-space distributions in quantum theory

In the accepted statistical interpretation of quantum theory, the possible values of a dynamical variable s are the eigenvalues s_i of the corresponding operator (observable)

s in the Hilbert space of the state vectors. The probability of finding s_i in a state ψ is then equal to the square of the modulus $|a_i|^2$ of the projection a_i of ψ on the corresponding eigenvector ψ_i. A *complete* or *irreducible* representation for a given mechanical system is given by a set of *commuting* observables **s** such that their eigenvectors ψ_i span the whole space, i.e. such that any $\psi = \sum_i a_i \psi_i$. Hence we obtain directly from ψ the joint distribution of the variables s. It is known, however, that these **s** are not sufficient in themselves to specify the system completely; we need, in addition, another complementary set, say **r**, which does not in general commute with **s**; for example, a complete representation is given by either the Cartesian coordinates **q** or their conjugate momenta **p**, but the complete dynamical specification of the system requires both **q**'s and **p**'s. Hence, the phase-space distributions of complete sets of dynamical variables, which are required for a statistical theory, are not given directly by ψ.

It has been argued (3) that such distributions do not exist, because of the impossibility of measuring non-commuting observables simultaneously. This argument is not conclusive for two reasons; one is that the impossibility of physical measurements does not preclude us from *considering* the proposition that there exists a well-defined probability for the two variables to take specified values or sets of values; in fact, the theory of probability is introduced to deal with such situations where exact measurement is impossible (see Jeffreys (4)). The other reason is that it is possible in principle to form operators **G** corresponding to functions $G(r, s)$ of non-commuting observables; the expectation value of **G** in a state ψ is then given by the scalar product $(\psi, \mathbf{G}\psi)$. But the joint distribution of r and s can be reconstructed from a set of such expectation values, e.g. the values of all the joint moments $\overline{r^k s^n}$. The formalism of quantum theory allows us therefore to derive the phase-space distributions indirectly *if a theory of functions of non-commuting observables is specified and conversely.*

There are serious difficulties to be met, however, in defining these distributions unambiguously. This may be seen, for example, in the case of the harmonic oscillator. The energy eigenvalues form a discrete set $E_n = (n+\frac{1}{2})h\nu$. The corresponding eigenfunctions $u_n(q)$, $v_n(p)$ are sets of Hermite functions, continuous in p and q. Hence any joint distribution for p and q in a state consistent with the individual distributions

$$\psi(q)\,\psi^*(q) = \sum_{i,k} a_i^* a_k u_i^*(q)\, u_k(q) \quad \text{and} \quad \phi(p)\,\phi^*(p) = \sum_{i,k} a_i^* a_k v_i^*(p)\, v_k(p)$$

must extend continuously over the whole (p, q) plane, while any joint distribution for the energy $H = \frac{1}{2}(p^2/m + 2\pi m\nu q^2)$ and the phase angle $\theta = \tan^{-1} p/q$ consistent with probabilities $a_n a_n^*$ for E_n, will be concentrated on a set of ellipses

$$\tfrac{1}{2}(p^2/m + 2\pi m\nu q^2) = (n+\tfrac{1}{2})h\nu.$$

We are thus forced to the conclusion that *phase-space distributions are not unique for a given state, but depend on the variables one is going to measure.* In Heisenberg's words (5), 'the statistical predictions of quantum theory are thus significant only when combined with experiments which are actually capable of observing the phenomena treated by the statistics'. Since the introduction of statistical concepts in atomic theory is justified by an analysis of the interaction between observed system and observer, it is perhaps not surprising that different distributions should arise according to the

experimental set-up. For example, measurement of the spectra of an atom corresponds to a distribution with discrete values for the energy and angular momenta. Direct transformation of this distribution to (p, q) space, corresponding to a distribution concentrated on discrete orbits, would not be appropriate for the treatment of collisions of the same atom with a beam of electrons; the appropriate distribution in the latter case arises from wave functions filling the whole space continuously, and is incompatible with discrete orbits.

The statistical interpretation of quantum kinematics will thus have to give methods for setting up the appropriate phase-space distributions of each *basic system of dynamical variables* in terms of the wave vectors, and for transforming such distribution into one another.

3. Phase-space distributions in terms of wave vectors

We denote by $\mathbf{r}$ a set of commuting observables or operators giving a complete representation, $\mathbf{s}$ the *complementary* set, such that $\mathbf{s}$ do not commute with $\mathbf{r}$ and that $\mathbf{r}$ and $\mathbf{s}$ together form a *basic set of dynamical variables*, characterizing a given system; r and s are their *possible values* or eigenvalues (these are, of course, ordinary commuting variables). The most natural way of obtaining the phase-space distribution $F(r, s)$ is to look for its Fourier inverse, i.e. the mean of $\exp\{i(\tau r + \theta s)\}$ (known in statistical terminology as the *characteristic function*). On forming the corresponding operator

$$\mathbf{M}(\tau, \theta) = \exp\{i(\tau\mathbf{r} + \theta\mathbf{s})\} = \sum_n \frac{i^n}{n!}(\tau\mathbf{r} + \theta\mathbf{s})^n, \tag{3·1}$$

the characteristic function in a state ψ is given by the scalar product

$$M(\tau, \theta) = (\psi, e^{i(\tau\mathbf{r} + \theta\mathbf{s})}\psi). \tag{3·2}$$

From well-known formulae for Fourier inversion, the phase-space distribution function is then

$$F(r, s) = \frac{1}{4\pi^2}\iint (\psi, e^{i(\tau\mathbf{r} + \theta\mathbf{s})}\psi)\, e^{-i(\tau r + \theta s)}\, d\tau\, d\theta \tag{3·3}$$

for continuous eigenvalues†, and

$$F(r_i, s_k) = \lim_{T\to\infty} \frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} (\psi, e^{i(\tau\mathbf{r} + \theta\mathbf{s})}\psi)\, e^{-i(\tau r_i + \theta s_k)}\, d\tau\, d\theta \tag{3·4}$$

for discrete eigenvalues r_i, s_k (Cramér (6))‡.

The operator (3·1) takes a specially simple form for canonically conjugate coordinates and momenta $\mathbf{q}$, $\mathbf{p}$ ($\mathbf{pq} - \mathbf{qp} = \hbar/i$),

$$\mathbf{M}(\tau, \theta) = e^{\frac{1}{2}i\hbar\tau\theta}\, e^{i\theta\mathbf{q}}\, e^{i\tau\mathbf{p}} = e^{-\frac{1}{2}i\tau\mathbf{p}}\, e^{i\theta\mathbf{q}}\, e^{\frac{1}{2}i\tau\mathbf{p}} \tag{3·5}$$

(cf. Kermack and McCrea (7)). From the second expression for $\mathbf{M}$, we find

$$M(\tau, \theta) = \int \psi^*(q - \tfrac{1}{2}\hbar\tau)\, e^{i\theta q}\, \psi(q + \tfrac{1}{2}\hbar\tau)\, dq, \tag{3·6}$$

† When no limits are specified, all integrals are to be taken as from $-\infty$ to $+\infty$.

‡ The term distribution function is used in this paper to denote the probability density of continuous eigenvalues, and the finite probability of discrete eigenvalues.

and hence by Fourier inversion

$$F(p,q) = \frac{1}{2\pi}\int \psi^*(q-\tfrac{1}{2}\hbar\tau)\, e^{-i\tau p}\, \psi(q+\tfrac{1}{2}\hbar\tau)\, d\tau, \tag{3·7}$$

an expression first given by Wigner (8). From the first operator form of **M** in (3·5), and by expressing $\psi(q)$ in terms of the momentum wave function $\phi(p)$

$$\psi(q) = h^{-\frac{1}{2}}\int \phi(p)\, e^{ipq/\hbar}\, dp, \tag{3·8}$$

we find, by a series of partial integrations,

$$\begin{aligned} M(\tau,\theta) &= h^{-\frac{1}{2}}\iint [\psi^*(q)\phi(p)\, e^{ipq/\hbar}]\, e^{-\frac{1}{2}i\hbar\tau\theta}\, e^{i(\tau p+\theta q)}\, dp\, dq \\ &= h^{-\frac{1}{2}}\iint e^{\frac{1}{2}(\hbar/i)\,\partial^2/\partial p\,\partial q}\, [\psi^*(q)\phi(p)\, e^{ipq/\hbar}]\, e^{i(\tau p+\theta q)}\, dp\, dq, \end{aligned} \tag{3·9}$$

and hence the alternative expression for the phase-space distribution

$$F(p,q) = h^{-\frac{1}{2}}\, e^{\frac{1}{2}(\hbar/i)\,\partial^2/\partial p\,\partial q}\, [\psi^*(q)\phi(p)\, e^{ipq/\hbar}]. \tag{3·10}$$

It is shown in Appendix 1 that the Heisenberg inequality $\Delta p\,\Delta q \geqslant \frac{1}{2}\hbar$ follows directly from the expression for $F(p,q)$ given above. In this sense, the expression of the phase-space distributions in terms of the wave vectors may be considered as a more complete formulation of the uncertainty principle than that given by the inequalities, since it should contain all possible restrictions on the probabilities and expectation values of non-commuting observables.

This choice of expression for the phase-space distributions constitutes a new hypothesis, not already included in the basic postulates of quantum theory as they are usually formulated. The discussion of certain difficulties associated with this choice, in particular the appearance of 'negative probabilities' for certain states, is made clearer by further developments of the theory, and will therefore be deferred to § 15. Other possible choices and the possibilities of experimental verification are discussed briefly in § 17.

4. PHASE-SPACE EIGENFUNCTIONS

If we insert the expansion of the wave vector ψ in terms of an orthonormal set of eigenvectors

$$\psi = \sum_l a_l \psi_l \tag{4·1}$$

in the expression (3·3) for $F(r,s)$, we find for the latter the expansion

$$F(r,s) = \sum_{l,k} a_l^* a_k f_{lk}(r,s), \tag{4·2}$$

where the functions $f_{lk}(r,s)$ are the Fourier inverses of the matrices

$$m_{lk}(\tau,\theta) = (\psi_l, e^{i(\tau\mathbf{r}+\theta\mathbf{s})}\psi_k) = m_{kl}^*(-\tau,-\theta) \tag{4·3}$$

of the operator (3·1) in the representation of the ψ_l. Explicitly, we have

$$f_{lk}(r,s) = \frac{1}{4\pi^2}\iint (\psi_l, e^{i(\tau\mathbf{r}+\theta\mathbf{s})}\psi_k)\, e^{-i(\tau r+\theta s)}\, d\tau\, d\theta, \tag{4·4}$$

$$f_{lk}(r_\alpha, s_\beta) = \lim_{T\to\infty}\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} (\psi_l, e^{i(\tau\mathbf{r}+\theta\mathbf{s})}\psi_k)\, e^{-i(\tau r_\alpha+\theta s_\beta)}\, d\tau\, d\theta, \tag{4·5}$$

where (4·4) refers to the case of continuous eigenvalues r, s and (4·5) to that of discrete eigenvalues r_α, s_β. The functions $f_{lk}(r,s)$ form a complete orthogonal set in the Hilbert space of the phase-space functions $F(r,s)$, satisfying the relations†

$$\iint f_{lk}(r,s) f^*_{l'k'}(r,s)\,dr\,ds = h^{-1}\delta_{ll'}\delta_{kk'}, \tag{4·6}$$

$$\sum_{l,k} f_{lk}(r,s) f^*_{lk}(r',s') = h^{-1}\delta(r-r')\,\delta(s-s'), \tag{4·7}$$

and also the 'self-orthogonality' relations

$$\iint f_{lk}(r,s)\,dr\,ds = \delta_{lk}, \tag{4·8}$$

$$\sum_{l} f_{ll}(r,s) = h^{-1}. \tag{4·9}$$

In the general case, this follows from the fact that (4·3) and (4·4) or (4·5) form a unitary transformation from a vector, say ψ_{lk}, of components ψ_l^*, ψ_k in the product space of the vectors ψ^* with the vectors ψ, to f_{lk}. The vectors ψ_{lk} form a complete orthogonal (and self-orthogonal) set, and these properties are invariant under a unitary transformation. Furthermore, it is easily seen from their definition that the f_{lk} form a Hermitian matrix with respect to their subscripts l, k

$$f_{lk}(r,s) = f^*_{kl}(r,s). \tag{4·10}$$

We shall see later (§§ 7 and 8) that the f_{lk} can be interpreted as the eigenfunctions of characteristic equations for the phase-space distribution functions, corresponding to the eigenvalue equations of the ψ's; we therefore call them *phase-space eigenfunctions*.

In the case of the canonical coordinates and momenta q and p, relations (4·6)–(4·9) can be proved by elementary methods (cf. Appendix 2), and the $f_{lk}(p,q)$ have the explicit expressions, corresponding to (3·7) and (3·8),

$$f_{lk}(p,q) = \frac{1}{2\pi}\int \psi_l^*(q-\tfrac{1}{2}\hbar\tau)\,e^{-i\tau p}\,\psi_k(q+\tfrac{1}{2}\hbar\tau)\,d\tau, \tag{4·11}$$

$$f_{lk}(p,q) = h^{-1}\,e^{\frac{1}{2}(\hbar/i)\,\partial^2/\partial p\,\partial q}[\psi_l^*(q)\,\phi_k(p)\,e^{ipq/\hbar}]. \tag{4·12}$$

Substituting the eigenfunctions $\psi_{p'}(q) = h^{-\frac{1}{2}}\,e^{ip'q/\hbar}$ in a p-representation, we find

$$f_{p'p''}(p,q) = h^{-1}\,\delta\left(p - \frac{p'+p''}{2}\right) e^{iq(p''-p')/\hbar}. \tag{4·13}$$

The expansion of $F(p,q)$ in terms of $f_{p'p''}$

$$\begin{aligned} F(p,q) &= h^{-1}\iint \phi^*(p')\,\phi(p'')\,\delta\left(p - \frac{p'+p''}{2}\right) e^{iq(p''-p')/\hbar}\,dp''\,dp' \\ &= \frac{1}{2\pi}\int \phi^*(p+\tfrac{1}{2}\hbar\theta)\,e^{-i\theta q}\,\phi(p-\tfrac{1}{2}\hbar\theta)\,d\theta, \end{aligned} \tag{4·14}$$

is the equivalent of (3·7) in terms of the momentum wave functions $\phi(p)$.

† Integration must be replaced by summation in what follows when the eigenvalues of **r**, **s** are discrete.

5. MEAN VALUES, OPERATORS AND MATRICES OF FUNCTIONS OF THE DYNAMICAL VARIABLES

The mean value of an ordinary function $G(r,s)$ taken with respect to the phase-space distribution $F(r,s)$ is

$$\begin{aligned}\bar{G} &= \iint G(r,s)\,F(r,s)\,dr\,ds \\ &= \iiiint G(r,s)\,(\psi, e^{i(\tau\mathbf{r}+\theta\mathbf{s})}\psi)\,e^{-i(\tau r+\theta s)}\,dr\,ds\,d\tau\,d\theta \\ &= \left(\psi, \left\{\iint \gamma(\tau,\theta)\,e^{i(\tau\mathbf{r}+\theta\mathbf{s})}\,d\tau\,d\theta\right\}\psi\right), \qquad (5{\cdot}1)\end{aligned}$$

where $\gamma(\tau,\theta)$ is the ordinary Fourier inverse of $G(r,s)$

$$\gamma(\tau,\theta) = \iint G(r,s)\,e^{-i(\tau r+\theta s)}\,dr\,ds. \qquad (5{\cdot}2)$$

$\bar{G}$ is thus the mean of the operator

$$\mathbf{G} = \iint \gamma(\tau,\theta)\,e^{i(\tau\mathbf{r}+\theta\mathbf{s})}\,d\tau\,d\theta, \qquad (5{\cdot}3)$$

which is thus the operator corresponding to the ordinary function $G(r,s)$ in our theory.

It now follows that the *matrix G_{lk} of G in any representation of eigenvectors ψ_l can be obtained by integration of the ordinary function $G(r,s)$ with respect to the corresponding phase-space eigenfunction $f_{lk}(r,s)$*

$$\begin{aligned}G_{lk} &= \iint G(r,s)\,f_{lk}(r,s)\,dr\,ds = \iiiint G(r,s)\,(\psi_l, e^{i(\tau\mathbf{r}+\theta\mathbf{s})}\psi_k)\,dr\,ds\,d\tau\,d\theta \\ &= (\psi_l, \mathbf{G}\psi_k). \qquad (5{\cdot}4)\end{aligned}$$

Since f_{lk} is a Hermitian matrix with respect to l and k, we see at once from (5·4) that G_{lk} will be Hermitian if $G(r,s)$ is real.

The operators and matrices corresponding to any function of the basic variables r, s are thus uniquely defined by the phase-space distributions. In other words, our theory of phase-space distributions is equivalent to a theory of functions of non-commuting operators. Inversely, this theory of functions defines the phase-space distributions uniquely.

In the special case of functions $G(p,q)$ of canonically conjugate coordinates and momenta, (5·3) coincides with an expression derived by Weyl (9) on group-theoretical considerations. An alternative expression corresponding to (3·10) for $F(p,q)$ is

$$\mathbf{G} = e^{\frac{1}{2}(\hbar/i)\,\partial^2/\partial p\,\partial q}\,\mathbf{G}_0(\mathbf{q},\mathbf{p}), \qquad (5{\cdot}5)$$

where $\mathbf{G}_0(\mathbf{q},\mathbf{p})$ is obtained directly from the ordinary function $G(p,q)$ by writing all the operators $\mathbf{p}$ to the right (e.g. $\mathbf{q}^n\mathbf{p}^m$), and this order is maintained when applying the operator $e^{\frac{1}{2}(\hbar/i)\,\partial^2/\partial p\,\partial q}$ to $\mathbf{G}_0$ (cf. Appendix 3 for the proof; see also McCoy (10)). The form of the usual operators of quantum theory: energy, angular momenta, radial momenta, etc., are not changed when they are derived by this method from the corresponding classical functions of p and q.

II. QUANTUM DYNAMICS

6. The laws of motion of general dynamical stochastic processes

We now come to the statistical interpretation of quantum dynamics. What we have to do for this purpose is to find the temporal transformation laws of the phase-space distributions of quantum theory corresponding to the quantum equations of motion. As mentioned in § 1, this cannot be done within the framework of classical statistical mechanics, which is a 'crypto-deterministic' theory, but appears rather as a special case in the general theory of dynamical stochastic processes. We start therefore with a brief survey of the integral and differential relations through which laws of motion can be expressed for such processes. The theory will be developed for Cartesian coordinates and momenta only.

The fundamental integral relation connecting the probability distributions $F(p,q;t)$ and $F_0(p_0,q_0;t_0)$ at times t and t_0 for a given mechanical system is

$$F(p,q;t) = \iint K(p,q \mid p_0,q_0;t-t_0)\, F_0(p_0,q_0;t_0)\, dp_0 dq_0, \tag{6·1}$$

where K is the distribution of p, q at t *conditional* in p_0, q_0 at t_0. K is therefore the temporal *transformation function*, and must express the laws of motion of the system. While F_0 and F depend on the initial and final *states* of the system, K must be independent of these states, and depend on the *inherent dynamical properties of the system*. Hence the assumption that K is *homogeneous*, i.e. invariant for a translation of the origin in t, and dependent only on the interval $t-t_0$ (as long as there are no external time-dependent forces acting on the system).

K gives the transformation for finite intervals. We now derive the corresponding infinitesimal transformation. The characteristic function Λ for the differences $q-\xi$, $p-\eta$ conditional in ξ, η is

$$\Lambda(\tau,\theta \mid \eta,\xi;t-t_0) = \iint e^{i[\theta(q-\xi)+\tau(p-\eta)]} K(p,q \mid \eta,\xi;t-t_0)\, dp\, dq. \tag{6·2}$$

We make the second assumption that in the stochastic processes of physics, the probability of a transition from ξ,η to $q \neq \xi$, $p \neq \eta$ in a small interval $t-t_0$ is of the order of $t-t_0$. For $t=t_0$, obviously $K = \delta(p-\eta)\,\delta(q-\xi)$ and $\Lambda = 1$. Hence $(\Lambda-1)/(t-t_0)$ tends to a finite limit L when $t \to t_0$

$$\lim_{t \to t_0} \frac{\Lambda-1}{t-t_0} = L(\tau,\theta \mid \eta,\xi). \tag{6·3}$$

We shall call L the *derivate characteristic function*. If $M(\tau,\theta;t_0)$ is the characteristic function at t_0

$$M(\tau,\theta;t_0) = \iint e^{i(\tau\eta+\theta\xi)} F_0(\eta,\xi;t_0)\, d\eta\, d\xi, \tag{6·4}$$

then the characteristic function at t is

$$M(\tau,\theta;t) = \iint e^{i(\tau\eta+\theta\xi)} \Lambda(\tau,\theta \mid \eta,\xi;t-t_0)\, F_0(\eta,\xi;t_0)\, d\eta\, d\xi.$$

Hence
$$\frac{\partial M}{\partial t} = \lim_{t_0 \to t} \iint \frac{\Lambda - 1}{t - t_0} e^{i(\tau\eta + \theta\xi)} F_0(\eta, \xi; t_0)\, d\eta\, d\xi$$
$$= \iint L(\tau, \theta \mid \eta, \xi)\, e^{i(\tau\eta + \theta\xi)} F(\eta, \xi; t)\, d\eta\, d\xi. \tag{6·5}$$

This can be expressed in the operational form
$$\frac{\partial M}{\partial t} = L\left(\tau, \theta \,\middle|\, \frac{1}{i}\frac{\partial}{\partial \tau}, \frac{1}{i}\frac{\partial}{\partial \theta}\right) M(\tau, \theta; t) \tag{6·6}$$

(first suggested to the author by Prof. M. S. Bartlett). (6·5) and (6·6) express the infinitesimal transformation corresponding to (6·1) in terms of characteristic functions; they can be inverted to express this transformation directly in terms of distribution functions. This may be achieved in two ways; if L admits a Fourier inverse
$$S(p, q \mid \eta, \xi) = \iint L(\tau, \theta \mid \eta, \xi)\, e^{i[\tau(\eta - p) + \theta(\xi - q)]}\, d\tau\, d\theta, \tag{6·7}$$

we obtain for F the integro-differential equation
$$\frac{\partial}{\partial t} F(p, q; t) = \iint S(p, q \mid \eta, \xi)\, F(\eta, \xi; t)\, d\eta\, d\xi. \tag{6·8}$$

If, on the other hand, it is possible to expand L in the form
$$L(\tau, \theta \mid \eta, \xi) = \lim_{t \to t_0} \iint \sum_{n=0}^{\infty} \sum_{r=0}^{n} \frac{(i\tau)^{n-r}(i\theta)^r}{(n-r)!\, r!} \frac{(p-\eta)^{n-r}(q-\xi)^r}{t - t_0} K(p, q \mid \eta, \xi; t - t_0)\, dp\, dq$$
$$= \sum_{n=0}^{\infty} \sum_{r=0}^{n} \frac{(i\tau)^{n-r}(i\theta)^r}{(n-r)!\, r!} \alpha_{nr}(\eta, \xi) \tag{6·9}$$

(where the $\alpha_{nr}(\eta, \xi)$ are called the *derivate moments* of the system), then F satisfies the differential equation of infinite order
$$\frac{\partial}{\partial t} F(p, q; t) = \sum_{n=0}^{\infty} \sum_{r=0}^{n} \frac{(-1)^n}{(n-r)!\, r!} \left(\frac{\partial}{\partial p}\right)^{n-r} \left(\frac{\partial}{\partial q}\right)^r [\alpha_{nr}(p, q)\, F(p, q; t)]. \tag{6·10}$$

This reduces to an equation of finite order if the expansion (6·9) for L terminates, i.e. if the derivate moments vanish above given powers of p and q.

7. Equations of the motion for the phase-space distributions of quantum theory

In order to derive the equations of motion for the quantum phase-space distributions, we look for the time derivatives of their characteristic functions. We find from the Poisson-bracket form of the quantum equations of motion
$$\frac{\partial M}{\partial t} = \int \psi^*(q)\, [\mathbf{M}, \mathbf{H}]\, \psi(q)\, dq = \frac{i}{\hbar} \int \psi^*(q)\, [\mathbf{MH} - \mathbf{HM}]\, \psi(q)\, dq, \tag{7·1}$$

where $\mathbf{M}(\tau, \theta)$ is the characteristic function operator (3·5), and $\mathbf{H}$ the Hamiltonian operator, expressed from (5·3) by
$$\mathbf{H} = \iint W(\sigma, \mu)\, e^{i(\sigma\mathbf{p} + \mu\mathbf{q})}\, d\sigma\, d\mu, \tag{7·2}$$

$W(\sigma,\mu)$ being the Fourier inverse of the corresponding classical Hamiltonian $H(p,q)$. Hence, using expression (3·5), we obtain

$$\frac{\partial M}{\partial t} = \frac{i}{\hbar}\iiint e^{\frac{1}{2}i\hbar(\tau\theta+\sigma\mu)}\left[e^{i\hbar\sigma\theta}-e^{i\hbar\tau\mu}\right]W(\sigma,\mu)\,\psi^*(q)e^{i(\theta+\mu)q}e^{i(\tau+\sigma)p}\psi(q)\,dq\,d\sigma\,d\mu$$

$$= \frac{2}{\hbar}\iiiint \sin\tfrac{1}{2}\hbar(\tau\mu-\sigma\theta)\,e^{i[(\tau+\sigma)p+(\theta+\mu)q]}W(\sigma,\mu)$$
$$\times e^{\frac{1}{2}(\hbar/i)\partial^2/\partial p\,\partial q}\left[h^{-1}\psi^*(q)\,\phi(p)\,e^{ipq/\hbar}\right]dp\,dq\,d\sigma\,d\mu.$$

Using expression (3·10) for $F(p,q;t)$, we obtain the two equivalent expressions

$$\frac{\partial M}{\partial t} = \frac{i}{\hbar}\iint\left[H(p+\tfrac{1}{2}\hbar\theta,q-\tfrac{1}{2}\hbar\tau)-H(p-\tfrac{1}{2}\hbar\theta,q+\tfrac{1}{2}\hbar\tau)\right]F(p,q;t)\,e^{i(\tau p+\theta q)}\,dp\,dq, \tag{7·3}$$

$$\frac{\partial M}{\partial t} = \iint e^{i(\tau p+\theta q)}\left\{\frac{2}{\hbar}\sin\frac{\hbar}{2}\left[\frac{\partial}{\partial p_F}\frac{\partial}{\partial q_H}-\frac{\partial}{\partial p_H}\frac{\partial}{\partial q_F}\right]H(p,q)\,F(p,q;t)\right\}dp\,dq, \tag{7·4}$$

where $\partial/\partial p_H$, $\partial/\partial q_H$ in the right hand of (7·4) operate only on H and $\partial/\partial p_F$, $\partial/\partial q_F$ only on F. The comparison of (7·3) with (6·5) gives the derivate characteristic function

$$L(\tau,\theta\mid p,q) = \frac{i}{\hbar}\left[H(p+\tfrac{1}{2}\hbar\theta,q-\tfrac{1}{2}\hbar\tau)-H(p-\tfrac{1}{2}\hbar\theta,q+\tfrac{1}{2}\hbar\tau)\right]. \tag{7·5}$$

If L possesses a Fourier transform

$$S(p,q\mid\eta,\xi) = \frac{i}{\hbar}\iint\left[H(\eta+\tfrac{1}{2}\hbar\theta,\xi-\tfrac{1}{2}\hbar\tau)-H(\eta-\tfrac{1}{2}\hbar\theta,\xi+\tfrac{1}{2}\hbar\tau)\right]e^{i[\tau(\eta-p)+\theta(\xi-q)]}\,d\tau\,d\theta, \tag{7·6}$$

then $F(p,q;t)$ satisfies an integro-differential equation of form (6·8)

$$\frac{\partial}{\partial t}F(p,q;t) = \iint S(p,q\mid\eta,\xi)\,F(\eta,\xi;t)\,d\eta\,d\xi, \tag{7·7}$$

with the kernel S given by (7·6). Similarly, we find from (7·4)

$$\frac{\partial}{\partial t}F(p,q;t) = \frac{2}{\hbar}\sin\frac{\hbar}{2}\left[\frac{\partial}{\partial p_F}\frac{\partial}{\partial q_H}-\frac{\partial}{\partial p_H}\frac{\partial}{\partial q_F}\right]H(p,q)\,F(p,q;t), \tag{7·8}$$

which is easily shown equivalent to (6·10) with derivate moments

$$\alpha_{2n+1,r}(p,q) = (-1)^{n+r}(\tfrac{1}{2}\hbar)^{2n}\left(\frac{\partial}{\partial p}\right)^r\left(\frac{\partial}{\partial q}\right)^{2n+1-r}H(p,q), \quad \alpha_{2n,r}(p,q)\equiv 0. \tag{7·9}$$

Inversely, the quantum equations of motion, and in particular the Schrödinger equation, may be derived from the equations above for $F(p,q;t)$ (cf. Appendix 4). There is thus complete equivalence between the two.

Finally, we may notice the analogy between the right-hand side of (7·8) and the classical Poisson bracket. This may be generalized in the following way. It may be shown by a method similar to that leading to (7·8), that the commutator $i\hbar[\mathbf{RG}-\mathbf{GR}]$ of two operators $\mathbf{R}$, $\mathbf{G}$ obtained (e.g. by (5·3) or (5·5)) from the ordinary functions $R(p,q)$, $G(p,q)$ is identical with the operator corresponding (by the same rules) to

$$\frac{2}{\hbar}\sin\frac{\hbar}{2}\left[\frac{\partial}{\partial p_G}\frac{\partial}{\partial q_R}-\frac{\partial}{\partial p_R}\frac{\partial}{\partial q_G}\right]R(p,q)\,G(p,q). \tag{7·10}$$

In other words, (7·10) is the analogue of the classical Poisson bracket when the laws of quantum mechanics are expressed in phase-space, and the commutator is the

corresponding operator in a q- or p-representation. It is also seen from this that operators whose classical analogue is 0 may correspond to non-vanishing phase-space functions in the present theory†.

8. THE CHARACTERISTIC EQUATIONS OF PHASE-SPACE EIGENFUNCTIONS

The expansion of the distributions $F(p,q;t)$ of a conservative system in terms of its energy phase-space eigenfunctions $f_{ik}(p,q)$ is, from (4·2),

$$F(p,q;t) = \sum_{i,k} a_i^* a_k f_{ik}(p,q)\, e^{i(E_i - E_k)t/\hbar}. \tag{8·1}$$

Substituting in (7·7) and identifying term-by-term, we see that the f_{ik} are the eigenfunctions of the homogeneous integral equation

$$f_{ik}(p,q) = \frac{i\hbar}{E_k - E_i}\iint S(p,q\,|\,\eta,\xi)\, f_{ik}(\eta,\xi)\, d\eta\, d\xi. \tag{8·2}$$

The kernel S can therefore be expanded in terms of the f_{ik}

$$S(p,q\,|\,\eta,\xi) = 2\pi i \sum_{i,k} (E_i - E_k)\, f_{ik}(p,q)\, f_{ik}^*(\eta,\xi). \tag{8·3}$$

Similar characteristic equations can be found for the eigenfunctions $g_{ik}(p,q)$ of any operator $\mathbf{G}$ corresponding to the classical function $G(p,q)$. Let γ_i be the eigenvalues of $\mathbf{G}$

$$\mathbf{G}u_i(q) = \gamma_i u_i(q). \tag{8·4}$$

Calculating the mean of the commutator $[\mathbf{G},\mathbf{M}]$ from the two sides of (8·4)

$$\begin{aligned}\int u_i^*(q)\,[\mathbf{GM}-\mathbf{MG}]\,u_k(q)\,dq &= (\gamma_i^* - \gamma_k)\iint e^{i(\tau p+\theta q)}\, g_{ik}(p,q)\, dp\, dq \\ &= \iint [G(p+\tfrac{1}{2}\hbar\theta, q-\tfrac{1}{2}\hbar\tau) - G(p-\tfrac{1}{2}\hbar\theta, q+\tfrac{1}{2}\hbar\tau)]\, g_{ik}(p,q)\, e^{i(\tau p+\theta q)}\, dp\, dq \\ &= \frac{2}{i}\iint e^{i(\tau p+\theta q)} \sin\frac{\hbar}{2}\left[\frac{\partial}{\partial p_g}\frac{\partial}{\partial q_G} - \frac{\partial}{\partial p_G}\frac{\partial}{\partial q_g}\right] G(p,q)\, g_{ik}(p,q)\, dp\, dq,\end{aligned} \tag{8·5}$$

we find the characteristic equations for g_{ik}

$$\begin{aligned} g_{ik}(p,q) &= \frac{i\hbar}{\gamma_k - \gamma_i^*}\iint S_G(p,q\,|\,\eta,\xi)\, g_{ik}(\eta,\xi)\, d\eta\, d\xi \\ &= \frac{2i}{\gamma_k - \gamma_i^*} \sin\frac{\hbar}{2}\left[\frac{\partial}{\partial p_g}\frac{\partial}{\partial q_G} - \frac{\partial}{\partial p_G}\frac{\partial}{\partial q_g}\right] G(p,q)\, g_{ik}(p,q), \end{aligned} \tag{8·6}$$

where the kernel

$$\begin{aligned} S_G(p,q\,|\,\eta,\xi) &= \frac{i}{\hbar}\iint [G(p+\tfrac{1}{2}\hbar\theta, q-\tfrac{1}{2}\hbar\tau) - G(p-\tfrac{1}{2}\hbar\theta, q+\tfrac{1}{2}\hbar\tau)]\, e^{i[\tau(\eta-p)+\theta(\xi-q)]}\, d\tau\, d\theta \\ &= 2\pi i \sum_{i,k} (\gamma_i^* - \gamma_k)\, g_{ik}(p,q)\, g_{ik}^*(\eta,\xi). \end{aligned} \tag{8·7}$$

† This question was raised by the referee.

9. Transformation equations for finite intervals

Having derived the infinitesimal transformations in phase-space, we now return to the transformation equations for a finite interval (cf. § 6)

$$\left.\begin{aligned} F(p,q;t) &= \iint K_{10}(p,q \mid p_0,q_0;\, t-t_0)\, F_0(p_0,q_0;\, t_0)\, dp_0 dq_0, \\ F_0(p_0,q_0;\, t_0) &= \iint K_{01}(p_0,q_0 \mid p,q;\, t_0-t)\, F(p,q;\, t)\, dp\, dq. \end{aligned}\right\} \tag{9·1}$$

We introduce the operator solutions of the Schrödinger equation

$$\mu_k(q;\, t-t_0) = e^{-i(t-t_0)\mathbf{H}/\hbar}\, u_k(q) \tag{9·2}$$

for an arbitrary orthonormal set of functions $u_k(q)$. The corresponding phase-space functions are

$$g_{ik}(p_0,q_0) = \frac{1}{2\pi}\int u_i^*(q_0-\tfrac{1}{2}\hbar\tau)\, e^{-i\tau p_0}\, u_k(q_0+\tfrac{1}{2}\hbar\tau)\, d\tau, \tag{9·3}$$

$$\begin{aligned} \gamma_{ik}(p,q;\, t-t_0) &= \frac{1}{2\pi}\int \mu_i^*(q-\tfrac{1}{2}\hbar\tau;\, t-t_0)\, e^{-i\tau p}\, \mu_k(q+\tfrac{1}{2}\hbar\tau;\, t-t_0)\, d\tau \\ &= \sum_{l,m} U_{il}(t_0-t)\, g_{lm}(p,q)\, U_{mk}(t-t_0), \end{aligned} \tag{9·4}$$

where

$$U_{ik}(t-t_0) = \int u_i^*(q)\, e^{-i(t-t_0)\mathbf{H}/\hbar}\, u_k(q)\, dq. \tag{9·5}$$

On substituting in (9·1) the expansions of $F(p,q;t)$ and $F_0(p_0,q_0;t_0)$ in terms of the g_{ik} and γ_{ik}, a term-by-term identification shows that

$$\begin{aligned} \gamma_{ik}(p,q;\, t-t_0) &= \iint K_{10}(p,q \mid p_0,q_0;\, t-t_0)\, g_{ik}(p_0,q_0)\, dp_0 dq_0, \\ g_{ik}(p_0,q_0) &= \iint K_{01}(p_0,q_0 \mid p,q;\, t-t_0)\, \gamma_{ik}(p,q;\, t-t_0)\, dp\, dq. \end{aligned} \tag{9·6}$$

The expansion of K_{10} in terms of the g_{ik}: $K_{10} = \Sigma\lambda_{ik} g_{ik}$ has coefficients

$$\lambda_{ik} = h\iint K_{10}(p,q \mid p_0,q_0;\, t-t_0)\, g_{ik}^*(p_0,q_0)\, dp_0 dq_0 = h\gamma_{ik}^*(p,q;\, t-t_0),$$

and similarly for K_{01}, so that the two are identical,

$$K_{01} = K_{10} = K(p,q \mid p_0,q_0;\, t-t_0) = h\sum_{i,k} g_{ik}(p_0,q_0)\, \gamma_{ik}^*(p,q;\, t-t_0). \tag{9·7}$$

We have thus found an expression for the transformation function K in terms of the g_{ik} and γ_{ik}; from it we see that K satisfies the iteration relation

$$K(p_2,q_2 \mid p_0,q_0;\, t-t_0) = \iint K(p_2,q_2 \mid p_1,q_1;\, t_2-t_1)\, K(p_1,q_1 \mid p_0,q_0;\, t_1-t_0)\, dp_1 dq_1. \tag{9·8}$$

The transformation (9·1) form therefore a continuous unitary group. Stochastic processes satisfying the iteration relations (9·8) are known as Markoff processes (cf. Hostinsky (**11**); see also Jeffreys (**12**)).

The energy eigenfunctions $f_{ik}(p,q)$ of a conservative system are easily seen from (9·4) and (9·6) to satisfy the homogeneous integral equation

$$f_{ik}(p,q) = e^{-i(E_i-E_k)(t-t_0)/\hbar}\iint K(p,q \mid p_0,q_0;\, t-t_0)\, f_{ik}(p_0,q_0)\, dp_0 dq_0. \tag{9·9}$$

The transformation function of a conservative system therefore forms a kernel symmetrical in p, q, antisymmetrical in t

$$K(p,q \mid p_0,q_0; t-t_0) = h\sum_{i,k} f_{ik}(p,q)\, f^*_{ik}(p_0,q_0)\, e^{i(E_i-E_k)(t-t_0)/\hbar}$$
$$= K(p_0,q_0 \mid p,q; t_0-t). \tag{9·10}$$

An alternative expression for K can be given in terms of the transformation wave function

$$\psi(q \mid q_0; t-t_0) = \sum_n u^*_n(q)\, u_n(q_0)\, e^{iE_n(t-t_0)/\hbar}. \tag{9·11}$$

Substituting expression (4·11) for f_{ik} in (9·10), we find

$$K(p,q \mid p_0,q_0; t-t_0) = \frac{\hbar}{2\pi}\iint e^{-i(\tau p-\tau_0 p_0)}\, \psi^*(q-\tfrac{1}{2}\hbar\tau \mid q_0-\tfrac{1}{2}\hbar\tau_0)\, \psi(q+\tfrac{1}{2}\hbar\tau \mid q_0+\tfrac{1}{2}\hbar\tau_0)\, d\tau\, d\tau_0. \tag{9·12}$$

10. The relation between infinitesimal and finite transformations: application to wave packet and collision problems and to the calculation of transition probabilities

It is seen from the expansions (8·3) and (9·10) of S and K in terms of the energy eigenfunctions that

$$S(p,q \mid p_0,q_0) = \lim_{t_0\to t} \frac{\partial}{\partial(t-t_0)} K(p,q \mid p_0,q_0; t-t_0). \tag{10·1}$$

Inversely, K can be expanded in terms of S

$$K(p,q \mid p_0,q_0; t-t_0)$$
$$= \delta(p-p_0)\,\delta(q-q_0) + \sum_{n=0}^{\infty} \frac{(t-t_0)^{n+1}}{(n+1)!} \int\!\!\cdots\!\!\int_{(n)} S(p,q \mid \eta_1,\xi_1)\, S(\eta_1,\xi_1 \mid \eta_2 \mid \xi_2)$$
$$\cdots S(\eta_n,\xi_n \mid p_0,q_0)\, d\eta_1 d\xi_1 \cdots d\eta_n d\xi_n. \tag{10·2}$$

This is easily verified by substituting from (8·3) for S and comparing with (9·10) for K.

Since S has a simple expression, obtained directly from the Hamiltonian, (10·2) supplies also a convenient method of approximation for K when the energy eigenfunctions are not known exactly. *We have thus a new method of solving problems in quantum mechanics, without having to solve the Schrödinger equation.*

The distribution $F(p,q;t)$ of a wave packet at any time t is obtained by the transformation (9·1) from the initial distribution $F_0(p_0,q_0;t_0)$ at t_0. We can apply this to the solution of collision problems by introducing a suitable initial distribution F_0 describing the motion of the two particles before the collision, and calculating the transformation function K by (10·2) and (7·6) from the Hamiltonian for the colliding particles.

These methods can also be applied to the calculation of transition probabilities. Let $f_{ik}(p,q)$ be the energy eigenfunctions corresponding to the unperturbed Hamiltonian H_0. We can approximate for K from (10·2), using the kernel S corresponding to the complete Hamiltonian $H = H_0 + H_1$ (where H_1 is the perturbing term). Taking a single diagonal eigenfunction $f_{kk}(p_0,q_0)$ as the initial distribution at $t=0$, the expansion of the transformed distribution $F_k(p,q;t)$ at time t in terms of the f_{ik} is

$$F_k(p,q;t) = \sum_{n,m} \alpha^*_{kn}\alpha_{km} f_{nm}(p,q). \tag{10·3}$$

The transition probabilities $c_{kn}(t)$ from state k to state n are the diagonal coefficients $c_{kn} = \alpha^*_{kn}\alpha_{kn}$ whose expression in terms of K will clearly be

$$c_{kn} = \alpha^*_{kn}\alpha_{kn} = \iiiint K(p,q \mid p_0,q_0;t) f_{kk}(p_0,q_0) f_{nn}(p,q)\, dp_0 dq_0 dp\, dq. \quad (10{\cdot}4)$$

11. The problem of determinism in quantum mechanics

The present theory should help to elucidate the question whether quantum mechanics is deterministic in the classical kinetic theory sense†, since it permits a direct comparison between the two. The infinitesimal time transformation of quantum phase-space distributions (7·8) may be written in the form

$$\frac{\partial F}{\partial t} + \frac{2}{\hbar}\sin\frac{\hbar}{2}\left\{\frac{\partial}{\partial p}, \frac{\partial}{\partial q}\right\} H(p,q)\, F(p,q;t) = 0, \quad (11{\cdot}1)$$

where $\{\partial/\partial p, \partial/\partial q\}$ is the phase-space differential operator giving the classical Poisson bracket. The corresponding transformation of classical kinetic theory is given by Liouville's theorem

$$\frac{\partial F}{\partial t} + \left\{\frac{\partial}{\partial p}, \frac{\partial}{\partial q}\right\} H(p,q)\, F(p,q;t) = 0. \quad (11{\cdot}2)$$

Its deterministic character may be seen from the fact that the characteristics of this first order partial differential equation are simply the classical paths in phase-space. Alternatively, we may say that F is an integral invariant of the transformation generated by the operator $\{\partial/\partial p, \partial/\partial q\}$; an element S_0 of phase-space will transform to S_t in the interval t, and

$$\int_{S_0} F(p_0,q_0)\, dp_0 dq_0 = \int_{S_t} F(p,q;t)\, dp\, dq. \quad (11{\cdot}3)$$

This no longer holds in the case of quantum theory; the transformation generated by the operator $(2/\hbar)\sin\frac{1}{2}\hbar\{\partial/\partial p, \partial/\partial q\}$ is equivalent to $\{\partial/\partial p, \partial/\partial q\}$ when applied to Hp, Hq, but *not* in general when applied to HF, so that while S_0 will transform into S_t exactly as for the corresponding classical system, yet generally

$$\int_{S_0} F(p_0,q_0)\, dp_0 dq_0 \neq \int_{S_t} F(p,q;t)\, dp\, dq. \quad (11{\cdot}4)$$

Hence the present theory leads to the conclusion that quantum theory is not generally deterministic in the classical sense.

In the correspondence principle limit, when $\hbar \to 0$, the quantum equation (11·1) is seen to reduce to the classical equation (11·2); this will equally well be the case if the Hamiltonian $H(p,q)$ is a second degree polynomial in q and p, leading to the surprising conclusion that systems such as a free or uniformly accelerated particle, or a harmonic oscillator, are deterministic in quantum theory: this should not be taken too seriously, since even small perturbations or non-linear terms would, according to (11·1), destroy this deterministic character.

The phase-space transformations with time of quantum theory form a continuous unitary group, which reduces therefore to the group of contact transformation of

† Cf. in this connexion Whittaker (2), Jeffreys (12) and also Reichenbach (25).

classical mechanics in the correspondence principle limit and for the 'deterministic' quantum systems whose Hamiltonian is a second degree polynomial; the transformation function K of § 9, which is the probability distribution of p and q at time t conditional in p_0, q_0, at time t_0, degenerates in the classical limit to a singular distribution, with complete concentration of the probability 'mass' on the classical path in phase-space issuing from p_0, q_0; K may then be expressed as a product of delta functions

$$K = \delta[p - p(p_0, q_0, t-t_0)]\,\delta[q - q(p_0, q_0, t-t_0)],$$

where p and q are the classical solutions as functions of the initial values p_0, q_0 and the interval $t-t_0$. The phase-space distributions F at t, will be obtained from F_0 at t_0 by substituting the classical solutions for p and q. This has been shown directly by Prof. M. S. Bartlett and the author in the 'deterministic' cases of the free and uniformly accelerated particle and the harmonic oscillator.

Owing to the fact that the transformation is unitary, the eigenvalues of the integral equations (9·8), (9·9) are all of modulus 1; in fact, of the form

$$\lambda_{ik} = e^{i(E_i - E_k)(t-t_0)/\hbar}.$$

In the theory of discrete Markoff processes (where the random variables have only a discrete and finite set of possible values) characteristic roots of modulus 1 for the transformation matrix correspond to deterministic processes (non-degenerate processes involving roots of the form $|e^{-\mu(t-t_0)}| < 1$). Yet we saw above that the quantum mechanical process is not deterministic in the classical sense. The explanation of this discrepancy must await the further study of unitary-Markoff processes of this type.

III. QUANTUM STATISTICS

12. GIBBS'S ENSEMBLES AND PHASE-SPACE DISTRIBUTIONS

A possible field of application for the statistical approach to quantum mechanics lies in the kinetic theories of matter, where the joint distributions of coordinates and momenta are required. As a first step in this direction, we shall study the equilibrium distributions in large assemblies of similar systems.

The notion of Gibbs's ensemble is translated into the quantum theory of statistical assemblies by introducing 'mixed' states, where the assembly has a probability P_n to be in a state ψ_n and the average of any dynamical variable G is given by the *diagonal* sum

$$\bar{G} = \sum_n (\psi_n, \mathbf{G}\psi_n)\, P_n \qquad (12{\cdot}1)$$

(Dirac (13)); the introduction of Gibbs's ensembles in quantum theory is due to von Neumann. The phase-space distribution corresponding to a Gibbs's ensemble may be found in accordance with the method of § 3, by calculating the mean of $e^{i\Sigma_\sigma(\tau_\sigma r_\sigma + \theta_\sigma s_\sigma)}$ from (12·1) (r_σ, s_σ being the dynamical variables characterizing the assembly), and taking its Fourier inverse. For the Cartesian coordinates and momenta of an assembly of N degrees of freedom

$$M(\tau_\sigma, \theta_\sigma) = \sum_n P_n \int_{(N)} \cdots \int \psi_n^*(q_\sigma)\, e^{i\Sigma_\sigma(\tau_\sigma \mathbf{p}_\sigma + \theta_\sigma \mathbf{q}_\sigma)}\, \psi_n(q_\sigma)\, dq_1 \ldots dq_N, \qquad (12{\cdot}2)$$

and the phase-space distribution ρ is a sum of diagonal eigenfunctions ρ_n (see §§ 3 and 4)

$$\rho(p_\sigma, q_\sigma) = \sum_n \rho_n(p_\sigma, q_\sigma) P_n, \tag{12·3}$$

$$\rho_n(p_\sigma, q_\sigma) = (2\pi)^{-N} \int_{(N)} \dots \int \psi_n^*(q_\sigma - \tfrac{1}{2}\hbar\tau_\sigma)\, e^{-i\Sigma_\sigma \tau_\sigma p_\sigma}\, \psi_n(q_\sigma + \tfrac{1}{2}\hbar\tau_\sigma)\, d\tau_1 \dots d\tau_N$$
$$= \hbar^{-\frac{1}{2}N}\, e^{\frac{1}{2}(\hbar/i)\Sigma_\sigma \partial^2/\partial p_\sigma \partial q_\sigma} [\psi_n^*(q_\sigma)\, \phi_n(p_\sigma)\, e^{i\Sigma_\sigma p_\sigma q_\sigma/\hbar}], \tag{12·4}$$

where $\psi_n(q_\sigma)$, $\phi_n(p_\sigma)$ are the eigenfunctions in q_σ, p_σ representations respectively.

Since each term ρ_n in the right-hand side of (12·3) is a solution of the phase-space equation of the motion (7·8), the transformation with time of ρ will be governed by the same equation, which now appears as a generalization of Liouville's theorem for the probability densities in phase-space of statistical assemblies. Introducing the phase-space differential operator of a Poisson bracket

$$\left\{\frac{\partial}{\partial p_\sigma}, \frac{\partial}{\partial q_\sigma}\right\} H\rho = \sum_\sigma \left[\frac{\partial H}{\partial p_\sigma}\frac{\partial \rho}{\partial q_\sigma} - \frac{\partial H}{\partial q_\sigma}\frac{\partial \rho}{\partial p_\sigma}\right], \tag{12·5}$$

we have symbolically

$$\frac{\partial \rho}{\partial t} + \frac{2}{\hbar}\sin\frac{\hbar}{2}\left\{\frac{\partial}{\partial p_\sigma}, \frac{\partial}{\partial q_\sigma}\right\} H\rho = 0. \tag{12·6}$$

It has been held that the existence of Gibbs's ensembles 'is rather surprising in view of the fact that phase-space has no meaning in quantum mechanics' (Dirac (13)). This apparent paradox is removed by the statistical approach to quantum theory, which leads, as seen above, to an interpretation of ensembles closely analogous to that of classical statistical mechanics.

13. Phase-space distributions of one member of a statistical assembly

We consider now an assembly of similar particles in weak interaction. For a given energy E_n of the whole assembly, we find complexions α_n with a_1 particles of energy ϵ_1, a_2 of energy ϵ_2, ..., a_k of energy ϵ_k, $N = \sum_1^k a_i$, and $E_n = \sum_1^k a_i \epsilon_i$. Assume at first that the energy eigenstates of individual particles are non-degenerate. The eigenfunctions corresponding to α_n are

$$\left.\begin{aligned}
&\text{M.B. case:} && \psi_{\alpha_n} = u_1(q_1)\, u_1(q_2) \dots u_1(q_{a_1})\, u_2(q_{a_1+1}) \dots u_k(q_N),\\
&\text{B.E. case:} && \psi_{\alpha_n} = (N!)^{-\frac{1}{2}} \sum_P P[u_1(q_1)\, u_1(q_2) \dots u_k(q_N)],\\
&\text{F.D. case:} && \psi_{\alpha_n} = (N!)^{-\frac{1}{2}} \sum_p \pm P[u_1(q_1)\, u_1(q_2) \dots u_k(q_N)],
\end{aligned}\right\} \tag{13·1}$$

where M.B. refers to a Maxwell-Boltzmann, B.E. to a Bose-Einstein (symmetrical), and F.D. to a Fermi-Dirac (antisymmetrical), assembly, P denotes all the permutations of the q_σ, and the + or − signs in the F.D. case refer to even or odd permutations. The numbers of distinct wave functions for each energy E_n are

$$\left.\begin{aligned}
&\text{M.B. case:} && C_{\alpha_n} = \frac{N!}{a_1!\, a_2! \dots a_k!},\\
&\text{B.E. case:} && C_{\alpha_n} = 1 \quad \text{for all } \alpha_n,\\
&\text{F.D. case:} && C_{\alpha_n} = \begin{cases} 1 & \text{when all } a_i = 0 \text{ or } 1,\\ 0 & \text{if any } a_i > 1. \end{cases}
\end{aligned}\right\} \tag{13·2}$$

The phase-space distribution $\rho(p_\sigma, q_\sigma)$ and eigenfunctions $\rho_{\alpha_n}(p_\sigma, q_\sigma)$ for the assembly are obtained by substituting from (13·1) in (12·3), (12·4). It is easily seen that in the M.B. case ρ_{α_n} is a product of diagonal eigenfunctions $f_{ii}(p, q)$ of the individual particles only, while in the B.E. and F.D. cases, non-diagonal eigenfunctions occur too.

The phase-space distribution for one particle is obtained by integrating over the coordinates and momenta of the remaining particles

$$f(p_1, q_1) = \int_{2(N-1)} \dots \int \rho(p_\sigma, q_\sigma)\, dp_2 dq_2 dp_3 dq_3 \dots dp_N dq_N. \tag{13·3}$$

Owing to this integration, all terms in ρ_{α_n} involving non-diagonal eigenfunction cancel, because $\iint f_{ik} dp dq = \delta_{ik}$. Hence in all three cases $f(p_1, q_1)$ appears as a sum of diagonal eigenfunctions

$$f(p_1, q_1) = \sum_i n_i f_{ii}(p_1, q_1). \tag{13·4}$$

It is easily shown that the n_i are simply the average frequencies of the occupation numbers a_i of (13·1). Introducing a canonical ensemble, where the P_n of (12·3) are proportional to $e^{-E_n/kT}$, we obtain

$$n_i = \sum_{n, \alpha_n} \frac{a_i}{N} C_{\alpha_n} e^{-E_n/kT} \Big/ \sum_{n, \alpha_n} C_{\alpha_n} e^{-E_n/kT}. \tag{13·5}$$

By substituting from (13·2) for the C_{α_n}, the n_i can be calculated by the method of 'sums-over-states' (Schrödinger (**14**)), leading to the well-known expressions

$$n_i = \frac{1}{(1/\xi)\, e^{\epsilon_i/kT} - \gamma}, \tag{13·6}$$

$$\text{M.B. case: } \gamma = 0; \quad \text{B.E. case: } \gamma = 1; \quad \text{F.D. case: } \gamma = -1, \tag{13·7}$$

which can be substituted in (13·4) to give an explicit expression for the phase-space distribution of one member of an assembly. As usual in equilibrium theory, all results are independent of the type of ensemble provided that the dispersion of the total energy is sufficiently small.

The effect of degeneracy of the individual energy eigenstates is to introduce non-diagonal terms in (13·4). As a result, the n_i must be multiplied by the corresponding order of degeneracy w_i, while the f_{ii} must each be replaced by

$$\overline{f_{ii}}(p_1, q_1) = \frac{1}{w_i} \sum_{k,l} f_{ii,kl}(p_1, q_1), \tag{13·8}$$

where the indices k, l refer to the degenerate phase-space eigenfunctions at the ith level, supposed orthogonal.

The foregoing may be used to justify the introduction of ensembles in quantum theory. If we think of an ensemble as an *assembly of similar assemblies*, then the distribution of one assembly will have the diagonal expansion (12·3) for the same reason that the distribution of one particle in an assembly has the diagonal expansion (13·4), even if the ensemble is in a pure state. If the ensemble consists of an infinite number of *distinguishable* assemblies, then the coefficients P_n of the expansion must be M.B. factors $e^{-E_n/kT}$ (E_n being now the energy of one whole assembly) and we thus have a canonical ensemble.

We may compare averaging over an ensemble to averaging over time. If an assembly is in a *pure* state, non-diagonal terms in the expansion of its distribution function

$$\rho(p_\sigma, q_\sigma; t) = \sum_{i,k} a_i^* a_k \rho_{ik}(p_\sigma, q_\sigma)\, e^{i(E_i - E_k)t/\hbar} \tag{13·9}$$

cancel in a time average, leaving a diagonal expansion similar to (12·3). This is analogous to the *ergodic principle* of classical theory.

14. Joint phase-space distribution for two members of an assembly

The distribution function for two particles is obtained by integrating ρ over the coordinates and momenta of the remaining particles.

$$f(p_1, q_1, p_2, q_2) = \int_{2(N-2)} \dots \int \rho(p_\sigma, q_\sigma)\, dp_3 dq_3 \dots dp_N dq_N. \tag{14·1}$$

In the M.B. case, the integration of each eigenfunction ρ_{α_n} yields only products of diagonal eigenfunctions of the form $f_{ii}(p_1, q_1) f_{kk}(p_2, q_2)$. In the other two cases, it is seen that if $i \neq k$, there will be in addition non-diagonal terms (obtained by permuting the two particles) $f_{ik}(p_1, q_1) f_{ki}(p_2, q_2)$, preceded by a + sign in the B.E. case, a − sign in the F.D. case. Other non-diagonal terms in ρ_{α_n} cancel by integration as in the case of a single particle. Hence we can write for all three cases

$$f(p_1, q_1, p_2, q_2) = \sum_{i,k} n_{ik} f_{ii}(p_1, q_1) f_{kk}(p_2, q_2) + \gamma \sum_{i \neq k} n_{ik} f_{ik}(p_1, q_1) f_{ki}(p_2, q_2), \tag{14·2}$$

where γ has the same meaning as in (13·7). The coefficients of this expansion are easily found to be for a canonical ensemble

$$\left.\begin{aligned} n_{ik} &= \sum_{n,\alpha_n} \frac{a_i a_k}{N(N-1)} C_{\alpha n} e^{-E_n kT} \Big/ \sum_{n,\alpha_n} C_{\alpha n} e^{-E_n/kT} \quad (i \neq k) \\ n_{ii} &= \sum_{n,\alpha_n} \frac{a_i(a_i - 1)}{N(N-1)} C_{\alpha n} e^{-E_n/kT} \Big/ \sum_{n,\alpha_n} C_{\alpha n} e^{-E_n/kT} \end{aligned}\right\}. \tag{14·3}$$

Carrying out the summations in (14·3) by the 'sum-over-states' method, we find that the non-diagonal coefficients ($i \neq k$) are

$$n_{ik} = n_i n_k, \tag{14·4}$$

where the n_i are the average frequencies of the a_i, as given in (13·6), while the diagonal coefficients are

$$\text{M.B. case: } n_{ii} = n_i^2, \quad \text{B.E. case: } n_{ii} = 2n_i^2, \quad \text{F.D. case: } n_{ii} = 0.\dagger \tag{14·5}$$

The last (F.D. case) is of course a result of the exclusion principle. Substituting in (14·2) we have

$$f(p_1, q_1, p_2, q_2) = \sum_{i,k} n_i n_k f_{ii}(p_1, q_1) f_{kk}(p_2, q_2) + \gamma \sum_{i,k} n_i n_k f_{ik}(p_1, q_1) f_{kk}(p_2, q_2), \tag{14·6}$$

which may be written, after comparison with (13·4),

$$f(p_1, q_1, p_2, q_2) = f(p_1, q_1) f(p_2, q_2) + \tfrac{1}{2}\gamma \sum_{i,k} [f_{ik}(p_1, q_1) f_{ki}(p_2, q_2) + f_{ki}(p_1, q_1) f_{ik}(p_2, q_2)]. \tag{14·7}$$

† Strictly speaking, the right-hand sides of (14·4) and (14·5) should be multiplied by a normalizing factor $(1 + \gamma \sum_i n_i^2)$.

We see thus that symmetry (or antisymmetry) conditions *introduce a probability dependence between any two particles in B.E. (or F.D.) assemblies even in the absence of any energy interaction*. For example the coordinates and momenta of the two particles will be correlated, with covariance

$$\left.\begin{aligned}\mu(q_1q_2) &= \overline{q_1q_2} - \overline{q_1}\,\overline{q_2} = \gamma \sum_{i,k} n_i n_k \,|\, Q_{nk} \,|^2, \\ \mu(p_1p_2) &= \overline{p_1p_2} - \overline{p_1}\,\overline{p_2} = \gamma \sum_{i,k} n_i n_k \,|\, P_{nk} \,|^2,\end{aligned}\right\} \tag{14·8}$$

where Q_{nk}, P_{nk} are the matrices of the individual $\mathbf{q}$'s and $\mathbf{p}$'s,

$$Q_{nk} = \iint q f_{nk}(p,q)\,dp\,dq, \quad P_{nk} = \iint p f_{nk}(p,q)\,dp\,dq.$$

It is this dependence which gives rise to the 'exchange energy' between the particles when they interact.

15. LIMITATIONS OF THE STATISTICAL APPROACH TO QUANTUM THEORY

The results obtained so far seem to offer a fairly complete scheme for treating quantum mechanics as a form of statistical dynamics. It is important now to return to the difficulties mentioned at the beginning of this paper, and discuss the limitations of this approach.

First, we notice that phase-space eigenfunctions must generally take negative as well as positive values, since they are orthogonal. Only one eigenfunction (generally the ground state one) may possibly be non-negative for all values of the dynamical variables, except for singular eigenfunctions involving delta functions, such as the momenta eigenfunctions (4·13). Hence, on taking for example Cartesian coordinates and momenta p, q as the basic system, the phase-space distribution in the nth energy eigenstate formed according to the method of § 3 would be the diagonal eigenfunction, $f_{nn}(p,q)$, which can be negative, and is therefore not a true probability. This is not really surprising, because we have seen in § 9 that the dynamical equations are those of a Markoff process. The existence of eigenfunction solutions for the fundamental equations (9·8), (9·9) of Markoff processes is well known (see Hostinsky (11)), and it is also known, that these eigenfunctions are not generally probabilities by themselves. Probability distributions are expressed as non-negative linear combinations of these eigenfunctions.

In the language of quantum theory, we may say that *true probability distributions of any given set of non-commuting variables do not exist for every state*; the physical interpretation would be that where the distribution, as calculated by the method of § 3, can take negative values, it is not an observable quantity. This is a restatement of the necessity, already discussed in § 2, for postulating the existence of different phase-space distributions according to the basic set of dynamical variables. Take, for example, a system composed of one proton and one electron. The distribution $F(p,q)$ corresponding to the $\psi(q)$ of a Gaussian wave-packet is positive for all p and q, and is hence an observable quantity. On the other hand, there would be no observable (p,q)

distributions for the energy eigenstates of a hydrogen atom, though an observable distribution may exist for some other set of variables.

It is usually accepted that a dynamical variable $\mathbf{G}$ is exactly equal to its eigenvalue g_n when the system is in the corresponding eigenstate. This means that the operator $\mathbf{W}$ corresponding to the function $W(G)$ should be equal to the function W of the operator $\mathbf{G}$, $\mathbf{W} = W(\mathbf{G})$, since if G is exactly equal to g_n the mean of W is $\overline{W} = W(g_n)$, and hence

$$\overline{W} = (\psi_n, \mathbf{W}\psi_n) = (\psi_n, W(\mathbf{G})\,\psi_n) = (\psi_n, W(g_n)\,\psi_n) = W(g_n). \tag{15·1}$$

Now it is easily seen (Appendix 5) that according to the theory of functions of § 5 this condition is fulfilled only when G is a function of some linear combination of the basic variables r, s: $G(ar+bs)$. This again is connected with the necessity for phase-space distributions adapted to the experimental situation; if the latter involves observation of G, then the distributions must be set up for some set of variables r, s such that $G = G(ar+bs)$.

In order for the scheme to be consistent, it should be possible to prove that if a state ψ admits a non-negative phase-space distribution F at the time $t = 0$, then F will be non-negative at any time t. This is easily seen for isolated systems possessing at least one cyclic coordinate θ. Suppose that θ and its conjugate g are obtained by a canonical transformation from the original system q_i, p_i, and let Q_i, P_i be the other (transformed) coordinates and momenta, $H(g, \theta, P_i, Q_i)$ the transformed Hamiltonian. Then

$$\frac{\partial H}{\partial \theta} = 0, \quad \frac{\partial H}{\partial g} = \text{constant} = \omega. \tag{15·2}$$

The transformed equation of the motion (7·8) can be written

$$\frac{\partial F}{\partial t} + \omega \frac{\partial F}{\partial \theta} + \frac{2}{\hbar} \sin \frac{\hbar}{2} \left(\frac{\partial}{\partial P_i}, \frac{\partial}{\partial Q_i} \right) HF = 0. \tag{15·3}$$

Separating the variables, we have

$$\left.\begin{aligned} F(g, \theta, P_i, Q_i; t) &= F_1(\theta, t)\, F_r(g, P_i, Q_i), \\ \frac{1}{F_1}\left(\frac{\partial F_1}{\partial t} + \omega \frac{\partial F_1}{\partial \theta}\right) &= 2i\mu \quad (\mu \text{ constant}), \\ F_1 &= e^{i\mu(t+\theta/\omega)}. \end{aligned}\right\} \tag{15·4}$$

Comparing with the expansion of F in energy eigenfunctions, we see that it must be of the form

$$F(g, \theta, P_i, Q_i; t) = \sum_{i,k} a_i^* a_k Q_{ik}(g, P_i, Q_i)\, e^{i\{(E_i - E_k)(t+\theta/\omega)\}/\hbar}. \tag{15·5}$$

Hence, if $F \geqslant 0$ for all θ at $t = 0$, it must be non-negative for all t. This proof was suggested to the author by Prof. M. S. Bartlett.

Finally, we may discuss the meaning in the present theory of observables having no classical analogue. §§ 2–5 on quantum kinematics are framed so as to apply to such observables as well as to those having a classical analogue. The phase-space distributions represent for both types the joint distributions of eigenvalues for non-commuting sets, and are subject to the same restrictions. The quantum equations of motion in phase-space, on the other hand, were expressed only for Cartesian coordinates and momenta, so as to bring out the relationship with the theory of general stochastic

processes. It is clear, however, that they can be extended to general quantum observables, say $\mathbf{r}$ and $\mathbf{s}$. If $F(r, s, t)$ is their joint distribution, then as in § 7, $\partial F/\partial t$ is obtained by Fourier inversion of

$$\frac{\partial M}{\partial t} = \left(\psi^*, \frac{i}{\hbar}[\mathbf{M}, \mathbf{H}]\psi\right), \tag{15·6}$$

where $\mathbf{M} = e^{i(\tau\mathbf{r}+\theta\mathbf{s})}$.

16. Practical applications of the theory

The foregoing restrictions are necessary as long as we require *probabilities* in phase-space. They may be relaxed in practical applications of the theory, where we introduce phase-space distributions as aids to calculation, and where the observable quantities we wish to calculate are necessarily non-negative, independently of whether the phase-space distribution takes negative values or not. It is not difficult to see that the phase-space distributions and eigenfunctions obtained by the rules of §§ 3 and 4, though not necessarily non-negative, obey the other fundamental rules of probability theory, i.e. the addition and multiplication laws. Bartlett (15) has discussed the introduction of such 'negative probabilities' as aids to calculation, and has shown that they can be manipulated according to the rules of the calculus of probabilities (with suitable precautions) provided we combine them in the end to give true (non-negative) probabilities. He remarks that 'where negative probabilities have appeared spontaneously in quantum theory, it is due to the mathematical segregation of systems or states which physically only exist in combination'.

Now this relaxation will be possible in practical applications, because the phase-space distributions contain more information than is generally required for comparison with observations. For example, if we wish to calculate the way the distribution in space $\rho(q;t)$ of a wave-packet varies with time, we may use the method of § 10, because $\rho(q;t) = \int F(p,q;t)\,dp = \psi(q;t)\,\psi^*(q;t)$ will never be negative, even if $F(p,q;t)$ can be negative. Similarly, transition-probabilities calculated by the method outlined in the same paragraph will always be non-negative, whether F takes negative values or not. Finally, we may use the methods of §§ 12–14 to calculate the phase-space distributions of members of an assembly even if the phase-space distribution for the whole assembly can be negative.

We conclude that in applications of the theory, we need not be concerned whether the phase-space distributions are true probabilities, provided that the final results, expressed either as linear combinations of these distributions or as integrals over part of their range, are necessarily true, non-negative probabilities.

17. Uniqueness of the theory and possibilities of experimental verification

The statistical approach to quantum theory involves the introduction of an additional postulate on the form of the phase-space distribution, which is equivalent to a theory of functions of non-commuting observables. The choice of this postulate is not unique. Dirac (16) has given a theory of functions of non-commuting observables which differs from the one obtained in § 5 of this paper; it has the advantage of being

independent of the basic set of variables, but, as might be expected from the foregoing discussion, it leads to complex quantities for the phase-space distributions which can never be interpreted as probabilities.†

It is natural to ask therefore whether any experimental evidence is obtainable on this subject. In so far as observable results calculated by such theories are equivalent to those obtained by orthodox methods, e.g. transition probabilities, or distributions of coordinates only, this is obviously impossible. However, though the simultaneous measurement of coordinates and momenta is not possible for single particles, there is some hope that experiments on large number of particles might be devised to verify the phase-space distributions predicted by the theory. Alternatively, one might hope to verify the corresponding theory of functions of non-commuting observables if experimental evidence became available on some Hamiltonian involving products of q and p.

APPENDICES

Appendix 1. *Space-conditional averages of the momenta and the uncertainty relations*

The space-conditional moments $\overline{\overline{p^n}}$‡ are the means of p^n when q is given. They may be obtained either from expression (4·14) for $F(p,q)$

$$\begin{aligned}\rho(q)\,\overline{\overline{p^n}} &= \int p^n F(p,q)\,dp \\ &= \iiint p^n \phi^*(p')\,\phi(p'')\,\delta\left(p-\frac{p'+p''}{2}\right) e^{iq(p''-p')/\hbar}\,dp\,dp'\,dp'' \\ &= \iint \phi^*(p')\,\phi(p'')\left(\frac{p'+p''}{2}\right)^n e^{iq(p''-p')/\hbar}\,dp'\,dp'' \\ &= \left(\frac{\hbar}{2i}\right)^n \left\{\left(\frac{\partial}{\partial q_1}-\frac{\partial}{\partial q_2}\right)^{(n)} \psi(q_1)\,\psi^*(q_2)\right\}_{q_1=q_2=q}, \end{aligned} \tag{A 1·1}$$

where $\rho(q) = \int F(p,q)\,dp = \psi(q)\,\psi^*(q)$, or from the characteristic function $M(\tau\,|\,q)$ of p conditional in q (see Bartlett (**17**)) which is seen, from (3·7), to be

$$M(\tau\,|\,q) = \frac{1}{\rho}\int F(p,q)\,e^{i\tau p}\,dp = \psi^*(q-\tfrac{1}{2}\hbar\tau)\,\psi(q+\tfrac{1}{2}\hbar\tau)/\psi^*(q)\,\psi(q). \tag{A 1·2}$$

On writing

$$\psi(q) = \rho^{\frac{1}{2}}(q)\,e^{iS(q)/\hbar} \tag{A 1·3}$$

the logarithm of $M(\tau\,|\,q)$ or 'cumulant' function (Kendall (**18**))

$$K(\tau\,|\,q) = \log M(\tau\,|\,q) = \tfrac{1}{2}\log\rho(q+\tfrac{1}{2}\hbar\tau)\,\rho(q-\tfrac{1}{2}\hbar\tau) - \log\rho(q) + \frac{i}{\hbar}[S(q+\tfrac{1}{2}\hbar\tau) - S(q-\tfrac{1}{2}\hbar\tau)] \tag{A 1·4}$$

leads to a simple expression for the 'cumulants' $\overline{\kappa}_n(q)$ (coefficients of $(i\tau)^n/n!$ in the Taylor expansion of K)

$$\overline{\kappa}_{2n+1}(q) = \left(\frac{\hbar}{2i}\right)^{2n}\left(\frac{\partial}{\partial q}\right)^{2n+1} S(q), \quad \overline{\kappa}_n(q) = \left(\frac{\hbar}{2i}\right)^{2n}\left(\frac{\partial}{\partial q}\right)^{2n} \log\rho(q). \tag{A 1·5}$$

† *Note added in proof.* Reference should also be made to a recent paper by Feynman (**26**) giving an alternative approach.

‡ The double bar $\overline{\overline{\ }}$ denotes a conditional moment, while a single bar $\overline{\ }$ denotes a mean over the distribution of both p and q.

The $\bar{\kappa}_n$ bear simple relations to the moments $\overline{\overline{p^n}}$ (Kendall (18)). In particular, the first moment by both methods is

$$\bar{\kappa}_1(q) = \bar{\bar{p}}(q) = \frac{\partial S}{\partial q}, \tag{A 1·6}$$

leading to the interpretation of the argument of the wave-function $\psi(q)$ as the potential $S(q)$ of the space-conditional mean $\bar{\bar{p}}(q)$. The conditional mean-square deviation is†

$$\bar{\kappa}_2(q) = \sigma^2_{p|q} = \overline{\overline{p^2}} - (\bar{\bar{p}})^2 = -\frac{\hbar^2}{4}\frac{\partial^2 \log \rho}{\partial q^2}. \tag{A 1·7}$$

We note also that the asymmetry of a distribution depends only on its odd cumulants; hence the asymmetry of the conditional distribution of p depends entirely on $S(q)$.

Formulae (A 1·6) and (A 1·7) lead directly to Heisenberg's inequality for the mean-square deviations of p and q. Let α, β be any two random variables with zero means. We have the well-known Schwarz inequality

$$|(\overline{\alpha^2\beta^2})| = \sigma_\alpha \sigma_\beta \geqslant |\overline{\alpha\beta}|. \tag{A 1·8}$$

Now take $\alpha = \bar{\bar{p}}(q)$, where we suppose $\bar{\bar{p}}$ to become random when we allow q to vary; take also $\beta = q$. Then from (A 1·8) above, and assuming (as can be done without loss of generality) that $\bar{p} = \bar{q} = 0$, we obtain

$$\sigma_q \sigma(\bar{\bar{p}}) \geqslant \left| \int q \bar{\bar{p}} \rho(q)\, dq \right| = |\overline{qp}|. \tag{A 1·9}$$

Take now $\quad \alpha = \partial \log \rho / \partial q, \quad \bar{\alpha} = \int \frac{\partial \log \rho}{\partial q} \rho\, dq = 0,$

$$\overline{\alpha^2} = \int \left(\frac{\partial \log \rho}{\partial q}\right)^2 \rho\, dq = -\int \frac{\partial^2 \log \rho}{\partial q^2} \rho\, dq = \frac{4}{\hbar^2} \int \sigma^2_{p|q} \rho\, dq,$$

$$\overline{\alpha q} = \int q \frac{\partial \log \rho}{\partial q} \rho\, dq = -1.$$

Hence, from (A 1·8), $$\sigma_q^2 \int \sigma^2_{p|q} \rho\, dq \geqslant \tfrac{1}{4}\hbar^2. \tag{A 1·10}$$

Since $$\sigma_p^2 = \int [\sigma^2_{p|q} + (\bar{\bar{p}})^2] \rho\, dq, \tag{A 1·11}$$

the sum of the two inequalities (A 1·9) and (A 1·10) gives Heisenberg's inequality

$$\sigma_p^2 \sigma_q^2 \geqslant (\overline{pq})^2 + \tfrac{1}{4}\hbar^2. \tag{A 1·12}$$

This derivation of Heisenberg's inequality was pointed out to the author by Prof. M. S. Bartlett.

† The fact that $\sigma^2_{p|q}$ can be negative according to (A 1·7) results from the possibility of the formal expression for $F(p, q)$ being negative in certain states. The restrictions thus imposed on the interpretation of $F(p, q)$ as a probability are discussed in § 15.

Appendix 2. *Orthogonality and completeness of the phase-space eigenfunctions for canonically conjugate variables*

The orthogonality relations of the phase-space eigenfunctions for canonically conjugate variables can be seen quite simply. We have, from (4·11),

$$\begin{aligned}\iint f_{lk}(p,q)\,f^*_{l'k'}(p,q)\,dp\,dq &= (2\pi)^{-2}\iiiint u^*_l(q-\tfrac{1}{2}\hbar\tau)\,u_k(q+\tfrac{1}{2}\hbar\tau)\,u_{l'}(q-\tfrac{1}{2}\hbar\tau')\,u^*_{k'}(q+\tfrac{1}{2}\hbar\tau')\,e^{-i(\tau-\tau')p}\,d\tau\,d\tau'\,dp\,dq\\ &= h^{-1}\iint u^*_l(x)\,u_{l'}(x)\,u_k(y)\,u^*_{k'}(y)\,dx\,dy = h^{-1}\delta_{ll'}\,\delta_{kk'} \qquad (\text{A } 2{\cdot}1)\end{aligned}$$

(the second line following from the change of variables $x = q-\frac{1}{2}\hbar\tau$, $y = q+\frac{1}{2}\hbar\tau$), and

$$\begin{aligned}\iint f_{lk}(p,q)\,dp\,dq &= (2\pi)^{-1}\iiint u^*_l(q-\tfrac{1}{2}\hbar\tau)\,u_k(q+\tfrac{1}{2}\hbar\tau)\,e^{-i\tau p}\,d\tau\,dp\,dq\\ &= \int u^*_l(q)\,u_k(q)\,dq = \delta_{lk}. \qquad (\text{A } 2{\cdot}2)\end{aligned}$$

The completeness relations follow from the corresponding relation for the $u_i(q)$

$$\begin{aligned}\sum_{l,k} f^*_{lk}(p,q)\,f_{lk}(p',q') &= (2\pi)^{-2}\iint \sum_{l,k} u^*_l(q-\tfrac{1}{2}\hbar\tau)\,u_l(q'-\tfrac{1}{2}\hbar\tau')\,u_k(q+\tfrac{1}{2}\hbar\tau)\,u^*_k(q'+\tfrac{1}{2}\hbar\tau')\,e^{i(\tau'p'-\tau p)}\,d\tau\,d\tau'\\ &= (2\pi)^{-2}\iint \delta[(q-q)'+\tfrac{1}{2}\hbar(\tau-\tau')]\,\delta[(q-q')-\tfrac{1}{2}\hbar(\tau-\tau')]\,e^{i(\tau'p'-\tau p)}\,d\tau\,d\tau'\\ &= h^{-1}\delta(q-q')\,\delta(p-p'), \qquad (\text{A } 2{\cdot}3)\end{aligned}$$

$$\begin{aligned}\sum_l f_{ll}(p,q) &= (2\pi)^{-1}\int \sum_l u^*_l(q-\tfrac{1}{2}\hbar\tau)\,u_l(q+\tfrac{1}{2}\hbar\tau)\,e^{-i\tau p}\,d\tau\\ &= (2\pi)^{-1}\int \delta(\hbar\tau)\,e^{-i\tau p}\,d\tau = h^{-1}. \qquad (\text{A } 2{\cdot}4)\end{aligned}$$

Appendix 3. *Operators corresponding to functions of canonically conjugate variables*

The proof of (5·5) follows from expression (3·10) for the phase-space distribution $F(p,q)$.

$$\begin{aligned}\overline{G(p,q)} &= \iint G(p,q)\,F(p,q)\,dp\,dq\\ &= h^{-1}\iint G(p,q)\,\{e^{\frac{1}{2}(\hbar/i)\partial^2/\partial p\partial q}\,[\psi^*(q)\,\phi(p)\,e^{ipq/\hbar}]\}\,dp\,dq\\ &= h^{-1}\iint \{e^{\frac{1}{2}(\hbar/i)\partial^2/\partial p\partial q}\,G(p,q)\}\,\psi^*(q)\,\phi(p)\,e^{ipq/\hbar}\,dp\,dq\\ &= \int \psi^*(q)\,\{e^{\frac{1}{2}(\hbar/i)\partial^2/\partial p\partial q}\,\mathbf{G}_0(\mathbf{q},\mathbf{p})\}\,\psi(q)\,dq\\ &= \int \psi^*(q)\,\mathbf{G}\psi(q)\,dq, \qquad (\text{A } 3{\cdot}1)\end{aligned}$$

and hence

$$\mathbf{G} = e^{\frac{1}{2}(\hbar/i)\partial^2/\partial p\partial q}\,\mathbf{G}_0(\mathbf{q},\mathbf{p}). \qquad (\text{A } 3{\cdot}2)$$

The operator corresponding to a function

$$G(p,q) = \sum_n \mu_n(q)\, p^n \tag{A 3·3}$$

is obtained very simply from (A 1·1). We have

$$\begin{aligned}\overline{\mu_n(q)\, p^n} &= \int \mu_n(q)\, \overline{\overline{p^n}}\rho(q)\, dq \\ &= \left(\frac{\hbar}{2i}\right)^n \int \mu_n(q) \left\{\left(\frac{\partial}{\partial q_1} - \frac{\partial}{\partial q_2}\right)^{(n)} \psi(q_1)\, \psi^*(q_2)\right\}_{q_1=q_2=q} dq \\ &= \int \psi^*(q) \left\{\sum_{k=0}^{n} \binom{n}{k} \mathbf{p}^k \mu_n(\mathbf{q})\, \mathbf{p}^{n-k}\right\} \psi(q)\, dq \end{aligned} \tag{A 3·4}$$

and hence

$$\mathbf{G} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{k} \mathbf{p}^k \mu_n(\mathbf{q})\, \mathbf{p}^{n-k}. \tag{A 3·5}$$

This could also be derived from (A 3·2) (cf. McCoy (10)).

Appendix 4. *Transport equations and the Schrödinger equation*

The 'transport' equation of any quantity $g(p,q,t)$ is defined as the equation governing the time variation of the mean $\overline{\overline{g}}(q,t)$ at every point q (space-conditional mean). It is obtained from (7·7) or (7·8) by integrating over the momenta p and making use of the expressions in Appendix 1 for the conditional moments of p. In the case of a particle of mass m, charge e in an electromagnetic field, whose classical Hamiltonian is

$$H(p_i, q_i) = \frac{1}{2m} \sum_i \left(p_i - \frac{e}{c} A_i\right)^2 + V(q_i, t) \quad (i = 1, 2, 3) \tag{A 4·1}$$

($A_i(q_k, t)$ being the vector, $V(q_k, t)$ the scalar, potentials) integration of (7·8) and substitution of $\overline{\overline{p_i}} = \partial S/\partial q_i$ from (A 1·6) lead to the continuity equation

$$\frac{\partial \rho}{\partial t} + \sum_i \frac{\partial}{\partial q_i}\left(\rho \frac{\partial S}{\partial q_i}\right) = 0, \tag{A 4·2}$$

where $\rho(q_i)$ is the distribution function of the coordinates. Multiplying (7·8) by p_k and integrating gives the transport equation for $\overline{\overline{p_k}}$

$$\frac{\partial}{\partial t}(\rho \overline{\overline{p_k}}) + \sum_i \frac{\partial}{\partial q_i}\left(\rho \overline{\overline{p_k \frac{\partial H}{\partial q_i}}}\right) + \rho \overline{\overline{\frac{\partial H}{\partial q_k}}} = 0. \tag{A 4·3}$$

Substituting in the above from (A 1·6) and (A 1·7), and combining with (A 4·2), we find

$$\frac{\partial}{\partial q_k}\left\{\frac{\partial S}{\partial t} + \overline{\overline{H}} - \frac{\hbar^2}{8m\rho} \sum_i \frac{\partial^2 \rho}{\partial q_i^2}\right\} = 0 \quad (k = 1, 2, 3). \tag{A 4·4}$$

Hence the quantum-mechanical equivalent of the classical Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + \overline{\overline{H}} = \frac{\hbar^2}{8m\rho} \nabla^2 \rho. \tag{A 4·5}$$

Substituting $\rho = \psi\psi^*$ and $S = \hbar/2i \log(\psi/\psi^*)$ and adding and subtracting (A 4·2) and (A 4·5) we find the Schrödinger equation of a charged particle in the field

$$\frac{1}{2m} \sum_i \left(\frac{\hbar}{i} \frac{\partial}{\partial q_i} - \frac{e}{c} A_i\right)^2 \psi + V\psi = i\hbar \frac{\partial \psi}{\partial t}. \tag{A 4·6}$$

Appendix 5. *Operators corresponding to functions of linear combinations of the basic variables*

According to (5·2) and (5·3), the operator corresponding to $G(ar+bs)$, where a and b are constants, is

$$\mathbf{G} = \iint e^{i(\tau\mathbf{r}+\theta\mathbf{s})}\,d\tau\,d\theta \iint G(ar+bs)\,e^{-i(\tau r+\theta s)}\,dr\,ds. \tag{A 5·1}$$

Changing to the variables

$$\xi = ar+bs, \quad \eta = ar-bs, \quad \lambda = \frac{\tau}{2a}+\frac{\theta}{2b}, \quad \mu = \frac{\tau}{2a}-\frac{\theta}{2b}, \tag{A 5·2}$$

we find

$$\mathbf{G} = \iint e^{i[(\lambda+\mu)a\mathbf{r}+(\lambda-\mu)b\mathbf{s}]}\,d\lambda\,d\mu \iint G(\xi)\,e^{-i(\lambda\xi+\mu\eta)}\,d\xi\,d\eta$$

$$= \int e^{i\lambda(a\mathbf{r}+b\mathbf{s})}\,d\lambda \int G(\xi)\,e^{-i\lambda\xi}\,d\xi = G(a\mathbf{r}+b\mathbf{s}). \tag{A 5·3}$$

I should like to acknowledge my indebtedness to Profs. P. A. M. Dirac, H. Jeffreys and the late R. H. Fowler for their criticisms, suggestions and encouragement in carrying out this work, and my gratitude to Prof. M. S. Bartlett for many invaluable discussions and the communication of his various results referred to in the text. M. J. Bass and Dr H. J. Groenewold have studied the same subject independently (cf. Bass (19) (20), Groenewold (21)), and I have benefited from discussions and correspondence with them. The papers of Powell (22), Stueckelberg (23). Dedebant (24) and Reichenbach's book (25) also have a bearing on the questions discussed in the present paper (I am indebted to Prof. Bartlett for these last references).

SUMMARY

An attempt is made to interpret quantum mechanics as a statistical theory, or more exactly as a form of non-deterministic statistical dynamics. The paper falls into three parts. In the first, the distribution functions of the complete set of dynamical variables specifying a mechanical system (phase-space distributions), which are fundamental in any form of statistical dynamics, are expressed in terms of the wave vectors of quantum theory. This is shown to be equivalent to specifying a theory of functions of non-commuting operators, and may hence be considered as an interpretation of *quantum kinematics*. In the second part, the laws governing the transformation with time of these phase-space distributions are derived from the equations of motion of *quantum dynamics* and found to be of the required form for a dynamical stochastic process. It is shown that these phase-space transformation equations can be used as an alternative to the Schrödinger equation in the solution of quantum mechanical problems, such as the evolution with time of wave packets, collision problems and the calculation of transition probabilities in perturbed systems; an approximation method is derived for this purpose. The third part, *quantum statistics*, deals with the phase-space distribution of members of large assemblies, with a view to applications of quantum mechanics to kinetic theories of matter. Finally, the limitations of the theory, its uniqueness and the possibilities of experimental verification are discussed.

REFERENCES

(1) DIRAC, P. A. M. *Commun. Dublin Inst. Adv. Stud.* A, 1 (1943), 11.
(2) WHITTAKER, E. T. *Proc. Phys. Soc., London*, 55 (1943), 459–70.
(3) DIRAC, P. A. M. *The principles of quantum mechanics* (2nd ed., Cambridge, 1935), p. 48.
(4) JEFFREYS, H. *Phil. Mag.* 32 (1941), 195.
(5) HEISENBERG, W. *The physical principles of the quantum theory* (Cambridge, 1930), p. 34.
(6) CRAMÉR, H. *Random variables and probability distributions* (Cambridge, 1937), p. 24.
(7) KERMACK, W. O. and MCCREA, W. H. *Proc. Edinburgh Math. Soc.* 2, series 2 (1931), 224.
(8) WIGNER, E. *Phys. Rev.* 40 (1932), 749.
(9) WEYL, H. *The theory of groups and quantum mechanics* (London, 1931), p. 274.
(10) MCCOY, N. H. *Proc. Nat. Acad. Sci., Wash.*, 18 (1932), 674.
(11) HOSTINSKY, M. B. *Méthodes générales du calcul des probabilités* (Paris, 1931), p. 47.
(12) JEFFREYS, H. *Phil. Mag.* 33 (1942), 815.
(13) DIRAC, P. A. M. *The principles of quantum mechanics* (2nd ed., Cambridge, 1935), pp. 140–2.
(14) SCHRÖDINGER, E. *Statistical thermodynamics* (Cambridge, 1946), p. 43.
(15) BARTLETT, M. S. *Proc. Cambridge Phil. Soc.* 41 (1944), 71.
(16) DIRAC, P. A. M. *Rev. Mod. Phys.* 17 (1945), 195–9.
(17) BARTLETT, M. S. *J. London Math. Soc.* 13 (1938), 62–7.
(18) KENDALL, M. G. *The advanced theory of statistics*, 1 (London, 1943), 60, 90.
(19) BASS, J. *Comptes Rendus*, 221 (1945), 46–9.
(20) BASS, J. *Rev. Scient.* 3240 (1945), 11–18.
(21) GROENEWOLD, H. J. *Physica*, 12 (1946), 405–60.
(22) POWELL, F. C. *Proc. Cambridge Phil. Soc.* 41 (1945), 57.
(23) STUECKELBERG, E. C. G. *Helv. Phys. Acta*, 18 (1945), 195.
(24) DEDEBANT, G. *Portugaliae Physica*, 2 (1946), 149.
(25) REICHENBACH, H. *Philosophic foundations of quantum mechanics* (Berkeley and Los Angeles, 1946).
(26) FEYNMAN, R. P. *Rev. Mod. Phys.* 20 (1948), 367–87.

DEPARTMENT OF MATHEMATICAL PHYSICS
THE QUEEN'S UNIVERSITY, BELFAST