



Argonne
NATIONAL
LABORATORY

... for a brighter future



U.S. Department
of Energy

UChicago ►
Argonne_{LLC}



Office of
Science
U.S. DEPARTMENT OF ENERGY

A U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC

Cosmas Zachos

A scrapbook story of Phase-Space Quantization

Characters

20s, 30s: Weyl, von Neumann,
Dirac, Wigner
(Terletski, Blokhintzev,
Husimi)

40s: Groenewold, Moyal,
(Yvon, Bass, Bartlett)

50s, 60s: Baker, Takabayasi,
Fairlie, (Kundt, Bopp,
Kubo, Sudarshan)

Istanbul, Sep 2014

THE WEYL CORRESPONDENCE BRIDGE

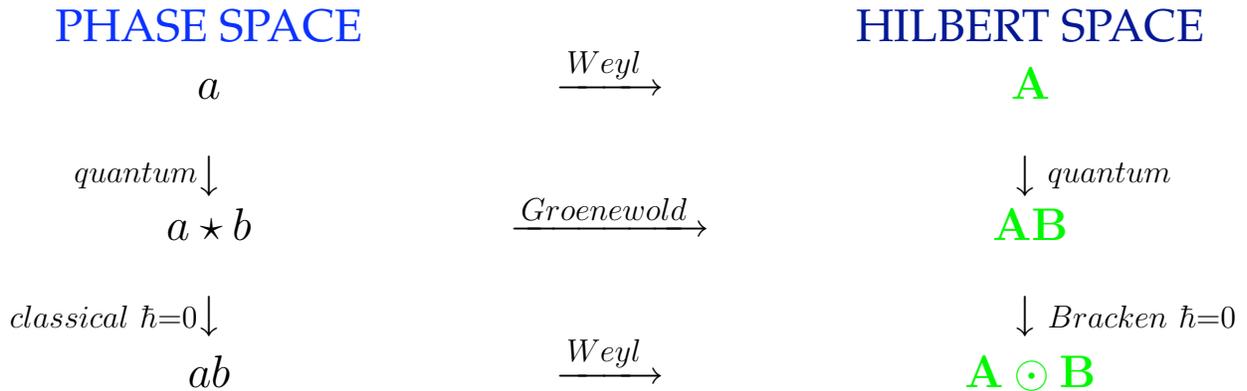
Weyl's correspondence map, by itself, merely provides **a change of representation between phase space and Hilbert space**

↷ Mutual language to contrast classical to quantum mechanics on common footing.

$$\mathbf{A}(\mathbf{x}, \mathbf{p}) = \frac{1}{(2\pi)^2} \int d\tau d\sigma dx dp a(x, p) \exp(i\tau(\mathbf{p} - p) + i\sigma(\mathbf{x} - x)),$$

Inverse map (Wigner):

$$a(x, p) = \frac{1}{2\pi} \int dy e^{-iyp} \left\langle x + \frac{\hbar}{2}y \left| \mathbf{A}(\mathbf{x}, \mathbf{p}) \right| x - \frac{\hbar}{2}y \right\rangle .$$



↷ A plethora of choice-of-ordering quantum mechanics problems reduce to purely \star -product algebraic ones: varied deformations (ordering choices) can be surveyed systematically in phase space. (Curtright & Zachos, New J Phys 4 (2002) 83.1-83.16 [hep-th/0205063])

0.23 Synopses of Selected Papers

The decisive contributors to the development of the formulation are Hermann Weyl (1885–1955), Eugene Wigner (1902–1995), Hilbrand Groenewold (1910–1996), and Jose Moyal (1910–1998). The bulk of the theory is implicit in Groenewold’s and Moyal’s seminal papers.

But confidence in the autonomy of the formulation accreted slowly and fitfully. As a result, an appraisal of critical milestones cannot avoid subjectivity. Nevertheless, here we provide summaries of a few papers that we believe remedied confusion about the logical structure of the formulation.

H Weyl (1927)^{Wey27} introduces the correspondence of “Weyl-ordered” operators to phase-space (c-number) kernel functions. The correspondence is based on Weyl’s formulation of the Heisenberg group, appreciated through a discrete QM application of Sylvester’s (1883)^{Syl82} clock and shift matrices. The correspondence is proposed as a general quantization prescription, unsuccessfully, since it fails, e.g., with angular momentum squared.

J von Neumann (1931)^{Neu31}, expatiates on a Fourier transform version of the \star -product, in a technical aside off an analysis of the uniqueness of Schrödinger’s representation, based on Weyl’s Heisenberg group formulation. This then effectively promotes Weyl’s correspondence rule to full isomorphism between Weyl-ordered operator multiplication and \star -convolution of kernel functions. Nevertheless, this result is not properly appreciated in von Neumann’s celebrated own book on the Foundations of QM.

E Wigner (1932)^{Wig32}, the author’s first paper in English, introduces the eponymous phase-space distribution function controlling quantum mechanical diffusive flow in phase space. It notes the negative values, and specifies the time evolution of this function and applies it to quantum statistical mechanics. (Actually, Dirac (1930)^{Dir30} has already considered a formally identical construct, and an implicit Weyl correspondence, for the approximate electron density in a multi-electron Thomas–Fermi atom; but, interpreting negative values as a failure of that semiclassical approximation, he crucially hesitates about the full quantum object.)

H Groenewold (1946)^{Gro46}, a seminal but inadequately appreciated paper, is based on Groenewold’s thesis work. It achieves full understanding of the Weyl correspondence as an invertible transform, rather than as a consistent quantization rule. It articulates and recognizes the WF as the phase-space (Weyl) kernel of the density matrix. It reinvents and streamlines von Neumann’s construct into the standard \star -product, in a systematic exploration of the isomorphism between Weyl-ordered operator products and their kernel function compositions. It thus demonstrates how Poisson Brackets contrast crucially

to quantum commutators—“Groenewold’s Theorem”. By way of illustration, it further works out the harmonic oscillator WF.

J Moyal (1949)^{Moy49} enunciates a grand synthesis: It establishes an independent formulation of quantum mechanics in phase space. It systematically studies all expectation values of Weyl-ordered operators, and identifies the Fourier transform of their moment-generating function (their characteristic function) with the Wigner Function. It further interprets the subtlety of the “negative probability” formalism and reconciles it with the uncertainty principle and the diffusion of the probability fluid. Not least, it recasts the time evolution of the Wigner Function through a deformation of the Poisson Bracket into the Moyal Bracket (the commutator of \star -products, i.e., the Wigner transform of the Heisenberg commutator), and thus opens up the way for a systematic study of the semi-classical limit. Before publication, Dirac contrasts this work favorably to his own ideas on functional integration, in Bohr’s Festschrift^{Dir45}, despite private reservations and lengthy arguments with Moyal. Various subsequent scattered observations of French investigators on the statistical approach^{Yv46}, as well as Moyal’s, are collected in J Bass (1948)^{Bas48}, which further stretches to hydrodynamics. Earlier Soviet efforts include^{Ter37,Blo40}.

M Bartlett and J Moyal (1949)^{BM49} applies this language to calculate propagators and transition probabilities for oscillators perturbed by time-dependent potentials.

T Takabayasi (1954)^{Tak54} investigates the fundamental projective normalization condition for pure state Wigner functions, and exploits Groenewold’s link to the conventional density matrix formulation. It further illuminates the diffusion of wavepackets.

G Baker (1958)^{Bak58} (Baker’s thesis paper) envisions the logical autonomy of the formulation, sustained by the projective normalization condition as a basic postulate. It resolves measurement subtleties in the correspondence principle and appreciates the significance of the anticommutator of the \star -product as well, thus shifting emphasis to the \star -product itself, over and above its commutator.

D Fairlie (1964)^{Fai64} (also see refs^{Kun67,Coh76,Dah83,Bas48}) explores the time-independent counterpart to Moyal’s evolution equation, which involves the \star -product, beyond mere Moyal Bracket equations, and derives (instead of postulating) the projective orthonormality conditions for the resulting Wigner functions. These now allow for a unique and full solution of the quantum system, in principle (without any reference to the conventional Hilbert-space formulation). Autonomy of the formulation is fully recognized.

R Kubo (1964)^{Kub64} elegantly reviews, in modern notation, the representation change between Hilbert space and phase space—although in ostensible ignorance of Weyl’s and Groenewold’s specific papers. It applies the phase-space picture to the description of electrons in a uniform magnetic field, initiating gauge-invariant formulations and pioneering

“noncommutative geometry” applications to diamagnetism and the Hall effect.

N Cartwright (1976)^{Car76} notes that the WF smoothed by a phase-space Gaussian (i.e., Weierstrass transformed) as wide or wider than the minimum uncertainty packet is positive-semidefinite. Actually, this convolution result goes further back to at least de Bruijn (1967)^{deB67} and Iagolnitzer (1969)^{Iag69}, if not Husimi (1940)^{Hus40}.

M Berry (1977)^{Ber77} elucidates the subtleties of the semiclassical limit, ergodicity, integrability, and the singularity structure of Wigner function evolution. Complementary results are featured in Voros (1976-78)^{Vo78}.

F Bayen, M Flato, C Fronsdal, A Lichnerowicz, and D Sternheimer (1978)^{BFF78} analyzes systematically the deformation structure and the uniqueness of the formulation, with special emphasis on spectral theory, and consolidates it mathematically. (Also see Berezin^{Ber75}.) It provides explicit illustrative solutions to standard problems and utilizes influential technical tools, such as the \star -exponential (already known in^{Imr67, GLS68}).

A Royer (1977)^{Roy77} interprets WFs as the expectation value of the operators effecting reflections in phase space. (Further see refs^{Kub64, Gro76, BV94}.)

G García-Calderón and M Moshinsky (1980)^{GM80} implements the transition from Hilbert space to phase space to extend classical propagators and canonical transformations to quantum ones in phase space. (The most conclusive work to date is ref^{BCW02}. Further see^{HKN88, Hie82, DKM88, CFZ98, DV97, GR94, Hak99, KL99, DP01}.)

J Dahl and M Springborg (1982)^{DS82} initiates a thorough treatment of the hydrogen and other simple atoms in phase space, albeit not from first principles—the WFs are evaluated in terms of Schrödinger wave-functions.

M De Wilde and P Lecomte (1983)^{deW83} consolidates the deformation theory of \star -products and MBs on general real symplectic manifolds, analyzes their cohomology structure, and confirms the absence of obstructions.

M Hillery, R O’Connell, M Scully, and E Wigner (1984)^{HOS84} has done yeoman service to the physics community as the classic introduction to phase-space quantization and the Wigner function.

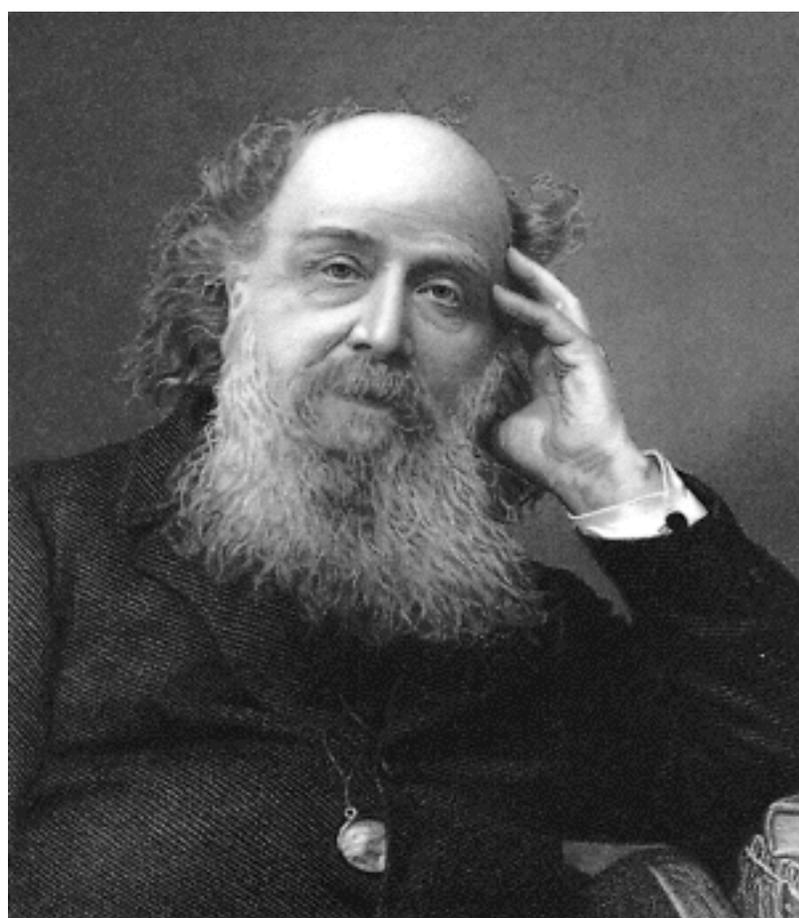
Y Kim and E Wigner (1990)^{KW90} is a classic pedagogical discussion of the spread of wavepackets in phase space, uncertainty-preserving transformations, coherent and squeezed states.

B Fedosov (1994)^{Fed94} initiates an influential geometrical construction of the \star -product

on all symplectic manifolds.

T Curtright, D Fairlie, and C Zachos (1998)^{CFZ98} illustrates more directly the equivalence of the time-independent \star -genvalue problem to the Hilbert space formulation, and hence its logical autonomy; formulates Darboux isospectral systems in phase space; works out the covariant transformation rule for general nonlinear canonical transformations (with reliance on the classic work of P Dirac (1933)^{Dir33}); and thus furnishes explicit solutions of practical problems on first principles, without recourse to the Hilbert space formulation. Efficient techniques for perturbation theory are based on generating functions for complete sets of Wigner functions in T Curtright, T Uematsu, and C Zachos (2001)^{CUZ01}. A self-contained derivation of the uncertainty principle in phase space is given in T Curtright and C Zachos (2001)^{CZ01}.

M Hug, C Menke, and W Schleich (1998)^{HMS98} introduce and exemplify techniques for numerical solution of \star -equations on a basis of Chebyshev polynomials. Dynamical scattering of wavepacket WFs off Gaussian barrier potentials on a similar basis is detailed in ref ^{SLC11}.



THE COLLECTED
MATHEMATICAL PAPERS

OF

JAMES JOSEPH SYLVESTER

F.R.S., D.C.L., LL.D., Sc.D.,

Honorary Fellow of St John's College, Cambridge ;

Sometime Professor at University College, London ; at the University of Virginia ;
at the Royal Military Academy, Woolwich ; at the Johns Hopkins University, Baltimore
and Savilian Professor in the University of Oxford

VOLUME III

(1870—1883)

Cambridge

At the University Press

1909

A WORD ON NONIONS.

[*Johns Hopkins University Circulars*, I. (1882), pp. 241, 242;
II. (1883), p. 46.]

IN my lectures on Multiple Algebra I showed that if u, v are two matrices of the second order, and if the determinant of the matrix $(z + yv + xu)$ be written as

$$z^2 + 2bxz + 2cyz + dx^2 + 2exy + fy^2$$

then the necessary and sufficient conditions for the equation $vu + uv = 0$ are the following, namely,

$$b = 0, \quad c = 0, \quad e = 0.$$

If to these conditions we superadd $d = 1, f = 1$, and write $uv = w$, then

$$u^2 = -1, \quad v^2 = -1, \quad w^2 = -1, \quad uv = -vu = w, \quad vw = -wv = u, \quad wu = -uw = v;$$

and $1, u, v, w$ form a quaternion system. The conditions above stated will be satisfied if

$$\text{Det. } (z + yv + xu) = z^2 + y^2 + x^2,$$

which will obviously be the case if

$$v = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}, \quad u = \begin{vmatrix} 0 & \theta \\ \theta & 0 \end{vmatrix},$$

where $\theta = \sqrt{-1}$. For then

$$z + yv + xu = \begin{vmatrix} z & y + x\theta \\ -y + x\theta & z \end{vmatrix}.$$

Hence the matrices

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} \begin{vmatrix} 0 & \theta \\ \theta & 0 \end{vmatrix} \begin{vmatrix} -\theta & 0 \\ 0 & \theta \end{vmatrix}$$

construed as complex quantities are a linear transformation of the ordinary

quaternion system $1, i, j, k$; that is to say, if we form the multiplication table

	λ	μ	ν	τ
λ	λ	μ	0	0
μ	0	0	λ	μ
ν	ν	τ	0	0
τ	0	0	ν	τ

$$\begin{aligned} \lambda + \tau &= 1 & -\mu + \nu &= i \\ -\theta\lambda + \theta\tau &= k & \theta\mu + \theta\nu &= j. \end{aligned}$$

Since u, v contain between them 8 letters subject to the satisfaction of 5 conditions, the most general values of λ, μ, ν, τ ought to contain 3 arbitrary constants; but it is well-known that any particular (i, j, k) system may be superseded by a $\lambda(i', j', k')$ system, where i', j', k' are orthogonally related linear functions of i, j, k ; and as this substitution introduces just 3 arbitrary constants, we may, by aid of it, pass from the system of matrices above given, to the most general form. The general expression for the matrices containing 3 arbitrary constants may also be found directly by the method given in my lectures, which will be reproduced in the memoir on Multiple Algebra in the *Mathematical Journal*. What goes before is by way of introduction to the *word* on Nonions which follows.

Just as the necessary and sufficient condition that u, v , two matrices of the second order, may satisfy the equations $vu = -uv, u^2 = 1, v^2 = 1$, is that the determinant to $z + yv + xu$ may be $z^2 + y^2 + x^2$, so I have proved that the necessary and sufficient condition, in order that we may have $vu = \rho uv, u^3 = 1, v^3 = 1$ (u, v being matrices of the third order, and ρ an imaginary cube root of unity) is that the determinant to $z + yu + xv$ may be $z^3 + y^3 + x^3$; but if we make

$$u = \begin{vmatrix} 0 & 0 & 1 \\ \rho & 0 & 0 \\ 0 & \rho^2 & 0 \end{vmatrix}, \quad v = \begin{vmatrix} 0 & 0 & 1 \\ \rho^2 & 0 & 0 \\ 0 & \rho & 0 \end{vmatrix},$$

$$\text{then } z + yu + xv = \begin{vmatrix} z & 0 & y+x \\ \rho y + \rho^2 x & z & 0 \\ 0 & \rho^2 y + \rho x & z \end{vmatrix}$$

of which the determinant is

$$z^3 + (y+x)(\rho y + \rho^2 x)(\rho^2 y + \rho x) = z^3 + y^3 + x^3.$$

Hence there will be a system of Nonions (precisely analogous to the known

system of quaternions) represented by the 9 matrices u^2

		1
u		v
	uv	v^2
u^2v		uv^2
		u^2v^2

and just as in the preceding case the 8 terms $\pm 1, \pm u, \pm v, \pm uv$ form a closed group, so here the 27 terms obtained by multiplying each of the above 9 by $1, \rho, \rho^2$ will form a closed group. The values of the 9 matrices will easily be found to be

$$\begin{array}{c} \left| \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right| \\ \left| \begin{array}{ccc} 0 & 0 & 1 \\ \rho & 0 & 0 \\ 0 & \rho^2 & 0 \end{array} \right| \quad \left| \begin{array}{ccc} 0 & 0 & 1 \\ \rho^2 & 0 & 0 \\ 0 & \rho & 0 \end{array} \right| \\ \left| \begin{array}{ccc} 0 & \rho^2 & 0 \\ 0 & 0 & \rho \\ 1 & 0 & 0 \end{array} \right| \quad \left| \begin{array}{ccc} 0 & \rho & 0 \\ 0 & 0 & \rho \\ \rho & 0 & 0 \end{array} \right| \quad \left| \begin{array}{ccc} 0 & \rho & 0 \\ 0 & 0 & \rho^2 \\ 1 & 0 & 0 \end{array} \right| \\ \left| \begin{array}{ccc} \rho & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{array} \right| \quad \left| \begin{array}{ccc} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & \rho \end{array} \right| \\ \left| \begin{array}{ccc} 0 & 0 & \rho \\ \rho & 0 & 0 \\ 0 & \rho & 0 \end{array} \right| \end{array}$$

These forms can be derived from an algebra given by Mr Charles S. Peirce (*Logic of Relatives*, 1870).

I will only stay to observe that as the condition of the Determinant to $z + uy + vx$ (which for general values of u, v is a general cubic with the coefficient of z^3 unity) assuming the form $z^3 + y^3 + x^3$, implies the satisfaction of 9 conditions, and as u, v between them contain 18 constants, the most general form of a system of Nonions must contain $18 - 9$, or 9 arbitrary constants; but how these can be obtained from the particular form of the system above given, remains open for further examination.

[*Note.* For the remark made above] "These forms can be derived from an algebra given by Mr Charles S. Peirce (*Logic of Relatives*, 1870)," read "Mr C. S. Peirce informs me that these forms can be derived from his *Logic of Relatives*, 1870." I know nothing whatever of the fact of my own personal knowledge*. I have not read the paper referred to, and am not

* I have also a great repugnance to being made to speak of Algebras in the plural; I would as lief acknowledge a plurality of Gods as of Algebras.

14.

ON QUATERNIONS, NONIONS, SEDENIONS, ETC.

[*Johns Hopkins University Circulars*, III. (1884), pp. 7—9.]

(1) SUPPOSE that m and n are two matrices of the second order.

Then if we call the determinant of the matrix $x + my + nz$,

$$x^2 + 2bxy + 2caz + dy^2 + 2eyz + fz^2,$$

the necessary and sufficient conditions for the subsistence of the equation $nm = -mn$ is that $b = 0, c = 0, e = 0$, and if we superadd the equations $m^2 + 1 = 0, n^2 + 1 = 0$, then $d = 1$ and $f = 1$, or in other words in order to satisfy the equations $mn = -nm, m^2 = -1, n^2 = -1$, where it will of course be understood that in these (as in the equations $m^2 + 1 = 0, n^2 + 1 = 0$) 1 is the abbreviated form of the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\bar{1}$ of* the form $\begin{pmatrix} \bar{1} & 0 \\ 0 & \bar{1} \end{pmatrix}$, the necessary and sufficient condition is that the determinant of $x + my + nz$ shall be equal to $x^2 + y^2 + z^2$.

The simplest mode of satisfying this condition is to write $m = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$, $n = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, i meaning $\sqrt{-1}$, which gives $mn = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$ and $nm = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$.

It is easy to express any matrix of the second order as a linear function of 1 (meaning $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$) m, n, p , where p stands for mn .

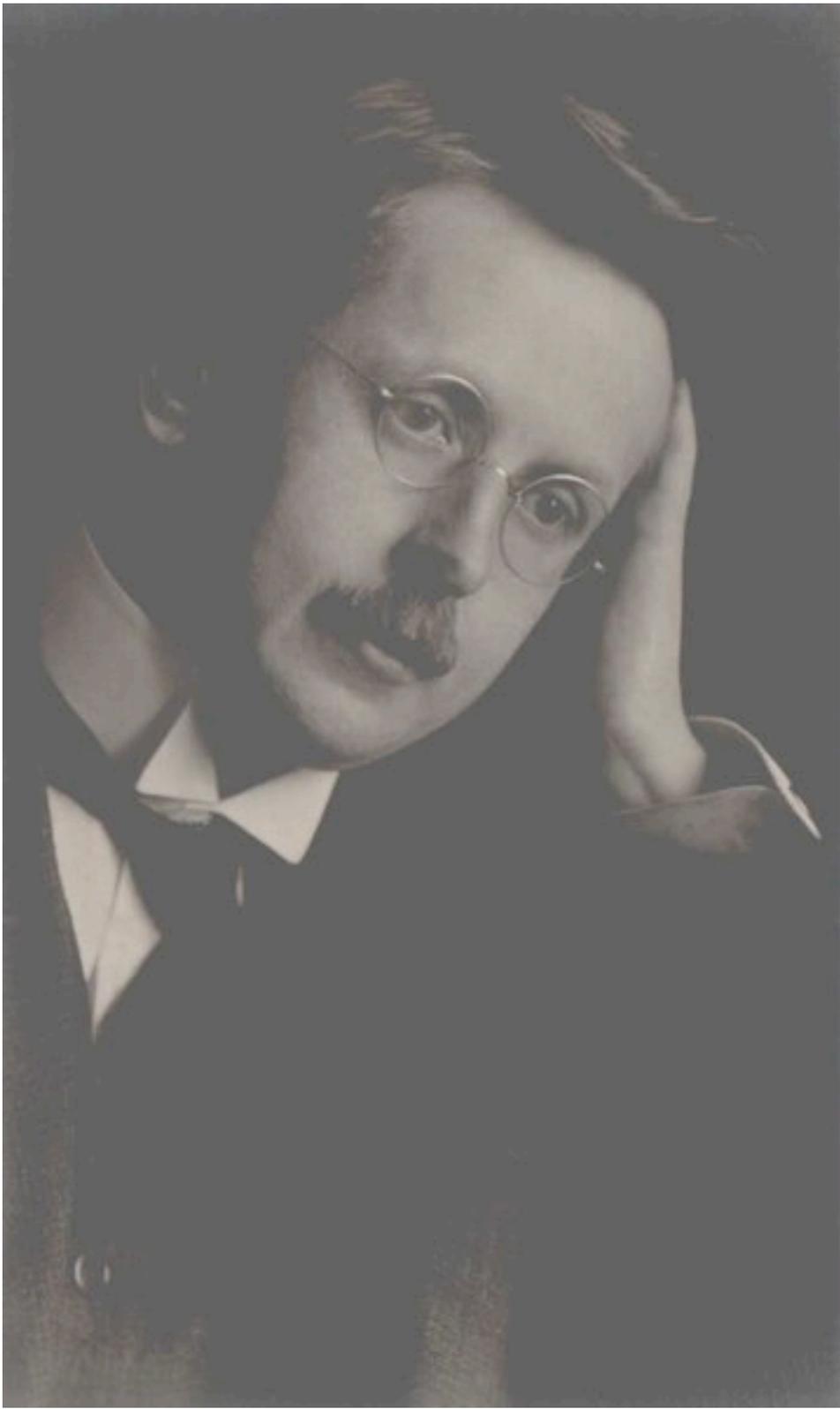
For if $\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}$ be any such matrix it is only necessary to write

$$\begin{aligned} a &= f + ig, & b &= -h - ki, \\ d &= f - ig, & c &= -h + ki, \end{aligned}$$

and then $\begin{pmatrix} a, & b \\ c, & d \end{pmatrix} = f + gm + hn + kp$.

The most general solution of the equations $MN = -NM, M^2 = N^2 = -1$, must contain three arbitrary constants, namely, the difference between the number of terms in m and n , and the number of conditions $b = 0, c = 0, e = 0, d = 1, f = 1$, which are to be satisfied.

[* $\bar{1}$ denotes -1 .]



Quantenmechanik und Gruppentheorie.

Von **H. Weyl** in Zürich.

Mit 1 Abbildung. (Eingegangen am 13. Oktober 1927.)

Einleitung und Zusammenfassung. — I. Teil. Bedeutung der Repräsentation von physikalischen Größen durch Hermitesche Formen. § 1. Mathematische Grundbegriffe, die Hermiteschen Formen betreffend. § 2. Der physikalische Begriff des reinen Falles. § 3. Die physikalische Bedeutung der repräsentierenden Hermiteschen Form. § 4. Statistik der Gemenge. — II. Teil. Kinematik als Gruppe. § 5. Über Gruppen und ihre unitären Darstellungen. § 6. Übertragung auf kontinuierliche Gruppen. § 7. Ersatz der kanonischen Variablen durch die Gruppe. Das Elektron. § 8. Übergang zu Schrödingers Wellentheorie. — III. Teil. Das dynamische Problem. § 9. Das Gesetz der zeitlichen Veränderung. Die Zeitgesamtheit. § 10. Kinetische Energie und Coulombsche Kraft in der relativistischen Quantenmechanik. — Mathematischer Anhang.

Einleitung und Zusammenfassung.

In der Quantenmechanik kann man zwei Fragen deutlich voneinander trennen: 1. Wie komme ich zu der Matrix, der Hermiteschen Form, welche eine gegebene Größe in einem seiner Konstitution nach bekannten physikalischen System repräsentiert? 2. Wenn einmal die Hermitesche Form gewonnen ist, was ist ihre physikalische Bedeutung, was für physikalische Aussagen kann ich ihr entnehmen? Auf die zweite Frage hat v. Neumann in einer kürzlich erschienenen Arbeit* eine klare und weitreichende Antwort gegeben. Aber sie spricht noch nicht alles aus, was sich darüber sagen läßt, umfaßt auch nicht alle Ansätze, die bereits in der physikalischen Literatur mit Erfolg geltend gemacht worden sind. Ich glaube, daß ich in dieser Hinsicht zu einem gewissen Abschluß gelangt bin durch die Aufstellung des Begriffs des reinen Falles**. Ein reiner Fall von Atomen z. B. liegt dann vor, wenn der betrachtete Atomschwarm den höchsten Grad von Homogenität besitzt, der sich realisieren läßt. Der monochromatische polarisierte Lichtstrahl ist ein Beispiel aus anderem Gebiet. Der reine Fall wird repräsentiert durch die Variablen der Hermiteschen Form; die Form selber gibt Aufschluß darüber, welcher Werte die durch sie repräsentierte Größe fähig ist, und mit welcher Wahrscheinlichkeit oder Häufigkeit diese Werte in irgend

* Mathematische Begründung der Quantenmechanik, Nachr. Gesellsch. d. Wissensch. Göttingen 1927, S. 1.

** Wie mir Herr v. Neumann mitteilt, ist auch er inzwischen zur Aufstellung dieses Begriffs gelangt [Zusatz bei der Korrektur].

bilden für sich eine f -parametrische Abelsche Gruppe unitärer (Vektor-) Abbildungen, ebenso die

$$V(\boldsymbol{\tau}) = e(\boldsymbol{\tau}_1 Q_1 + \boldsymbol{\tau}_2 Q_2 + \cdots + \boldsymbol{\tau}_f Q_f).$$

Hingegen ist

$$U(\boldsymbol{\sigma}) V(\boldsymbol{\tau}) U^{-1}(\boldsymbol{\sigma}) V^{-1}(\boldsymbol{\tau}) = e(\boldsymbol{\sigma}_1 \boldsymbol{\tau}_1 + \cdots + \boldsymbol{\sigma}_f \boldsymbol{\tau}_f) \cdot \mathbf{1}$$

und

$$\begin{aligned} & e(\boldsymbol{\sigma}_1 P_1 + \cdots + \boldsymbol{\sigma}_f P_f + \boldsymbol{\tau}_1 Q_1 + \cdots + \boldsymbol{\tau}_f Q_f) \\ &= e\left(\frac{1}{2} \sum_{i=1}^f \boldsymbol{\sigma}_i \boldsymbol{\tau}_i\right) V(\boldsymbol{\tau}) U(\boldsymbol{\sigma}) = e\left(-\frac{1}{2} \sum_{i=1}^f \boldsymbol{\sigma}_i \boldsymbol{\tau}_i\right) U(\boldsymbol{\sigma}) V(\boldsymbol{\tau}). \end{aligned} \quad (37)$$

§ 7. Ersatz der kanonischen Variablen durch die Gruppe. Das Elektron. Unsere Entwicklungen sind bis zu dem Punkte gedielen, wo die Verbindung mit der Quantenmechanik in die Augen springt. Liegt ein mechanisches System von f Freiheitsgraden vor, so genügen ja die Hermiteschen Matrizen, welche die kanonischen Variablen repräsentieren, gerade den Relationen (36), bis auf den Faktor $h/2\pi$, von dem noch die Rede sein wird und den wir einstweilen in die Maßeinheiten hineinstecken. Nehmen wir die Zahl der Freiheitsgrade f zunächst $= 1$ und bezeichnen in der üblichen Weise die kanonischen Variablen mit p, q , ihre repräsentierenden Formen mit P, Q , so sagt die Relation

$$i(PQ - QP) = \mathbf{1} \quad (38)$$

aus, daß die beiden durch die Matrizen iP, iQ gekennzeichneten infinitesimalen Drehungen des Strahlenkörpers vertauschbar sind. Die durch sie erzeugte Abelsche Drehungsgruppe besteht aus den Drehungen

$$U(\boldsymbol{\sigma}, \boldsymbol{\tau}) = e(P\boldsymbol{\sigma} + Q\boldsymbol{\tau}) \quad (39)$$

($\boldsymbol{\sigma}, \boldsymbol{\tau}$ reelle Parameter, die sich bei Zusammensetzung additiv verhalten). Die reelle Größe im Gruppengebiet, deren Komponenten $\xi(\boldsymbol{\sigma}, \boldsymbol{\tau})$ der Gleichung (19) oder

$$\bar{\xi}(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \xi(-\boldsymbol{\sigma}, -\boldsymbol{\tau}) \quad (40)$$

genügen, erscheint als die Hermitesche Form

$$F = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e(P\boldsymbol{\sigma} + Q\boldsymbol{\tau}) \xi(\boldsymbol{\sigma}, \boldsymbol{\tau}) d\boldsymbol{\sigma} d\boldsymbol{\tau}. \quad (41)$$

Eine physikalische Größe ist durch ihren Funktionsausdruck $f(p, q)$ in den kanonischen Variablen p, q mathematisch definiert. Es blieb ein Problem, wie ein derartiger Ausdruck auf die Matrizen zu übertragen war. Ohne weiteres klar war das nur für die Potenzen p^k, q^l und damit für Polynome. Freilich trat schon hier die Schwierigkeit auf, daß man nicht wußte, ob man einen Term wie $p^2 q$ als $P^2 Q$ oder $Q P^2$ oder $P Q P$ usw.

zu interpretieren hatte. Der Ansatz ist offenbar viel zu formal. Unsere gruppentheoretische Auffassung zeigt sogleich den rechten Weg: die Hermitesche Form (41) repräsentiert die Größe

$$f(p, q) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e(p\sigma + q\tau) \xi(\sigma, \tau) d\sigma d\tau. \quad (42)$$

Nach dem Fourierschen Integraltheorem läßt sich ja jede Funktion $f(p, q)$ in dieser Form eindeutig entwickeln, und wenn f eine reellwertige Funktion der reellen Veränderlichen p, q ist, genügt $\xi(\sigma, \tau)$ gerade der Bedingung (40). Die Integralentwicklung (42) ist nicht immer ganz wörtlich zu verstehen; das wesentliche ist nur, daß rechts eine lineare Kombination der $e(p\sigma + q\tau)$ steht, in denen σ und τ beliebige reelle Werte annehmen können. Wenn z. B. q eine zyklische Koordinate ist, die nur mod. 2π zu verstehen ist, so daß alle in Betracht kommenden Funktionen periodisch in q mit der Periode 2π sind, wird die Integration nach τ ersetzt werden müssen durch eine Summation über alle ganzen Zahlen τ ; wir haben dann den Fall einer gemischten kontinuierlich-diskreten Gruppe. Die Einschränkungen, denen $f(p, q)$ unterworfen sein muß, damit sie eine Entwicklung des Typus (42) gestattet, könnten noch Bedenken erregen. Nun wissen wir aber, daß es eigentlich gilt, $e(kf(p, q))$ so zu entwickeln (k irgend eine reelle Konstante), und in dieser Fassung läßt sich die Aufgabe nach neueren Untersuchungen von N. Wiener, Bochner und Hardy in zwingender Weise eindeutig erledigen*.

Die Übertragung auf f Freiheitsgrade liegt auf der Hand. Insbesondere sahen wir, wie aus der Forderung der Irreduzibilität im Falle der kontinuierlichen Gruppen die charakteristische kanonische Paarung entspringt. Für endliche Gruppen freilich existiert nicht ein so einheitliches Schema. Das ist im Einklang mit den physikalischen Tatsachen. Denn aus den Entwicklungen von P. Jordan** ging bereits hervor, daß beim magnetischen Elektron σ_y so gut wie σ_z als

* N. Wiener, On representations of functions by trigonometrical integrals, *Math. ZS.* **24**, 575, 1926; S. Bochner und G. H. Hardy, Note on two theorems of N. Wiener, *Journ. Lond. Math. Soc.* **1**, 240, 1926; S. Bochner, Darstellung reell variabler und analytischer Funktionen durch verallgemeinerte Fourier- und Laplaceintegrale, *Math. Ann.* **97**, 635, 1927; vgl. dazu ferner die von H. Bohr stammende Theorie der fastperiodischen Funktionen; am einfachsten bei H. Weyl, *Math. Ann.* **97**, 338, 1926.

** *ZS. f. Phys.* **44**, 21—25, 1927. Nach P. Jordan, Über die Polarisation der Lichtquanten, ebenda, S. 292, ist die Kinematik der Lichtquanten die gleiche.

die „kanonische Konjugierte“ von σ_x angesehen werden kann. Höchstens von einem Tripel, nicht von einem Paar kanonisch konjugierter Größen könnte hier vernünftigerweise die Rede sein. Bestätigen wir, daß gerade auch in diesem diskreten, dem Kontinuierlichen am meisten entgegengesetzten Falle unsere Formulierung genau das Richtige trifft! Sie lautet, um das noch einmal zusammenzufassen, so: Der kinematische Charakter eines physikalischen Systems findet seinen Ausdruck in einer irreduziblen Abelschen Drehungsgruppe, deren Substrat der Strahlenkörper der „reinen Fälle“ ist. Die reellen Größen dieses Gruppengebietes sind die physikalischen Größen; die Hermiteschen Matrizen, als welche sie vermöge der Darstellung der abstrakten Gruppe durch Drehungen erscheinen, sind die Repräsentanten der physikalischen Größen, deren Bedeutung im I. Teil auseinandergesetzt wurde.

Nun: die früher beschriebene zweidimensionale Drehungsgruppe \mathfrak{B} , welche der Vierergruppe isomorph ist, kennzeichnet, wie der Vergleich mit § 2, (12) lehrt, die Kinematik des magnetischen Elektrons. Da $n = 2$ ist, sind alle Größen nur zweier Werte fähig. Die einzigen physikalischen Größen, welche existieren, sind die mit Hilfe reeller Zahlkoeffizienten gebildeten linearen Kombinationen von $1, \sigma_x, \sigma_y, \sigma_z$. Aber das magnetische Elektron ergibt sich nicht nur als Sonderfall der Theorie, sondern die ihm eigentümliche Kinematik ist überhaupt die einzig mögliche, wenn alle Größen nur zweier Werte fähig sein sollen, wenn $n = 2$ ist. Beweis: Wir wissen schon, daß unter dieser Voraussetzung jedes Gruppenelement a außer dem Einheitselement von der Ordnung 2 ist. Die beiden Eigenwerte der korrespondierenden zweidimensionalen Matrix A sind daher entgegengesetzt gleich. Wählen wir ein bestimmtes $a \neq 1$, so können wir das zugehörige A samt einem normalen Koordinatensystem so festlegen, daß

$$A = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} \quad (43)$$

wird. Die mit A vertauschbaren Matrizen U unserer Gruppe haben notwendig die Gestalt $\begin{vmatrix} c & 0 \\ 0 & c' \end{vmatrix}$; wenn sie nicht $\simeq 1$ sind, ist $c' = -c$, U also $\simeq A$. Es gibt Gruppenelemente, deren Matrix B nicht mit A vertauschbar ist. Wir wissen, daß in der Gleichung

$$AB = \varepsilon BA$$

ε eine zweite Einheitswurzel, darum $\varepsilon = -1$ sein muß. Daraus folgt, daß B die Gestalt

$$B = \begin{vmatrix} 0 & b \\ b' & 0 \end{vmatrix} \quad (44)$$

hat. Die Zahlen b, b' sind vom absoluten Betrag 1. Wir wählen ein bestimmtes solches B , das gemäß $B^2 = \mathbf{1}$ geeicht sei: $bb' = 1$. Außerdem kann man b zu 1 machen, indem man das bisherige normale Koordinatensystem e_1, e_2 durch $e_1, b e_2$ ersetzt; (43) wird dadurch nicht angegriffen:

$$B = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}. \quad (45)$$

Jede Matrix U unserer Gruppe, welche mit A vertauschbar ist, ist $\simeq \mathbf{1}$ oder $\simeq A$. Wenn sie nicht mit A vertauschbar ist, hat sie die Form (44), und demnach ist ihre Zusammensetzung UB mit dem durch (45) gegebenen bestimmten B eine Diagonalmatrix. Als solche ist sie mit A vertauschbar, also $\simeq \mathbf{1}$ oder $\simeq A$. Das Resultat ist, daß jedes $U \simeq$ einer von den vier Matrizen $\mathbf{1}, A, B, AB$ ist. Es liegt in der Tat die Vierergruppe vor und die Darstellung \mathfrak{B} derselben.

§ 8. Übergang zu Schrödingers Wellentheorie. In ähnlicher Weise, wie soeben der Fall $n = 2$ behandelt wurde, wollen wir jetzt zeigen, daß die zweiparametrischen kontinuierlichen Gruppen nur einer irreduziblen Darstellung in unserem Sinne (außer der identischen) fähig sind. Wir erhalten jene Gruppen durch Grenzübergang aus den zweibasigen endlichen. Die irreduzible Abelsche Drehungsgruppe mit der Basis A, B habe die Dimensionszahl n . In der Kommutatorgleichung

$$AB = \varepsilon BA \quad (46)$$

ist ε eine n -te Einheitswurzel. Diese Gleichung gilt es jetzt näher zu untersuchen. Die Kommutatorzahl ε sei eine primitive m -te Einheitswurzel, d. h. ε^m sei die niederste Potenz, welche $= 1$ ist; m ist Teiler von n . Die Drehungen A, B sind von einer in n aufgehenden Ordnung: $A^n \simeq \mathbf{1}, B^n \simeq \mathbf{1}$, und die Matrizen können daher so geeicht werden, daß $A^n = B^n = \mathbf{1}$ ist. Durch geeignete Wahl des normalen Koordinatensystems sei B auf Hauptachsen gebracht; die Glieder in der Hauptdiagonale, b_i , sind lauter n -te Einheitswurzeln. Die Gleichung (46) liefert für die Koeffizienten von $A = \|a_{ik}\|$:

$$\frac{b_k}{b_i} a_{ik} = \varepsilon a_{ik}. \quad (47)$$

q zugleich eine Variable verstehend, welche den Wertbereich der physikalischen Größe q durchläuft, und $\sqrt{\xi}$. $\psi(q)$ an Stelle von x_k . $\psi(q)$ ist eine willkürliche komplexwertige Funktion, welche der Normierungsgleichung

$$\int |\psi(q)|^2 dq = 1 \quad (49)$$

unterworfen ist. Ihre Werte sind aufzufassen als die den verschiedenen Werten von q entsprechenden Komponenten eines „reinen Falles“ in demjenigen normalen Koordinatensystem, das aus den Eigenvektoren der Größe q besteht. — An Stelle der zweiten Gleichung (48) erhalten wir im Limes

$$\psi' = \psi V_\tau: \quad \psi'(q) = e^{i\tau q} \cdot \psi(q): \quad (50)$$

das ist die unitäre Abbildung V_τ , welche die Größe $e^{i\tau q}$ darstellt. Der gleiche Grenzübergang an der ersten Gleichung liefert die unitäre Abbildung

$$\psi' = \psi U_\sigma: \quad \psi'(q) = \psi(q - \sigma), \quad (51)$$

welche $e^{i\sigma p}$ repräsentiert. Beide Abbildungen sind in der Tat unitär, weil sie die Gleichung (49) invariant lassen; sie bilden, den verschiedenen Werten von σ bzw. τ entsprechend, zwei einparametrische Abelsche Gruppen linearer Funktionaltransformationen:

$$U_{\sigma+\sigma'} = U_\sigma U_{\sigma'}, \quad V_{\tau+\tau'} = V_\tau V_{\tau'}.$$

$\psi U_\sigma V_\tau$ ist die Funktion $e^{i\tau q} \cdot \psi(q - \sigma)$, $\psi V_\tau U_\sigma$ aber $= e^{i\tau(q-\sigma)} \cdot \psi(q - \sigma)$, so daß, wie es sein muß, die Kommutatorgleichung gilt:

$$\psi U_\sigma V_\tau = e^{i\sigma\tau} \cdot \psi V_\tau U_\sigma.$$

Der Größe $e^{(\sigma p + \tau q)}$ entspricht nach (37) die Abbildung

$$\psi(q) \rightarrow \psi'(q) = e^{-1/2 i\sigma\tau} \cdot e^{i\tau q} \psi(q - \sigma).$$

Geht man endlich auf die infinitesimalen Operationen zurück — was freilich im allgemeinen nicht zweckmäßig ist —, so bekommt man als Repräsentation von

$$p: \quad \delta\psi = i \frac{d\psi(q)}{dq}, \quad \text{von } q: \quad \delta\psi = q \cdot \psi(q). \quad (52)$$

Damit sind wir bei der Schrödingerschen Fassung angelangt. Die Eigenfunktionen $\psi_n(q)$ seiner Wellengleichung haben danach die Bedeutung, daß sie die unitäre Transformation angeben, welche zwischen den beiden Hauptachsensystemen der Größe q und der Energie E vermittelt. Im Hinblick auf den ersten Teil ergeben sich daraus die bekannten Paulischen Ansätze für ihre Wahrscheinlichkeitsbedeutung.

Die Übertragung auf mehrere Freiheitsgrade ist mühelos durchführbar. Die Kinematik eines Systems, die durch eine konti-



Die Eindeutigkeit der Schrödingerschen Operatoren.

Von

J. v. Neumann in Berlin.

1. Die sogenannte Vertauschungsrelation

$$PQ - QP = \frac{h}{2\pi i} 1$$

ist in der neuen Quantentheorie von fundamentaler Bedeutung, sie ist es, die den „Koordinaten-Operator“ R und den „Impuls-Operator“ P im wesentlichen definiert¹⁾. Mathematisch gesprochen, liegt darin die folgende Annahme: Seien P, Q zwei Hermitesche Funktionaloperatoren des Hilbertschen Raumes, dann werden sie durch die Vertauschungsrelation bis auf eine Drehung des Hilbertschen Raumes, d. i. eine unitäre Transformation U , eindeutig festgelegt²⁾. Es liegt im Wesen der Sache, daß noch der Zusatz gemacht werden muß: vorausgesetzt, daß P, Q ein irreduzibles System bilden (vgl. weiter unten Anm. ⁶⁾). Wird nun, wie es sich durch die Schrödingersche Fassung der Quantentheorie als besonders günstig erwies, der Hilbertsche Raum als Funktionenraum interpretiert — der Einfachheit halber etwa als Raum aller komplexen Funktionen $f(q)$ ($-\infty < q < +\infty$) mit endlichem $\int_{-\infty}^{+\infty} |f(q)|^2 dq$ —, so gibt es nach Schrödinger ein besonders einfaches Lösungssystem der Vertauschungsrelation

$$Q: f(q) \rightarrow qf(q), \quad P: f(q) \rightarrow \frac{h}{2\pi i} \frac{d}{dq} f(q) \text{ } ^3).$$

¹⁾ Vgl. Born-Heisenberg-Jordan, Zeitschr. f. Phys. **34** (1925), S. 858—888, ferner Dirac, Proc. Roy. Soc. **109** (1925) u. f. Besonders in der letztgenannten Darstellung ist die Rolle dieser Relation fundamental. Einen interessanten Versuch zur Begründung des im folgenden zu diskutierenden Eindeutigkeitsatzes machte Jordan, Zeitschr. f. Phys. **37** (1926), S. 383—386. Indessen beruht dieser auf Konvergenzannahmen über Potenzreihen unbeschränkter Operatoren, deren Gültigkeitsbereich fraglich ist.

²⁾ Dieselbe bewirkt ein Ersetzen von P, Q durch UPU^{-1}, UQU^{-1} , wodurch weder der Hermitesche Charakter noch das Bestehen der Vertauschungsrelation berührt wird.

³⁾ Vgl. Schrödinger, Annalen d. Phys. **79** (1926), S. 734—756.

Sind nun dies die im wesentlichen einzigen (irreduziblen) Lösungen der Vertauschungsrelation?

Indessen ist die Aufgabe in dieser Form nicht genügend präzise formuliert. Denn als P, Q sind, wie es die Schrödingerschen Lösungen zeigen, auch unbeschränkte, nicht überall definierte Operatoren ins Auge zu fassen, und für diese wird der Operator $PQ - QP$ nicht überall definiert sein, während es der (auf der anderen Seite der Vertauschungsrelation stehende) Operator $\frac{\hbar}{2\pi i} 1$ ist. Die beiden Seiten können also nur gleichgesetzt werden, wenn ihre Definitionsbereiche (d. h. der der linken Seite) näher umschrieben werden. Dieser Schwierigkeit kann man folgendermaßen aus dem Wege gehen:

Durch formale Operatorenrechnung folgt aus der Vertauschungsrelation ($F(x)$ analytisch, $F'(x)$ seine Ableitung, vgl. Anm. 4))

$$PF(Q) - F(Q)P = \frac{\hbar}{2\pi i} F'(Q),$$

und hieraus für $F(x) = e^{\frac{2\pi i}{\hbar} \beta x}$

$$e^{-\frac{2\pi i}{\hbar} \beta Q} P e^{\frac{2\pi i}{\hbar} \beta Q} = P + \beta 1.$$

Hieraus folgt wieder formal

$$e^{-\frac{2\pi i}{\hbar} \beta Q} F(P) e^{\frac{2\pi i}{\hbar} \beta Q} = F(P + \beta 1),$$

und somit für $F(x) = e^{\frac{2\pi i}{\hbar} \alpha x}$

$$e^{\frac{2\pi i}{\hbar} \alpha P} e^{\frac{2\pi i}{\hbar} \beta Q} = e^{\frac{2\pi i}{\hbar} \alpha \beta} \cdot e^{\frac{2\pi i}{\hbar} \beta Q} e^{\frac{2\pi i}{\hbar} \alpha P}.$$

Diese Gleichung ist von Weyl aufgestellt und als Ersatz der Vertauschungsrelation vorgeschlagen worden⁴⁾. Ihr großer Vorzug besteht in folgendem: Es ist unter Umständen möglich, mit Hilfe der Operatoren P, Q einparametrische Scharen $U(\alpha) = e^{\frac{2\pi i}{\hbar} \alpha P}$, $V(\beta) = e^{\frac{2\pi i}{\hbar} \beta Q}$ zu definieren, die unitär sind, und dem Multiplikationsgesetz

$$U(\alpha)U(\beta) = U(\alpha + \beta), \quad V(\alpha)V(\beta) = V(\alpha + \beta)$$

genügen⁵⁾. Dann stehen auf beiden Seiten der Weylschen Gleichung

$$U(\alpha)V(\beta) = e^{\frac{2\pi i}{\hbar} \alpha \beta} \cdot V(\beta)U(\alpha)$$

⁴⁾ Vgl. Weyl, Zeitschr. f. Phys. 46 (1928), Seite 1—46.

⁵⁾ Vgl. Weyl, Anm. 4), ferner Stone, Proc. of Nat. Academy 1930. Im Schrödingerschen Falle wird, wie man leicht erkennt:

$$U(\alpha): f(q) \rightarrow f(q + \alpha), \quad V(\beta): f(q) \rightarrow e^{\frac{2\pi i}{\hbar} \beta q} f(q).$$

Orthogonalsystem:

$$\begin{aligned} (A(\alpha) B(\alpha) f, g) &= (B(\alpha) f, A(\alpha)^* g) = \sum_{n=1}^{\infty} (B(\alpha) f, \varphi_n) (\varphi_n, A^*(\alpha) g) \\ &= \sum_{n=1}^{\infty} \overline{(A(\alpha) g, \varphi_n)} (B(\alpha) f, \varphi_n). \end{aligned}$$

Dasselbe gilt, wenn an der Stelle von α mehrere Variable α, β, \dots stehen. Wir kehren nun zu unserem Problem zurück, ersetzen aber in $V(\beta)$ β durch $\frac{\hbar}{2\pi}\beta$. Dann lautet es so:

Alle $U(\alpha), V(\beta)$ seien unitäre Operatoren, die meßbar von α, β abhängen. Es gelten die Relationen

$$\begin{aligned} U(\alpha) U(\beta) &= U(\alpha + \beta), & V(\alpha) V(\beta) &= V(\alpha + \beta), \\ U(\alpha) V(\beta) &= e^{i\alpha\beta} V(\beta) U(\alpha). \end{aligned}$$

Alle derartigen Systeme sind zu bestimmen.

Wenn wir die (von α, β meßbar abhängende, unitäre) Operatorenschar

$$S(\alpha, \beta) = e^{-\frac{1}{2}i\alpha\beta} U(\alpha) V(\beta) = e^{\frac{1}{2}i\alpha\beta} V(\beta) U(\alpha)$$

einführen, so können wir die obigen Relationen zu

$$S(\alpha, \beta) S(\gamma, \delta) = e^{\frac{1}{2}i(\alpha\delta - \beta\gamma)} S(\alpha + \gamma, \beta + \delta)$$

zusammenfassen. Infolgedessen ist $S(0, 0)$ die Einheit, und daher $S(-\alpha, -\beta)$ zu $S(\alpha, \beta)$ reziprok, also $S(\alpha, \beta)^* = S(-\alpha, -\beta)$. Es sollen nun Linearaggregate der $S(\alpha, \beta)$ betrachtet werden, diese werden folgendermaßen definiert: Sei $\alpha(\alpha, \beta)$ eine über die ganze α, β -Ebene absolut integrierbare Funktion, dann ist wegen der Schwarzschen Ungleichheit

$$|(S(\alpha, \beta) f, g)| \leq |S(\alpha, \beta) f| \cdot |g| = |f| \cdot |g|,$$

d. h. beschränkt, also auch das Integral

$$\iint \alpha(\alpha, \beta) (S(\alpha, \beta) f, g) d\alpha d\beta$$

absolut konvergent. Und zwar ist es, wenn wir $c = \iint |\alpha(\alpha, \beta)| d\alpha d\beta$ setzen, absolut $\leq c \cdot |f| |g|$. Dabei ist es in f linear und in g konjugiert-linear. Daher ist ein Satz von F. Riesz anwendbar⁹⁾, wonach bei festem f ein f^* existiert, so daß dieser Ausdruck für jedes $g = (f^*, g)$ ist, und zwar ist $|f^*| \leq c \cdot |f|$. f^* ist durch f bestimmt, und zwar ist die Abhängigkeit linear, wir können also einen linearen Operator A durch $Af = f^*$ definieren, nach der obigen Formel ist A auch beschränkt. Wir schreiben symbolisch

$$A = \iint \alpha(\alpha, \beta) S(\alpha, \beta) d\alpha d\beta,$$

obwohl die Definition eigentlich

$$(Af, g) = \iint \alpha(\alpha, \beta) (S(\alpha, \beta) f, g) d\alpha d\beta$$

lautet. $\alpha(\alpha, \beta)$ heiße der Kern von A .

⁹⁾ Vgl. auch a. a. O. Anm. 7), Math. Annalen 102 (1930), S. 94, Anm. 52).

Wir beweisen einige Rechenregeln für diese Operatoren. Daß aA den Kern $a\alpha(\alpha, \beta)$ hat, ist klar, A^* hat wegen $S(\alpha, \beta)^* = S(-\alpha, -\beta)$ den Kern $\overline{a(-\alpha, -\beta)}$, $AS(u, v)$ und $S(u, v)A$ wegen der Multiplikationsregel der $S(\alpha, \beta)$ den Kern

$$e^{\frac{1}{2}i(\alpha v - \beta u)} \alpha(\alpha - u, \beta - v) \quad \text{bzw.} \quad e^{-\frac{1}{2}i(\alpha v - \beta u)} \alpha(\alpha - u, \beta - v).$$

Haben A, B die bzw. Kerne $\alpha(\alpha, \beta), \mathfrak{b}(\alpha, \beta)$, so hat $A + B$ offenbar $\alpha(\alpha, \beta) + \mathfrak{b}(\alpha, \beta)$, bei AB dagegen ist eine kleine Rechnung notwendig:

$$\begin{aligned} (ABf, g) &= (Bf, A^*g) = \iint \mathfrak{b}(\alpha, \beta) (S(\alpha, \beta)f, A^*g) d\alpha d\beta \\ &= \iint \mathfrak{b}(\alpha, \beta) (AS(\alpha, \beta)f, g) d\alpha d\beta \\ &= \iiint \mathfrak{b}(\alpha, \beta) e^{\frac{1}{2}i(\gamma\beta - \delta\alpha)} \alpha(\gamma - \alpha, \delta - \beta) (S(\gamma, \delta)f, g) d\alpha d\beta d\gamma d\delta \\ &= \iint \left[\iint e^{\frac{1}{2}i(\gamma\beta - \delta\alpha)} \alpha(\gamma - \alpha, \delta - \beta) \mathfrak{b}(\alpha, \beta) d\alpha d\beta \right] (S(\gamma, \delta)f, g) d\gamma d\delta. \end{aligned}$$

Der Kern von AB ist also (statt γ, δ schreiben wir wieder α, β , statt α, β ξ, η) $\iint e^{\frac{1}{2}i(\alpha\eta - \beta\xi)} \alpha(\alpha - \xi, \beta - \eta) \mathfrak{b}(\xi, \eta) d\xi d\eta$. (Die absolute Integrierbarkeit folgt aus der Deduktion.)

Schließlich zeigen wir: wenn A verschwindet, so ist auch sein Kern (bis auf eine Lebesguesche Nullmenge) gleich 0. Aus $A = 0$ folgt nämlich $S(-u, -v)AS(u, v) = 0$, also, da dieses den Kern $e^{i(\alpha v - \beta u)} \alpha(\alpha, \beta)$ hat,

$$\iint e^{i(\alpha v - \beta u)} \alpha(\alpha, \beta) (S(\alpha, \beta)f, g) d\alpha d\beta = 0.$$

Somit ist jedenfalls

$$\iint P(\alpha, \beta) \alpha(\alpha, \beta) (S(\alpha, \beta)f, g) d\alpha d\beta = 0,$$

wenn $P(\alpha, \beta)$ ein Linearaggregat von endlich vielen $e^{i(k\alpha + l\beta)}$ ist, also für jedes trigonometrische Polynom mit einer Periode $p > 0$ in α, β . Da der zweite Faktor absolut integrierbar ist, und der dritte beschränkt, können wir mit dem ersten ($P(\alpha, \beta)$) Grenzübergänge ausführen, falls dieser dabei gleichmäßig beschränkt bleibt. So können wir die Klasse der $P(\alpha, \beta)$ sukzessiv erweitern: 1. zu allen stetigen Funktionen mit einer Periode $p > 0$ in α, β , 2. zu allen beschränkten stetigen Funktionen, 3. zu allen beschränkten Funktionen der ersten Baireschen Klasse. Wenn also \mathfrak{R} ein beliebiges (endliches) Rechteck in der α, β -Ebene ist, so können wir $P(\alpha, \beta)$ in \mathfrak{R} gleich 1 und außerhalb = 0 setzen, es wird:

$$\iint_{\mathfrak{R}} \alpha(\alpha, \beta) (S(\alpha, \beta)f, g) d\alpha d\beta = 0$$

für alle diese \mathfrak{R} . Daher ist (mit Ausnahme einer α, β -Nullmenge) $\alpha(\alpha, \beta) (S(\alpha, \beta)f, g) = 0$. Dies gilt bei festem f, g , ist aber nur f fest, während g ein vollständiges normiertes Orthogonalsystem durchläuft, so gilt es für dieses f und alle genannten g auch noch mit Ausnahme einer

MATHEMATISCHE GRUNDLAGEN DER QUANTENMECHANIK

VON

JOHANN v. NEUMANN

UNVERÄNDERTER NACHDRUCK
DER ERSTEN AUFLAGE VON 1932

MIT 4 ABBILDUNGEN



SPRINGER-VERLAG
BERLIN · HEIDELBERG · NEW YORK

1968

rules regarding the relationship between the

$$\mathfrak{R}(q_1, \dots, q_k, p_1, \dots, p_k)$$

and the R , we stated the following special rules in III.1. and III.3.:

L. If the operators R, S correspond to the simultaneously observable quantities $\mathfrak{R}, \mathfrak{S}$, then the operator $aR + bS$ (a, b real numbers) corresponds to the quantity $a\mathfrak{R} + b\mathfrak{S}$.

F. If the operator R corresponds to the quantity \mathfrak{R} , then the operator $F(R)$ corresponds to the quantity $F(\mathfrak{R})$ [$F(\lambda)$ an arbitrary real function].

L., F. permit a certain generalization. That of **F.** is rigorously implied, and runs as follows:

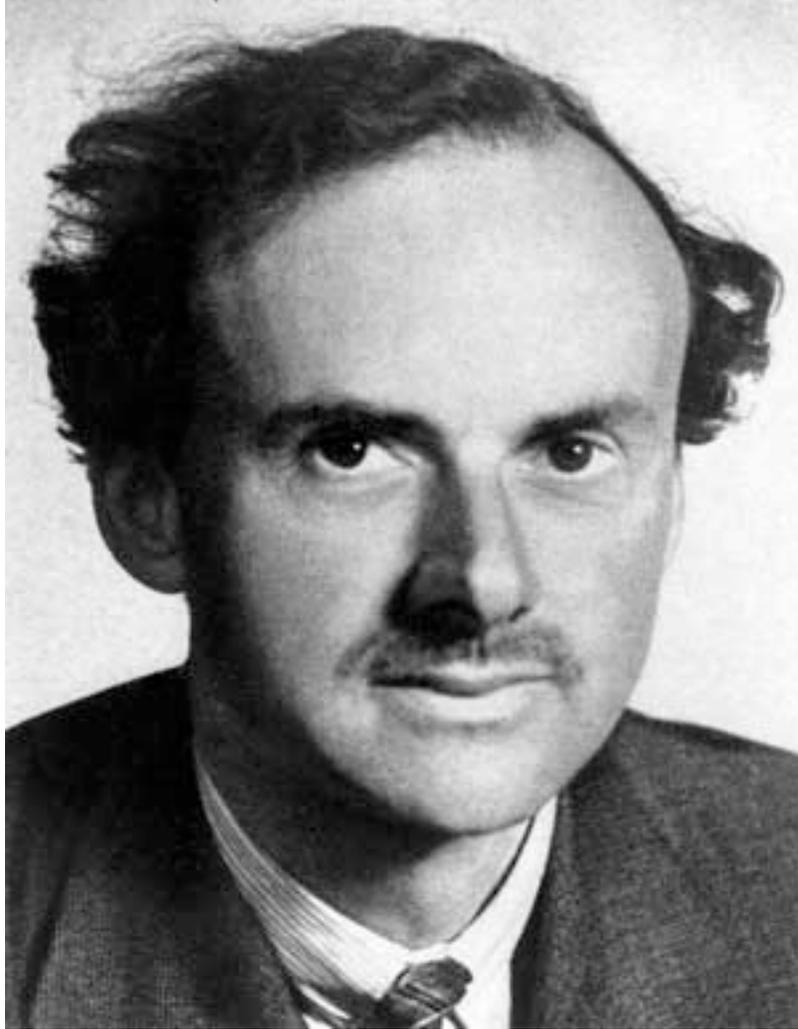
F*. If the operators R, S, \dots correspond to the simultaneously measurable quantities $\mathfrak{R}, \mathfrak{S}, \dots$ (which are consequently commutative; we assume that their number is finite), then the operator $F(R, S, \dots)$ corresponds to the quantity $F(\mathfrak{R}, \mathfrak{S}, \dots)$.

In this case, we shall assume that $F(\lambda, \mu, \dots)$ is a real polynomial in λ, μ, \dots , so that the meaning of $F(R, S, \dots)$ may be clear (R, S, \dots commutative) although **F*** could also be defined for arbitrary $F(\lambda, \mu, \dots)$ [for the definition of the general $F(R, S, \dots)$, see the reference in Note 94]. Now since each polynomial is obtained by repetition of the three operations $a\lambda, \lambda + \mu, \lambda\mu$, it suffices to consider these, and since $\lambda\mu = \frac{1}{4} [(\lambda + \mu)^2 - (\lambda - \mu)^2]$, i.e., is equal to

$$\frac{1}{4} \cdot (\lambda + \mu)^2 + (-\frac{1}{4}) \cdot (\lambda + (-1) \cdot \mu)^2$$

we can also replace these three operations by $a\lambda, \lambda + \mu, \lambda^2$. But the first two fall under **L.**, and the latter under **F.** Consequently, **F*** is proved.

On the other hand, **L.** is extended in quantum mechanics even to the case where $\mathfrak{R}, \mathfrak{S}$ are not simultaneously measurable. We shall discuss this question later



Note on Exchange Phenomena in the Thomas Atom. By P. A. M. DIRAC, Ph.D., St John's College.

[Read 19 May 1930.]

Introduction.

For dealing with atoms involving many electrons the accurate quantum theory, involving a solution of the wave equation in many-dimensional space, is far too complicated to be practicable. One must therefore resort to approximate methods. The best of these is Hartree's method of the self-consistent field*. Even this, however, is hardly practicable when one has to deal with very many electrons, so that one then requires a still simpler and rougher method. Such a method is provided by Thomas' atomic model†, in which the electrons are regarded as forming a perfect gas satisfying the Fermi statistics and occupying the region of phase space of lowest energy. This region of phase space is assumed to be saturated, with two electrons with opposite spins in each volume $(2\pi h)^3‡$, and the remainder is assumed to be empty. Although this model hitherto has not been justified theoretically, it seems to be a plausible approximation for the interior of a heavy atom and one may expect it to give with some accuracy the distribution of electric charge there.

The method of the self-consistent field has recently been established on a very much better theoretical basis in a paper by Fock§, which shows how one can take into account the exchange phenomena between the equivalent electrons. Fock shows that if one takes the best approximation to the many-dimensional wave function that is of the form of a product of a number of three-dimensional wave functions, one for each electron, then the three-dimensional wave functions will satisfy just Hartree's equations. In this way a theoretical justification for Hartree's method is obtained. The exclusion principle of Pauli, however, requires that the wave function representing a number of electrons shall always be antisymmetrical. One would therefore expect to get a better approximation if one first made the product of a number of three-dimensional wave functions antisymmetrical, by applying permutations and taking a linear combination, and then made it approach as closely as possible to

* Hartree, *Proc. Camb. Phil. Soc.*, Vol. 24, p. 111 (1927).

† Thomas, *Proc. Camb. Phil. Soc.*, Vol. 23, p. 542 (1926). See also Fermi, *Zeit. für Phys.*, Vol. 48, p. 73 (1928).

‡ h denotes Planck's constant divided by 2π .

§ Fock, *Zeit. für Phys.*, Vol. 61, p. 126 (1930).

where D denotes the diagonal sum (or integral) of the matrix following it. To verify the constancy of this quantity, we observe that, from (11),

$$\begin{aligned} D(\rho \dot{B}) &= \iint (q'' | \rho | q') (q' | \dot{B} | q'') dq' dq'' \\ &= \iiint \iiint (q'' | \rho | q') (q' q''' | V | q'' q^{1v}) (q^{1v} | \dot{\rho} | q''') dq' dq'' dq''' dq^{1v} \\ &= \iiint \iiint (q^{1v} | \dot{\rho} | q''') (q''' q' | V | q^{1v} q'') (q'' | \rho | q') dq' dq'' dq''' dq^{1v} \\ &= D(\dot{\rho} B), \end{aligned}$$

and similarly $D(\rho \dot{A}) = D(\dot{\rho} A)$.

Hence
$$\begin{aligned} \frac{d}{dt} D\{\rho(H_0 + \frac{1}{2}B - \frac{1}{2}A)\} &= D\{\dot{\rho}(H_0 + \frac{1}{2}B - \frac{1}{2}A)\} + \frac{1}{2}D(\rho B) - \frac{1}{2}D(\rho \dot{A}) \\ &= D\{\dot{\rho}(H_0 + \frac{1}{2}B - \frac{1}{2}A)\} + \frac{1}{2}D(\dot{\rho} B) - \frac{1}{2}D(\dot{\rho} A) \\ &= D(\dot{\rho} H). \end{aligned}$$

Thus
$$\begin{aligned} ih \frac{d}{dt} D\{\rho(H_0 + \frac{1}{2}B - \frac{1}{2}A)\} &= D\{(H\rho - \rho H)H\} \\ &= D(H\rho H) - D(\rho H^2) = 0, \end{aligned}$$

since the diagonal sum of a product is not changed by a cyclic permutation of the factors. Hence $D\{\rho(H_0 + \frac{1}{2}B - \frac{1}{2}A)\}$ is a constant of the motion. It may be interpreted as the energy integral, $D(\rho H_0)$ being the proper energy of the electrons (their kinetic energy plus their potential energy in the field of the nucleus), $\frac{1}{2}D(\rho B)$ being their interaction energy and $-\frac{1}{2}D(\rho A)$ being a correction for exchange effects.

Reduction to a classical density function.

We shall now examine what the equation of motion (9) becomes when the electron density ρ is spread over such a large volume of phase space that we can neglect the fact that the momenta p do not commute with the coordinates q and reduce our description of the atom to a classical one. We shall also now neglect the spin variables. Each element $(q' | \alpha | q'')$ of the matrix representing any dynamical variable α will now be connected with a certain Fourier component in the p -variables of $\alpha(qp)$ considered as a function of commuting q 's and p 's. We shall have, in fact,

$$(q' | \alpha | q'') = (2\pi h)^{-3} \int \alpha(qp) e^{i(q'-q'')p/h} dp \dots\dots\dots(14),$$

the connection being most accurate when the q on the right-hand side is taken to be the mean of q' and q'' . The converse equation will be

$$\alpha(qp) = \int (q' | \alpha | q'') e^{-i(q'-q'')p/h} d(q' - q'') \dots\dots(15).$$

If we apply equation (15) to the matrix $(q' | B | q'')$ defined by the first of equations (13), we obtain for the corresponding classical function $B(qp)$ of commuting q 's and p 's,

$$\begin{aligned} B(qp) &= e^2 \int \delta(q' - q'') e^{-i(q'-q'')p/h} d(q' - q'') \cdot \int \frac{(q''' | \rho | q''')}{r(qq''')} dq''' \\ &= e^2 \int \frac{(q''' | \rho | q''')}{r(qq''')} dq''' \\ &= \frac{e^2}{(2\pi h)^3} \int \frac{dq'}{r(qq')} \int \rho(q'p') dp' \dots\dots\dots(16), \end{aligned}$$

with the help of (14) applied to ρ . Thus $B(qp)$ is independent of p , as it must be since the matrix $(q' | B | q'')$ is diagonal, and it is just that function of q which corresponds to the potential arising from a distribution of electrons of density $\rho(qp)$ per volume $(2\pi h)^3$ of phase space.

If we now apply (15) to the matrix $(q' | A | q'')$ defined by the second of equations (13), we obtain

$$\begin{aligned} A(qp) &= e^2 \int \frac{(q' | \rho | q'')}{r(q'q'')} e^{-i(q'-q'')p/h} d(q' - q'') \\ &= \frac{e^2}{(2\pi h)^3} \int \rho(qp') dp' \int \frac{e^{-i(q'-q'')(p-p')/h}}{r(q'q'')} d(q' - q''), \end{aligned}$$

with the help of (14) applied to ρ . The second integral here can be evaluated. The $q' - q''$ and $p - p'$ appearing in the exponential are really vectors and their product should be understood to mean their scalar product. If we denote by θ the angle between these two vectors and by σ the magnitude of $q' - q''$, which is the same as $r(q'q'')$, we have for this second integral

$$\begin{aligned} \int \frac{e^{-i(q'-q'')(p-p')/h}}{r(q'q'')} d(q' - q'') &= \int_0^\infty \sigma d\sigma \int_{-1}^1 e^{-i\sigma|p-p'|\cos\theta/h} 2\pi d(\cos\theta) \\ &= \int_0^\infty 4\pi h d\sigma \sin\{\sigma|p-p'|/h\} / |p-p'| \\ &= 4\pi h^2 / |p-p'|^2. \end{aligned}$$

Hence
$$A(qp) = \frac{e^2}{2\pi^2 h} \int \frac{\rho(qp')}{|p-p'|^2} dp' \dots\dots\dots(17).$$

These expressions for $B(q)$ and $A(qp)$ substituted in (12) will give us a classical Hamiltonian governing the motion of the distribution of electrons. To proceed with the solution we shall now make an assumption concerning the form of ρ , which seems to be plausible when we are dealing with the state of lowest energy of the atom, namely the assumption that for each value of q the phase space is saturated, with two electrons per volume $(2\pi\hbar)^3$, in a region for which the magnitude of the momentum p is less than a certain value P , and is empty outside this region. In symbols

$$\begin{aligned} \rho(qp) &= 2 && |p| < P \\ &= 0 && |p| > P, \end{aligned}$$

where P is a certain function of q . This assumption gives us at once from (16)

$$B(q) = \frac{e^2}{3\pi^2\hbar^3} \int \frac{P^3(q') dq'}{r(qq')} \dots\dots\dots(18),$$

and from (17), after a straightforward integration,

$$A(qp) = \frac{e^2}{\pi\hbar} \left[\frac{P^2 - |p|^2}{|p|} \log \left| \frac{P + |p|}{P - |p|} \right| + 2P \right] \dots(19).$$

For a stationary state of the atom ρ must be constant, so that the Poisson bracket of ρ with H must vanish. With ρ of the form which we have assumed, this condition becomes that H must be constant along the boundary between the saturated and unoccupied phase space, *i.e.* $H(qP)$ must be constant. Now

$$H(qP) = H_0(qP) + B(q) - 2e^2/\pi\hbar \cdot P.$$

The constancy of this gives us a condition for the unknown function P .

For an atom with atomic number Z , we shall have

$$H_0 = -\frac{Ze^2}{r} + \frac{p^2}{2m},$$

where r denotes distance from the nucleus. We may assume spherical symmetry, so that P is a function of r only. We now get for the value of H on the boundary

$$\begin{aligned} H(qP) &= -\frac{Ze^2}{r} + \frac{P^2}{2m} - \frac{2e^2P}{\pi\hbar} \\ &\quad + \frac{4e^2}{3\pi\hbar^3} \left[\frac{1}{r} \int_0^r r'^2 P^3(r') dr' + \int_r^\infty r' P^3(r') dr' \right]. \end{aligned}$$

By equating to zero the differential coefficient of this with respect to r , we get

$$\frac{Ze^2}{r^2} + \frac{d}{dr} \left(\frac{P^2}{2m} - \frac{2e^2P}{\pi\hbar} \right) - \frac{4e^2}{3\pi\hbar^2} \frac{1}{r^2} \int_0^r r'^2 P^3(r') dr' = 0.$$

Multiplication by r^2 and a further differentiation then give

$$\frac{d}{dr} \left\{ r^2 \frac{d}{dr} \left(\frac{P^2}{2m} - \frac{2e^2}{\pi h} P \right) \right\} = \frac{4e^2}{3\pi h^2} r^2 P^3 \dots\dots\dots(20).$$

This is a differential equation which determines P , the maximum momentum for an electron, as a function of r . It differs from the corresponding equation in Thomas' theory only on account of its having a term linear in P on the left-hand side, which term may be considered to represent the exchange effects. This term will not be very important in the interior of a heavy atom, since the ratio of its coefficient to that of the P^2 term is $4e^2m/\pi h$, which is $4/\pi$ times the momentum of an electron in the first Bohr orbit in the hydrogen atom. For a sufficiently large value of r , however, the extra term causes P to become negative, and then to oscillate with decreasing amplitude and increasing period as $r \rightarrow \infty$. A negative value for P has, of course, no physical meaning. The fact that our theory gives this meaningless result for the outside of the atom is hardly surprising, since the approximation we made of regarding ρ as a function of commuting q 's and p 's is certainly not valid for this region. We may expect equation (20) to be more accurate than Thomas' equation in the interior of the atom, in spite of the fact that it is inapplicable outside.



On the Quantum Correction For Thermodynamic Equilibrium

By E. WIGNER

Department of Physics, Princeton University

(Received March 14, 1932)

The probability of a configuration is given in classical theory by the Boltzmann formula $\exp[-V/kT]$ where V is the potential energy of this configuration. For high temperatures this of course also holds in quantum theory. For lower temperatures, however, a correction term has to be introduced, which can be developed into a power series of h . The formula is developed for this correction by means of a probability function and the result discussed.

1

IN classical statistical mechanics the relative probability for the range p_1 to p_1+dp_1 ; p_2 to p_2+dp_2 ; \dots ; p_n to p_n+dp_n for the momenta and x_1 to x_1+dx_1 ; x_2 to x_2+dx_2 ; \dots ; x_n to x_n+dx_n for the coordinates is given for statistical equilibrium by the Gibbs-Boltzmann formula

$$P(x_1, \dots, x_n; p_1, \dots, p_n) dx_1 \dots dx_n dp_1 \dots dp_n = e^{-\beta\epsilon} dx_1 \dots dx_n dp_1 \dots dp_n \quad (1)$$

where ϵ is the sum of the kinetic and potential energy V

$$\epsilon = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + \dots + \frac{p_n^2}{2m_n} + V(x_1 \dots x_n) \quad (2)$$

and β is the reciprocal temperature T divided by the Boltzmann constant

$$\beta = 1/kT. \quad (3)$$

In quantum theory there does not exist any similar simple expression for the probability, because one cannot ask for the simultaneous probability for the coordinates and momenta. Moreover, it is not possible to derive a simple expression even for the relative probabilities of the coordinates alone—as is given in classical theory by $e^{-\beta V(x_1 \dots x_n)}$. One sees this by considering that this expression would give at once the square of the wave function of the lowest state $|\psi_0(x_1 \dots x_n)|^2$ when $\beta = \infty$ is inserted and on the other hand we know that it is not possible, in general, to derive a closed formula for the latter.

The thermodynamics of quantum mechanical systems is in principle, however, given by a formula of Neumann,¹ who has shown that the mean value of any physical quantity is, (apart from a normalizing constant depending only on temperature), the sum of the diagonal elements of the matrix

$$Qe^{-\beta H} \quad (4)$$

where Q is the matrix (operator) of the quantity under consideration and H is the Hamiltonian of the system. As the diagonal sum is an invariant under

¹ J. von Neumann, Gött. Nachr. p. 273, 1927.

transformations, one can choose any matrix or operator-representation for the Q and H . In building the exponential of H one must, of course, take into account the non-commutability of the different parts of H .

2

It does not seem to be easy to make explicit calculations with the form (4) of the mean value. One may resort therefore to the following method.

If a wave function $\psi(x_1 \cdots x_n)$ is given one may build the following expression²

$$P(x_1, \cdots, x_n; p_1, \cdots, p_n) = \left(\frac{1}{h\pi}\right)^n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dy_1 \cdots dy_n \psi(x_1 + y_1 \cdots x_n + y_n)^* \psi(x_1 - y_1 \cdots x_n - y_n) e^{2i(p_1 y_1 + \cdots + p_n y_n)/h} \quad (5)$$

and call it the probability-function of the simultaneous values of $x_1 \cdots x_n$ for the coordinates and $p_1 \cdots p_n$ for the momenta. In (5), as throughout this paper, h is the Planck constant divided by 2π and the integration with respect to the y has to be carried out from $-\infty$ to ∞ . Expression (5) is real, but not everywhere positive. It has the property, that it gives, when integrated with respect to the p , the correct probabilities $|\psi(x_1 \cdots x_n)|^2$ for the different values of the coordinates and also it gives, when integrated with respect to the x , the correct quantum mechanical probabilities

$$\left| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \psi(x_1 \cdots x_n) e^{-i(p_1 x_1 + \cdots + p_n x_n)/h} dx_1 \cdots dx_n \right|^2$$

for the momenta p_1, \cdots, p_n . The first fact follows simply from the theorem about the Fourier integral and one gets the second by introducing $x_k + y_k = u_k; x_k - y_k = v_k$ into (5).

Hence it follows, furthermore, that one may get the correct expectation values of any function of the coordinates or the momenta for the state ψ by the normal probability calculation with (5). As expectation values are additive this even holds for a sum of a function of the coordinates and a function of the momenta as, e.g., the energy H . In formulas, it is

$$\begin{aligned} & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dx_1 \cdots dx_n dp_1 \cdots dp_n [f(p_1 \cdots p_n) + g(x_1 \cdots x_n)] \\ & P(x_1 \cdots x_n; p_1 \cdots p_n) \\ & = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \psi(x_1 \cdots x_n)^* \left[f\left(\frac{h}{i} \frac{\partial}{\partial x_1}, \cdots, \frac{h}{i} \frac{\partial}{\partial x_n}\right) \right. \\ & \quad \left. + g(x_1 \cdots x_n) \right] \psi(x_1 \cdots x_n) dx_1 \cdots dx_n \end{aligned} \quad (6)$$

for any ψ, f, g , if P is given by (5).

² This expression was found by L. Szilard and the present author some years ago for another purpose.

Of course $P(x_1, \dots, x_n; p_1, \dots, p_n)$ cannot be really interpreted as the simultaneous probability for coordinates and momenta, as is clear from the fact, that it may take negative values. But of course this must not hinder the use of it in calculations as an auxiliary function which obeys many relations we would expect from such a probability. It should be noted, furthermore, that (5) is not the only bilinear expression in ψ , which satisfies (6). There must be a great freedom in the expression (5), as it makes from a function ψ of n variables one with $2n$ variables. It may be shown, however, that there does not exist any expression $P(x_1 \dots x_n; p_1 \dots p_n)$ which is bilinear in ψ , satisfies (6) and is everywhere (for all values of $x_1, \dots, x_n, p_1, \dots, p_n$) positive, so (5) was chosen from all possible expressions, because it seems to be the simplest.

If $\psi(x_1, \dots, x_n)$ changes according to the second Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = - \sum_{k=1}^n \frac{\hbar^2}{2m_k} \frac{\partial^2 \psi}{\partial x_k^2} + V(x_1, \dots, x_n) \psi \tag{7}$$

the change of $P(x_1, \dots, x_n; p_1, \dots, p_n)$ is given by

$$\frac{\partial P}{\partial t} = - \sum_{k=1}^n \frac{p_k}{m_k} \frac{\partial P}{\partial x_k} + \sum \frac{\partial^{\lambda_1 + \dots + \lambda_n} V}{\partial x_1^{\lambda_1} \dots \partial x_n^{\lambda_n}} \frac{(h/2i)^{\lambda_1 + \dots + \lambda_n - 1}}{\lambda_1! \dots \lambda_n!} \frac{\partial^{\lambda_1 + \dots + \lambda_n} P}{\partial p_1^{\lambda_1} \dots \partial p_n^{\lambda_n}} \tag{8}$$

where the last summation has to be extended over all positive integer values of $\lambda_1, \dots, \lambda_n$ for which the sum $\lambda_1 + \lambda_2 + \dots + \lambda_n$ is odd. In fact we get for $\partial P/\partial t$ by (5) and (7)

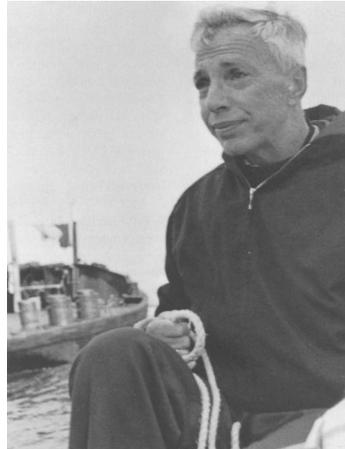
$$\begin{aligned} \frac{\partial P}{\partial t} &= \frac{1}{(h\pi)^n} \int \dots \int dy_1 \dots dy_n e^{2i(p_1 y_1 + \dots + p_n y_n)/\hbar} \\ &\cdot \left\{ \sum_k \frac{i\hbar}{2m_k} \left[- \frac{\partial^2 \psi(x_1 + y_1, \dots, x_n + y_n)^*}{\partial x_k^2} \psi(x_1 - y_1, \dots, x_n - y_n) \right. \right. \\ &+ \psi(x_1 + y_1, \dots, x_n + y_n)^* \frac{\partial^2 \psi(x_1 - y_1, \dots, x_n - y_n)}{\partial x_k^2} \\ &+ \frac{i}{\hbar} [V(x_1 + y_1, \dots, x_n + y_n) \\ &\left. \left. - V(x_1 - y_1, \dots, x_n - y_n)] \psi(x_1 + y_1, \dots, x_n + y_n)^* \psi(x_1 - y_1, \dots, x_n - y_n) \right\}. \end{aligned} \tag{9}$$

Here one can replace the differentiations with respect to x_k by differentiations with respect to y_k and perform in the first two terms one partial integration with respect to y_k . In the last term we can develop $V(x_1 + y_1, \dots, x_n + y_n)$ and $V(x_1 - y_1, \dots, x_n - y_n)$ in a Taylor series with respect to the y and get

$$\begin{aligned} \frac{\partial P}{\partial t} &= \frac{1}{(\pi\hbar)^n} \int \dots \int dy_1 \dots dy_n e^{2i(p_1 y_1 + \dots + p_n y_n)/\hbar} \\ &\cdot \left\{ \sum_k \frac{p_k}{m_k} \left[- \frac{\partial \psi(x_1 + y_1, \dots, x_n + y_n)^*}{\partial y_k} \psi(x_1 - y_1, \dots, x_n - y_n) \right. \right. \end{aligned}$$

o.6 Hilbrand Johannes Groenewold

29 June 1910 – 23 November 1996^o



H Groenewold

Hip Groenewold was born in Muntendam, The Netherlands. He studied at the University of Groningen, from which he graduated in physics with subsidiaries in mathematics and mechanics in 1934.

In that same year, he went of his own accord to Cambridge, drawn by the presence there of the mathematician John von Neumann, who had given a solid mathematical foundation to quantum mechanics with his book *Mathematische Grundlagen der Quantenmechanik*. This period had a decisive influence on Groenewold's scientific thinking. During his entire life, he remained especially interested in the interpretation of quantum mechanics (e.g. some of his ideas are recounted in Saunders et al.^p). It is therefore not surprising that his PhD thesis, which he completed eleven years later, was devoted to this subject ^{Gro46}. In addition to his revelation of the star product, and associated technical details, Groenewold's achievement in his thesis was to escape the cognitive straightjacket of the mainstream view that the defining difference between classical mechanics and quantum mechanics was the use of c-number functions and operators, respectively. He understood that these were only habits of use and in no way restricted the physics.

Ever since his return from England in 1935 until his permanent appointment at theoretical physics in Groningen in 1951, Groenewold experienced difficulties finding a paid job in physics. He was an assistant to Zernike in Groningen for a few years, then he went to the Kamerlingh Onnes Laboratory in Leiden, and taught at a grammar school in the Hague from 1940 to 1942. There, he met the woman whom he married in 1942. He spent the remaining war years at several locations in the north of the Netherlands. In July 1945, he began work for another two years as an assistant to Zernike. Finally, he

^oThe material presented here contains statements taken from a previously published obituary, N Hugenholtz, "Hip Groenewold, 29 Juni 1910-23 November 1996", *Nederlands Tijdschrift voor Natuurkunde* 2 (1997) 31.

^pS Saunders, J Barrett, A Kent, and D Wallace, *Many Worlds?*, Oxford University Press (2010).

worked for four years at the KNMI (Royal Dutch Meteorological Institute) in De Bilt.

During all these years, Groenewold never lost sight of his research. At his suggestion upon completing his PhD thesis, in 1946, Rosenfeld, of the University of Utrecht, became his promoter, rather than Zernike. In 1951, he was offered a position at Groningen in theoretical physics: First as a lecturer, then as a senior lecturer, and finally as a professor in 1955. With his arrival at the University of Groningen, quantum mechanics was introduced into the curriculum.

In 1971 he decided to resign as a professor in theoretical physics in order to accept a position in the Central Interfaculty for teaching Science and Society. However, he remained affiliated with the theoretical institute as an extraordinary professor. In 1975 he retired.

In his younger years, Hip was a passionate puppet player, having brought happiness to many children's hearts with beautiful puppets he made himself. Later, he was especially interested in painting. He personally knew several painters, and owned many of their works. He was a great lover of the after-war CoBrA art. This love gave him much comfort during his last years.

ON THE PRINCIPLES OF ELEMENTARY QUANTUM MECHANICS

by H. J. GROENEWOLD

Natuurkundig Laboratorium der Rijks-Universiteit te Groningen

Summary

Our problems are about

α the correspondence $a \longleftrightarrow \mathbf{a}$ between physical quantities a and quantum operators \mathbf{a} (quantization) and

β the possibility of understanding the statistical character of quantum mechanics by averaging over uniquely determined processes as in classical statistical mechanics (interpretation).

α and β are closely connected. Their meaning depends on the notion of observability.

We have tried to put these problems in a form which is fit for discussion. We could not bring them to an issue. (We are inclined to restrict the meaning of α to the trivial correspondence $\mathbf{a} \rightarrow a$ (for $\lim \hbar \rightarrow 0$) and to deny the possibility suggested in β).

Meanwhile special attention has been paid to the measuring process (coupling, entanglement; ignorance, infringement; selection, measurement).

For the sake of simplicity the discussion has been confined to elementary non-relativistic quantum mechanics of scalar (spinless) systems with one linear degree of freedom without exchange. Exact mathematical rigour has not been aimed at.

1. Statistics and correspondence.

1.01 *Meaning.* When poring over

α the correspondence $a \longleftrightarrow \mathbf{a}$ between observables a and the operators \mathbf{a} , by which they are represented in elementary quantum mechanics,

β the statistical character of elementary quantum mechanics (we need α for β), we run a continuous risk of lapsing into meaningless problems. One should keep in mind the meaning of the conceptions and statements used. We only consider

M_o : *observational meaning*, determined by the relation with what is (in a certain connection) understood as observation,

M_f : *formal meaning*, determined with respect to the mathematical formalism without regard to observation.

Only M_o is of physical interest, M_f is only of academic interest. Dealing with M_f may sometimes suggest ideas, fruitful in the sense of M_o , but may often lead one astray.

1.02 *Quantization*. Very simple systems suffice for demonstrating the essential features of α and β . In elementary classical point mechanics a system is described by the coordinates q of the particles and the conjugate momenta p . We only write down a single set p, q , corresponding to one degree of freedom. Any other measurable quantity (observable) a of the system is a function $a(p, q)$ of p and q (and possibly of the time t). The equations of motion can be expressed in terms of P o i s s o n brackets

$$(a, b) = \frac{\partial a}{\partial p} \frac{\partial b}{\partial q} - \frac{\partial a}{\partial q} \frac{\partial b}{\partial p}. \quad (1.01)$$

When the same system is treated in elementary quantum mechanics, the (real) quantities a are replaced by (H e r m i t i a n) operators \mathbf{a} , which now represent the observables. In the equations of motion the P o i s s o n brackets (1.01) are replaced by the operator brackets

$$[\mathbf{a}, \mathbf{b}] = \frac{i}{\hbar} (\mathbf{a}\mathbf{b} - \mathbf{b}\mathbf{a}) \quad (\hbar = \frac{h}{2\pi}, h \text{ P l a n c k's constant of action}). \quad (1.02)$$

Problem α_1 is to find the correspondence $a \rightarrow \mathbf{a}$ (other problems α are stated further on).

1.03 *Statistical character*. The statements of quantum mechanics on observations are in general of statistical character. Problem β is whether the statistical quantum processes could be described by a statistical average over uniquely determined processes (statistical description of the 1st kind, type S^1) or not (statistical description of the 2nd kind, type S^2). The observability of the uniquely determined processes may be required (proper statistical description, type S_o) or not (formal statistical description, type S_f). (Classical statistical mechanics, e.g. are properly of the 1st kind, type S_o^1).

1.04 *Transition operator*. Before going on we have to deal for a moment with the operators and the wave functions.

The Hermitian operators \mathbf{a} form a non-commutative ring. The normalized elements (wave functions) of (generalized) Hilbert space on which they act from the left are denoted by φ_μ , the adjoint elements on which they act from the right are denoted by φ_μ^\dagger . Unless otherwise stated the inner product of φ_μ^\dagger and φ_ν is simply written $\varphi_\mu^\dagger \varphi_\nu$. The outer product of φ_μ^\dagger and φ_ν defines the transition operator

$$\mathbf{k}_{\nu\mu} = \varphi_\nu \varphi_\mu^\dagger, \quad \mathbf{k}_{\mu\nu}^\dagger = \mathbf{k}_{\mu\nu}. \quad (1.03)$$

Take a complete system of orthonormal wave functions φ_ν . The orthonormality is expressed by

$$\varphi_\mu^\dagger \varphi_\nu = \delta_{\mu\nu}, \quad (1.04)$$

the completeness by

$$\sum_\mu \varphi_\mu \varphi_\mu^\dagger = \mathbf{1}. \quad (1.05)$$

In continuous regions of the parameter μ the Weierstrass δ -symbol must be replaced by the Dirac δ -function and the sum by an integral. (1.04) and (1.05) show that every (normalizable) function φ can be expanded into

$$\varphi = \sum_\mu f_\mu \varphi_\mu \quad \text{with} \quad f_\mu = \varphi_\mu^\dagger \varphi. \quad (1.06)$$

$\mathbf{k}_{\nu\mu}$ and $\mathbf{k}_{\nu\mu}^\dagger$ transform $\varphi_{\mu'}$ and $\varphi_{\mu'}^\dagger$ according to

$$\mathbf{k}_{\nu\mu} \varphi_{\mu'} = \varphi_\nu \delta_{\mu\mu'}, \quad \text{and} \quad \varphi_{\mu'}^\dagger \mathbf{k}_{\nu\mu}^\dagger = \delta_{\mu'\mu} \varphi_\nu^\dagger \quad (1.07)$$

(that is why they are called transition operators). (1.04) gives

$$\mathbf{k}_{\mu\nu} \mathbf{k}_{\nu'\mu'} = \mathbf{k}_{\mu\mu'} \delta_{\nu\nu'}. \quad (1.08)$$

In particular $\mathbf{k}_{\mu\mu}$ and $\mathbf{k}_{\nu\nu}$ are for $\mu \neq \nu$ orthogonal projection operators (belonging to φ_μ and φ_ν respectively).

The trace of an operator \mathbf{a} (resulting when \mathbf{a} acts towards the right upon itself from the left, or opposite; when it bites its tail) is (according to (1.05)) defined by

$$\text{Tr} \mathbf{a} = \sum_\mu \varphi_\mu^\dagger \mathbf{a} \varphi_\mu. \quad (1.09)$$

(Because the right hand member is invariant under unitary transformations of the φ_μ , this definition is independent of the special choice of the complete orthonormal system of φ_μ). This gives

$$\text{Tr}(\mathbf{k}_{\nu\mu} \mathbf{a}) = \varphi_\mu^\dagger \mathbf{a} \varphi_\nu. \quad (1.10)$$

(1.04) and (1.05) can be written

$$\text{Tr} \mathbf{k}_{\nu\mu} = \delta_{\nu\mu}, \quad (1.11)$$

$$\sum_\mu \mathbf{k}_{\mu\mu} = \mathbf{1} \quad (1.12)$$

and further imply

$$Tr(\mathbf{k}_{\mu\nu}\mathbf{k}_{\nu'\mu'}) = \delta_{\mu'\mu}\delta_{\nu\nu'}, \quad (1.13)$$

$$\sum_{\mu,\nu} \mathbf{k}_{\nu\mu} Tr(\mathbf{k}_{\mu\nu}\mathbf{a}) = \mathbf{a} \text{ (for every } \mathbf{a}\text{)}. \quad (1.14)$$

(1.13) and (1.14) show that every operator \mathbf{a} (with adjoint \mathbf{a}^\dagger) can be expanded into

$$\mathbf{a} = \sum_{\mu,\nu} \alpha_{\nu\mu} \mathbf{k}_{\mu\nu} \text{ with } \alpha_{\nu\mu} = Tr(\mathbf{k}_{\nu\mu}\mathbf{a}). \quad (1.15)$$

$\alpha_{\nu\mu}$ is the matrix element (1.10) of \mathbf{a} with respect to φ_ν and φ_μ . It follows further that if $Tr(\mathbf{a}\mathbf{c}) = 0$ for every \mathbf{a} , then $\mathbf{c} = 0$ and therefore (1.14) is equivalent to

$$\sum_{\mu,\nu} Tr(\mathbf{k}_{\nu\mu}\mathbf{b}) Tr(\mathbf{k}_{\mu\nu}\mathbf{a}) = Tr(\mathbf{a}\mathbf{b}) \text{ (for every } \mathbf{a} \text{ and } \mathbf{b}\text{)}. \quad (1.16)$$

Further

$$Tr(\mathbf{a}\mathbf{b}) = Tr(\mathbf{b}\mathbf{a}). \quad (1.17)$$

When \mathbf{a} is a Hermitian operator

$$\mathbf{a}^\dagger = \mathbf{a}, \quad \alpha_{\nu\mu}^* = \alpha_{\mu\nu} \quad (1.18)$$

(the asterik denotes the complex conjugate), the system of eigenfunctions φ_μ with eigenvalues a_μ

$$\mathbf{a}\varphi_\mu = a_\mu\varphi_\mu \quad (1.19)$$

can serve as reference system. In this representation (1.15) takes the diagonal form

$$\mathbf{a} = \sum_{\mu} a_{\mu} \mathbf{k}_{\mu\mu}. \quad (1.20)$$

1.05 *Statistical operator*¹⁾. The quantum state of a system is said to be pure, if it is represented by a wave function φ_μ . The statistical operator of the state is defined by the projection operator $\mathbf{k}_{\mu\mu}$ of φ_μ . We will see that the part of the statistical operator is much similar to that of a statistical distribution function. The most general quantum state of the system is a statistical mixture of (not necessarily orthogonal) pure states with projection operators $\mathbf{k}_{\mu\mu}$ and non-negative weights k_μ , which are normalized by

$$\sum_{\mu} k_{\mu} = 1. \quad (1.21)$$

(In some cases the sum diverges and the right member actually should symbolically be written as a δ -function). The statistical operator of the mixture is (in the same way as it would be done for

a distribution function) defined by

$$\mathbf{k} = \sum_{\mu} k_{\mu} \mathbf{k}_{\mu\mu} \quad (1.22)$$

and because of (1.21) normalized by

$$\text{Tr} \mathbf{k} = 1. \quad (1.23)$$

(we will always write 1 for the right member, though in some cases it actually should be written as a δ -function). For brevity we often speak of the state (or mixture) \mathbf{k} .

An arbitrary non-negative definite normalized Hermitian operator \mathbf{k} ($\text{Tr} \mathbf{k} = 1$) has non-negative eigenvalues k_{μ} , for which $\sum k_{\mu} = 1$ and corresponding eigenstates with projection operators $\mathbf{k}_{\mu\mu}$. Therefore \mathbf{k} can according to (1.20) be expanded in the form (1.22) and represents a mixture of its (orthogonal) eigenstates with weights given by the eigenvalues.

The statistical operator $\mathbf{k}_{\mu\mu}$ of a pure state is from the nature of the case idempotent ($\mathbf{k}_{\mu\mu}^2 = \mathbf{k}_{\mu\mu}$). If on the other hand an idempotent normalized Hermitian operator \mathbf{k} is expanded with respect to its eigenstates $\mathbf{k}_{\mu\mu}$ with eigenvalues k_{μ} , we get

$$\mathbf{k}^2 = \mathbf{k}, \quad k_{\mu}^2 = k_{\mu}; \quad \text{Tr} \mathbf{k} = 1, \quad \sum_{\mu} k_{\mu} = 1, \quad (1.24)$$

so that one eigenvalue k_{ν} is 1, all other are 0. Then \mathbf{k} is the projection operator of the pure state φ_{ν}

$$\mathbf{k} = \mathbf{k}_{\nu\nu}. \quad (1.25)$$

Therefore pure states and only these have idempotent statistical operators.

Suppose the normalized statistical operator \mathbf{k} of an arbitrary quantum state is expanded in some way into other normalized (but not necessarily orthogonal) statistical operators \mathbf{k}_r with non-negative weights k_r

$$\mathbf{k} = \sum_r k_r \mathbf{k}_r; \quad k_r \geq 0. \quad (1.26)$$

This gives

$$\mathbf{k} - \mathbf{k}^2 = \sum_r k_r (\mathbf{k}_r - \mathbf{k}_r^2) + \frac{1}{2} \sum_{r,s} k_r k_s (\mathbf{k}_r - \mathbf{k}_s)^2. \quad (1.27)$$

If we expand with respect to pure states \mathbf{k}_r ($\mathbf{k}_r^2 = \mathbf{k}_r$), (1.27) becomes

$$\mathbf{k} - \mathbf{k}^2 = \frac{1}{2} \sum_{r,s} k_r k_s (\mathbf{k}_r - \mathbf{k}_s)^2. \quad (1.28)$$

This shows that $\mathbf{k} - \mathbf{k}^2$ is a non-negative definite operator. If the given state is pure ($\mathbf{k}^2 = \mathbf{k}$) all terms at the right hand side of (1.27) (which are non-negative definite) must vanish separately. For the terms of the first sum this means that all states \mathbf{k}_r with non-vanishing weight ($k_r > 0$) must be pure, for the terms of the second sum it means further that all these states must be identical with each other and therefore also with the given state ($\mathbf{k}_r = \mathbf{k}$). The given state is then said to be indivisible. If the given state is a mixture, $\mathbf{k} - \mathbf{k}^2$ must be positive definite. Then at least one term at the right hand side of (1.28) must be different from zero. This means that at least two different states \mathbf{k}_r and \mathbf{k}_s ($\mathbf{k}_r \neq \mathbf{k}_s$) must have non-vanishing weight ($k_r > 0, k_s > 0$). The given state is then said to be divisible. Thus pure states and only these are indivisible. This has been proved in a more exact way by von Neumann¹⁾.

1.06 *Observation*. In order to establish the observational meaning M_o , one must accept a definite notion of observation. We deal with 3 different notions:

O_c : *the classical notion*: all observables $a(p, q)$ can be measured without fundamental restrictions and without disturbing the system,

O_q : *the quantum notion* (elucidated in 2): measurement of an observable, which is represented by an operator \mathbf{a} , gives as the value of the observable one of the eigenvalues a_μ of \mathbf{a} and leaves the system in the corresponding eigenstate $\mathbf{k}_{\mu\mu}$ (cf. (1.20)); if beforehand the system was in a state \mathbf{k} , the probability of this particular measuring result is $Tr(\mathbf{k}\mathbf{k}_{\mu\mu})$.

Suppose for a moment that the statistical description of quantum mechanics had been proven to be formally of the 1st kind S_f^1 , but with respect to O_q properly of the 2nd kind S_{oq}^2 . Then (if any) the only notion, which could give a proper sense to the formal description, would be

O_u : *the utopian notion*: the uniquely determined processes are observable by methods, hitherto unknown, consistent with and complementary to the methods of O_q .

With respect to quantum theory classical theory is incorrect, though for many purposes it is quite a suitable approximation (for $\lim \hbar \rightarrow 0$). With regard to the utopian conception quantum theory would be correct, but incomplete. In this a description is called correct if none of its statements is in contradiction with observational data. It is called complete if another correct description,

in finding the appropriate form of the required operators \mathbf{a} . This suggests the problem (not further discussed here) whether all or only a certain simple class of operators \mathbf{a} occur in quantum mechanics.

Suppose for a moment that all relevant quantum operators \mathbf{a} had been fixed in one or other way. Then one might ask for a rule $\mathbf{a} \rightarrow a(\phi, q)$, by which the corresponding classical quantities $a(\phi, q)$ are uniquely determined (problem α_2). Problem α_2 would be easily solved in zero order of \hbar , ambiguities might arise in higher order. Now (with respect to O_q) the classical quantities have only a meaning as approximations to the quantum operators for $\lim \hbar \rightarrow 0$. Therefore, whereas in zero order of \hbar it is hardly a problem, in higher order problem α_2 has no observational meaning M_{oq} (with respect to O_q).

Problems α_1 and α_2 could be combined into problem α_3 , asking for a rule of one-to-one correspondence $a(\phi, q) \longleftrightarrow \mathbf{a}$ between the classical quantities $a(\phi, q)$ and the quantum operators \mathbf{a} . Beyond the trivial zero order stage in \hbar , problem α_3 can (with respect to O_q) only have an observational meaning M_{oq} as a guiding principle for detecting the appropriate form of the quantum operators (i.e. as problem α_1). A formal solution of problem α_3 has been proposed by Weyl²⁾ (cf. 4.03). We incidentally come back to problem α_3 in 1.18.

1.10 *Quantum observables.* In this section a will not denote a classical quantity $a(\phi, q)$, but it will stand as a symbol for the observable, which (with regard to O_q) is represented by the quantum operator \mathbf{a} . According to O_q two or more observables a, b, \dots can be simultaneously measured or not, according as the corresponding operators $\mathbf{a}, \mathbf{b}, \dots$ respectively do or do not commute i.e. as they have all eigenstates in common or not. Problem α_4 deals with the (one-to-one) correspondence $a \longleftrightarrow \mathbf{a}$ between the symbols a and the operators \mathbf{a} . Problem α_4 has no sense as long as the symbols a are undefined. They may, however, be implicitly defined just by putting a rule of correspondence. (When the symbols a are identified with the classical quantities $a(\phi, q)$, problem α_4 becomes identical with problem α_3). Von Neumann¹⁾ has proposed the rules

$$\text{if } a \longleftrightarrow \mathbf{a}, \text{ then } f(a) \longleftrightarrow f(\mathbf{a}), \quad \text{I}$$

$$\text{if } a \longleftrightarrow \mathbf{a} \text{ and } b \longleftrightarrow \mathbf{b}, \text{ then } a + b \longleftrightarrow \mathbf{a} + \mathbf{b}. \quad \text{II}$$

$f(\mathbf{a})$ is defined as the operator, which has the same eigenstates as \mathbf{a} with eigenvalues $f(a_\mu)$, where a_μ are those of \mathbf{a} . Then I seems to be obvious. The observable $f(a)$ can be measured simultaneously with

a , its value is $f(a_\mu)$, where a_μ is that of a . When \mathbf{a} and \mathbf{b} commute, $\mathbf{a} + \mathbf{b}$ has the same eigenstates as \mathbf{a} and \mathbf{b} with eigenvalues $a_\mu + b_\mu$, where a_μ and b_μ are those of \mathbf{a} and \mathbf{b} . Then II seems also to be obvious. $a + b$ can be measured simultaneously with a and b , its value is $a_\mu + b_\mu$, where a_μ and b_μ are the values of a and b . When \mathbf{a} and \mathbf{b} do not commute, II is proposed with some hesitation. Because according to O_q the probability of finding a value a_μ for a in a state \mathbf{k} is $Tr(\mathbf{k}\mathbf{k}_{\mu\mu})$ (and because of 1.20)), the expectation value (average value) of a in this state is

$$Ex(a) = \sum_{\mu} Tr(\mathbf{k}\mathbf{k}_{\mu\mu})a_{\mu} = Tr(\mathbf{k}\mathbf{a}) \quad (1.31)$$

and similar for b . If one requires that for a certain pair of observables a and b always

$$Ex(a + b) = Ex(a) + Ex(b), \quad (1.32)$$

one must, because of

$$Tr(\mathbf{k}(\mathbf{a} + \mathbf{b})) = Tr(\mathbf{k}\mathbf{a}) + Tr(\mathbf{k}\mathbf{b}), \quad (1.33)$$

have that

$$Ex(a + b) = Tr(\mathbf{k}(\mathbf{a} + \mathbf{b})). \quad (1.34)$$

Because this has to hold for all states \mathbf{k} , a and b have to satisfy rule II. When II is given up for certain pairs a, b , the additivity of the expectation values of these pairs has also to be given up.

In 4.01 it will be shown that, if I and II shall be generally valid, the symbols a have to be isomorphic with the operators \mathbf{a} . But then there is no reason to introduce the former, their task (if any) can be left to the latter. Accordingly for the sake of brevity we shall henceforth speak of the (quantum) observable \mathbf{a} .

When on the other hand, the symbols a are intended as real commuting quantities, the general validity of I and II cannot be maintained. As long as the symbols a are not further defined, problem α_4 comes to searching for a one-to-one correspondence $a \longleftrightarrow \mathbf{a}$ between the commutative ring of real symbols a and the non-commutative ring of Hermitian operators \mathbf{a} . There may be no, one or more solutions. After the pleas for I and for II, one might be inclined to maintain I and to restrict II. In 1.13 we meet with a particular case (problem α_5) for which II has to be maintained and therefore I has to be restricted. Because we are further exclusively interested in problem α_5 , we will not examine the possibility of solutions for which II is restricted.

(the $\alpha_{\nu\mu}$ are the same as in (1.15)). These relations show, that the functions $a(\xi)$ can be regarded as elements of a (generalized) Hilbert space, in which the $k_{\mu\nu}(\xi)$ form a complete orthonormal system; (1.52) expresses the orthonormality, (1.53) the completeness.

We now show that the correspondence $\mathbf{a} \longleftrightarrow a(\xi)$ has to be a one-to-one correspondence. Suppose for a moment there are operators $\mathbf{k}_{\mu\nu}$ to which there correspond more than one functions $k_{\mu\nu}(\xi)$, which we distinguish by an index ρ , $\mathbf{k}_{\mu\nu} \longleftrightarrow k_{\mu\nu;\rho}(\xi)$. Then the expression

$$\sum_{\mu',\nu'} \int d\xi k_{\mu\nu;\rho}(\xi) k_{\mu';\nu';\rho'}^*(\xi) k_{\mu';\nu';\rho'}(\xi')$$

evaluated with (1.52) gives $k_{\mu\nu;\rho;\rho'}(\xi')$, evaluated with (1.53) it gives $k_{\mu\nu;\rho}(\xi')$. Therefore $k_{\mu\nu;\rho;\rho'}(\xi')$ and $k_{\mu\nu;\rho}(\xi')$ have to be identical. To each operator \mathbf{a} and only to this one there has to correspond one and only one superquantity $a(\xi)$. As a consequence the superquantities $a(\xi)$ must depend on the same number of parameters (at least if they are not too bizarre) as the operators \mathbf{a} , i.e. on twice as many as the wave functions φ .

Thus to each (normalizable) real function $a(\xi)$ and only to this one there corresponds one and only one Hermitian operator \mathbf{a} , which represents an observable quantity (with respect to O_q). In other words every real function $a(\xi)$ is a superquantity. Because this also holds for the (real and imaginary parts of the) parameters ξ themselves, none of them can be hidden in the sense defined above. (An observable quantity may occasionally be inobservable in a measuring device adapted to an incommensurable quantity; in this sense a parameter may occasionally be hidden). In particular all parameters must obey (1.48).

Comparing (1.15) and (1.54) we see that the correspondence $\mathbf{a} \longleftrightarrow a(\xi)$ can be expressed by

$$a(\xi) = Tr(\mathbf{m}(\xi)\mathbf{a}), \quad \mathbf{a} = \int d\xi \mathbf{m}(\xi)a(\xi), \quad (1.55)$$

with

$$\mathbf{m}(\xi) = \sum_{\mu,\nu} \mathbf{k}_{\mu\nu} k_{\mu\nu}^*(\xi); \quad \mathbf{m}^\dagger(\xi) = \mathbf{m}(\xi). \quad (1.56)$$

The Hermitian transformation nucleus $\mathbf{m}(\xi)$ satisfies the relations

$$Tr \mathbf{m}(\xi) = 1, \quad (1.57)$$

$$\int d\xi \mathbf{m}(\xi) = \mathbf{1}; \quad (1.58)$$

$$Tr(\mathbf{m}(\xi) \mathbf{m}(\xi')) = \delta(\xi - \xi'), \quad (1.59)$$

$$\int d\xi Tr(\mathbf{m}(\xi) \mathbf{a}) Tr(\mathbf{m}(\xi) \mathbf{b}) = Tr(\mathbf{ab}) \text{ (for every } \mathbf{a} \text{ and } \mathbf{b}) \quad (1.60)$$

and with (4.01)

$$ab \longleftrightarrow \mathbf{ab} \text{ (for all } a \text{ and } b) \text{ or } ab \longleftrightarrow \mathbf{ba} \text{ (for all } a \text{ and } b). \quad (4.05)$$

This means that the rings are isomorphous.

It follows that, if one ring is commutative and the other not, I and II are inconsistent ⁹⁾. (When the commutators are of the order of \hbar , the discrepancy is according to (4.03) of the order of \hbar^2).

4.02 *Bracket expressions.* Then V' . For the correspondence $a \longleftrightarrow \mathbf{a}$ between the commutative ring with generating elements p and q and the non-commutative ring with generating elements \mathbf{p} and \mathbf{q} with commutator (3.01) ($p \longleftrightarrow \mathbf{p}$ and $q \longleftrightarrow \mathbf{q}$) we show that the rule (cf. 1.18)

if $a(p, q) \longleftrightarrow \mathbf{a}$ and $b(p, q) \longleftrightarrow \mathbf{b}$, then $(a(p, q), b(p, q)) \longleftrightarrow [\mathbf{a}, \mathbf{b}]$ V' is self contradictory.

With

$$p^2 \longleftrightarrow \mathbf{x}_1, q^2 \longleftrightarrow \mathbf{x}_2; p^3 \longleftrightarrow \mathbf{y}_1, q^3 \longleftrightarrow \mathbf{y}_2 \quad (4.06)$$

we find from

$$\begin{aligned} \frac{1}{2} (p^2, q) = p &\longleftrightarrow \frac{1}{2} [\mathbf{x}_1, \mathbf{q}] = \mathbf{p}, \\ \frac{1}{2} (p^2, p) = 0 &\longleftrightarrow \frac{1}{2} [\mathbf{x}_1, \mathbf{p}] = 0 \end{aligned} \quad (4.07)$$

(and similar relations for q^2 and \mathbf{x}_2) that

$$p^2 \longleftrightarrow \mathbf{p}^2 + c_1, q^2 \longleftrightarrow \mathbf{q}^2 + c_2 \quad (4.08)$$

and from

$$\begin{aligned} \frac{1}{3} (p^3, q) = p^2 &\longleftrightarrow \frac{1}{3} [\mathbf{y}_1, \mathbf{q}] = \mathbf{p}^2 + c_1, \\ \frac{1}{3} (p^3, p) = 0 &\longleftrightarrow \frac{1}{3} [\mathbf{y}_1, \mathbf{p}] = 0 \end{aligned} \quad (4.09)$$

(and similar relations for q^3 and \mathbf{y}_2) that

$$p^3 \longleftrightarrow \mathbf{p}^3 + 3c_1\mathbf{p} + d_1, q^3 \longleftrightarrow \mathbf{q}^3 + 3c_2\mathbf{q} + d_2 \quad (4.10)$$

($c_1, c_2; d_1, d_2$ are undetermined constants). Further we get

$$\begin{aligned} \frac{1}{6} (p^3, q^2) = p^2 q &\longleftrightarrow \frac{1}{6} [(\mathbf{p}^3 + 3c_1\mathbf{p} + d_1), (\mathbf{q}^2 + c_2)] = \frac{1}{2} (\mathbf{p}^2\mathbf{q} + \mathbf{q}\mathbf{p}^2) + c_1\mathbf{q}, \\ pq^2 &\longleftrightarrow \frac{1}{2} (\mathbf{p}\mathbf{q}^2 + \mathbf{q}^2\mathbf{p}) + c_2\mathbf{p} \end{aligned} \quad (4.11)$$

and

$$\begin{aligned} \frac{1}{9} (p^3, q^3) = p^2 q^2 &\longleftrightarrow \frac{1}{9} [(\mathbf{p}^3 + 3c_1\mathbf{p} + d_1), (\mathbf{q}^3 + 3c_2\mathbf{q} + d_2)] \\ &= \frac{1}{2} (\mathbf{p}^2\mathbf{q}^2 + \mathbf{q}^2\mathbf{p}^2) + \frac{1}{3} \hbar^2 + c_1\mathbf{q}^2 + c_2\mathbf{p}^2 + c_1c_2. \end{aligned} \quad (4.12)$$

With (4.11) we get

$$\begin{aligned} \frac{1}{3}(\mathbf{p}^2 q, \mathbf{p} q^2) &= \mathbf{p}^2 q^2 \longleftrightarrow \frac{1}{3} [(\frac{1}{2}(\mathbf{p}^2 \mathbf{q} + \mathbf{q} \mathbf{p}^2) + c_1 \mathbf{q}), (\frac{1}{2}(\mathbf{p} \mathbf{q}^2 + \mathbf{q}^2 \mathbf{p}) + c_2 \mathbf{p})] \\ &= \frac{1}{2}(\mathbf{p}^2 \mathbf{q}^2 + \mathbf{q}^2 \mathbf{p}^2) + \frac{2}{3} \hbar^2 - c_1 \mathbf{q}^2 - c_2 \mathbf{p}^2 - \frac{1}{3} c_1 c_2. \end{aligned} \quad (4.13)$$

(4.12) and (4.13) can only be identical for $c_1 = c_2 = 0$ and $\hbar = 0$. Therefore V' is self inconsistent (the deficiency is of the order of \hbar^2).

4.03 *Weyl's correspondence.* And finally III and IV with parameters \mathbf{p} and \mathbf{q} (i.e. for the same rings as in 4.02). We denote the density function by $\rho(\mathbf{p}, \mathbf{q})$. The rules (cf. 1.13)

$$1 \longleftrightarrow \mathbf{1}, \quad \text{III}$$

$$\text{if } a(\mathbf{p}, \mathbf{q}) \longleftrightarrow \mathbf{a} \text{ and } b(\mathbf{p}, \mathbf{q}) \longleftrightarrow \mathbf{b},$$

$$\text{then } \iint d\mathbf{p} d\mathbf{q} \rho(\mathbf{p}, \mathbf{q}) a(\mathbf{p}, \mathbf{q}) b(\mathbf{p}, \mathbf{q}) = \text{Tr}(\mathbf{a}\mathbf{b}) \quad \text{IV}$$

can be satisfied by (1.55)

$$a(\mathbf{p}, \mathbf{q}) = \text{Tr}(\mathbf{m}(\mathbf{p}, \mathbf{q}) \mathbf{a}), \quad \mathbf{a} = \iint d\mathbf{p} d\mathbf{q} \rho(\mathbf{p}, \mathbf{q}) \mathbf{m}(\mathbf{p}, \mathbf{q}) a(\mathbf{p}, \mathbf{q}) \quad (4.14)$$

with a transformation nucleus $\mathbf{m}(\mathbf{p}, \mathbf{q})$, which satisfies (1.57), (1.58); (1.59), (1.60)

$$\text{Tr} \mathbf{m}(\mathbf{p}, \mathbf{q}) = 1, \quad (4.15)$$

$$\iint d\mathbf{p} d\mathbf{q} \rho(\mathbf{p}, \mathbf{q}) \mathbf{m}(\mathbf{p}, \mathbf{q}) = \mathbf{1}; \quad (4.16)$$

$$\text{Tr}(\mathbf{m}(\mathbf{p}, \mathbf{q}) \mathbf{m}(\mathbf{p}', \mathbf{q}')) = \rho^{-1}(\mathbf{p}, \mathbf{q}) \delta(\mathbf{p} - \mathbf{p}') \delta(\mathbf{q} - \mathbf{q}'), \quad (4.17)$$

$$\begin{aligned} \iint d\mathbf{p} d\mathbf{q} \rho(\mathbf{p}, \mathbf{q}) \text{Tr}(\mathbf{m}(\mathbf{p}, \mathbf{q}) \mathbf{a}) \text{Tr}(\mathbf{m}(\mathbf{p}, \mathbf{q}) \mathbf{b}) &= \\ &= \text{Tr}(\mathbf{a}\mathbf{b}) \text{ (for every } \mathbf{a} \text{ and } \mathbf{b}), \end{aligned} \quad (4.18)$$

When we replace in (1.56) the complete orthonormal systems $\mathbf{k}_{\mu\nu}^*(\mathbf{p}, \mathbf{q})$ of (1.54) and $\mathbf{k}_{\mu\nu}$ of (1.15) by the complete orthonormal systems

$$\frac{1}{\hbar} e^{-\frac{i}{\hbar}(x\mathbf{p} + y\mathbf{q})} \text{ of (3.10) and } e^{\frac{i}{\hbar}(x\mathbf{p} + y\mathbf{q})} \text{ of (3.20),}$$

we find a solution

$$\mathbf{m}(\mathbf{p}, \mathbf{q}) = \frac{1}{\hbar} \iint dx dy e^{\frac{i}{\hbar}(x\mathbf{p} + y\mathbf{q})} e^{-\frac{i}{\hbar}(x\mathbf{p} + y\mathbf{q})} \quad (4.19)$$

of (4.15), (4.16); (4.17), (4.18) with the density function

$$\rho(\mathbf{p}, \mathbf{q}) = \frac{1}{\hbar}. \quad (4.20)$$

Then we get for (4.14)

$$a(\mathbf{p}, \mathbf{q}) = \frac{1}{\hbar} \iint dx dy e^{\frac{i}{\hbar}(x\mathbf{p} + y\mathbf{q})} \text{Tr} \left(e^{-\frac{i}{\hbar}(x\mathbf{p} + y\mathbf{q})} \mathbf{a} \right), \quad (4.21)$$

$$\mathbf{a} = \frac{1}{\hbar} \iint dx dy e^{\frac{i}{\hbar}(x\mathbf{p} + y\mathbf{q})} \frac{1}{\hbar} \iint d\mathbf{p} d\mathbf{q} e^{-\frac{i}{\hbar}(x\mathbf{p} + y\mathbf{q})} a(\mathbf{p}, \mathbf{q}).$$

With the Fourier expansions (3.10) and (3.20) this correspondence reads

$$\iint dx dy \alpha(x,y) e^{\frac{i}{\hbar}(xp+yp)} \longleftrightarrow \iint dx dy \alpha(x,y) e^{\frac{i}{\hbar}(xp+yp)}, \quad (4.22)$$

which is Weyl's correspondence²⁾.

II is a consequence of IV and is therefore satisfied by the correspondence (4.21). We will see what is left of I and V'. If $a \longleftrightarrow \mathbf{a}$ and $b \longleftrightarrow \mathbf{b}$ according to (4.21) we find with (3.04)

$$\begin{aligned} \mathbf{ab} &= \frac{1}{\hbar^4} \iint \dots \iint dx dy dx' dy' dp dq dp' dq' \\ &\cdot e^{\frac{i}{\hbar}((x+x')p+(y+y')q)} e^{\frac{i}{2\hbar}(xy'-yx')} e^{-\frac{i}{\hbar}(xp+yp+x'p'+y'q')} a(p,q) b(p',q'). \end{aligned} \quad (4.23)$$

With the variables

$$\begin{aligned} \xi &= x + x', & \eta &= y + y', & \sigma &= \frac{p + p'}{2}, & \tau &= \frac{q + q'}{2}, \\ \xi' &= \frac{x - x'}{2}, & \eta' &= \frac{y - y'}{2}, & \sigma' &= p - p', & \tau' &= q - q', \end{aligned} \quad (4.24)$$

this becomes

$$\begin{aligned} \mathbf{ab} &= \frac{1}{\hbar^4} \iint \dots \iint d\xi d\eta d\xi' d\eta' d\sigma d\tau d\sigma' d\tau' e^{\frac{i}{\hbar}(\xi p + \eta q)} e^{\frac{i}{2\hbar}(-\xi\eta' + \eta\xi')} \\ &\cdot e^{-\frac{i}{\hbar}(\xi\sigma + \eta\tau + \xi'\sigma' + \eta'\tau')} a(\sigma + \frac{1}{2}\sigma', \tau - \frac{1}{2}\tau') b(\sigma - \frac{1}{2}\sigma', \tau + \frac{1}{2}\tau') \\ &= \frac{1}{\hbar^2} \iiint d\xi d\eta d\sigma d\tau e^{\frac{i}{\hbar}(\xi p + \eta q)} e^{-\frac{i}{\hbar}(\xi\sigma + \eta\tau)} \\ &\quad a(\sigma + \frac{1}{4}\eta, \tau - \frac{1}{4}\xi) b(\sigma - \frac{1}{4}\eta, \tau + \frac{1}{4}\xi) \\ &= \frac{1}{\hbar^2} \iiint d\xi d\eta d\sigma d\tau e^{\frac{i}{\hbar}(\xi p + \eta q)} e^{-\frac{i}{\hbar}(\xi\sigma + \eta\tau)} \\ &\quad \left(e^{\frac{1}{4}(\eta \frac{\partial}{\partial \sigma} - \xi \frac{\partial}{\partial \tau})} a(\sigma, \tau) \right) \left(e^{-\frac{1}{4}(\eta \frac{\partial}{\partial \sigma} - \xi \frac{\partial}{\partial \tau})} b(\sigma, \tau) \right). \end{aligned} \quad (4.25)$$

The expressions in brackets at the end are a symbolical representation of Taylor expansion. With the substitution

$$\xi \rightarrow x, \eta \rightarrow y, \sigma \rightarrow p, \tau \rightarrow q \quad (4.26)$$

we get by partial integration

$$\begin{aligned} \mathbf{ab} &= \frac{1}{\hbar} \iint dx dy e^{\frac{i}{\hbar}(xp+yp)} \frac{1}{\hbar} \iint dp dq \\ &\cdot e^{-\frac{i}{\hbar}(xp+yp)} \left(a(p,q) e^{\frac{\hbar}{2i}(\frac{\partial}{\partial p} \frac{\partial}{\partial q} - \frac{\partial}{\partial q} \frac{\partial}{\partial p})} b(p,q) \right). \end{aligned} \quad (4.27)$$

This gives for the Hermitian operators $\frac{1}{2}(\mathbf{ab} + \mathbf{ba})$ and $\frac{i}{2}(\mathbf{ab} - \mathbf{ba})$ the correspondence

$$a(p, q) \cos \frac{\hbar}{2} \left(\frac{\delta}{\delta p} \frac{\partial}{\partial q} - \frac{\delta}{\delta q} \frac{\partial}{\partial p} \right) b(p, q) \longleftrightarrow \frac{1}{2}(\mathbf{ab} + \mathbf{ba}), \quad (4.28)$$

$$a(p, q) \sin \frac{\hbar}{2} \left(\frac{\delta}{\delta p} \frac{\partial}{\partial q} - \frac{\delta}{\delta q} \frac{\partial}{\partial p} \right) b(p, q) \longleftrightarrow \frac{i}{2}(\mathbf{ab} - \mathbf{ba}). \quad (4.29)$$

To the neglect of terms of order of \hbar^2 and higher (4.28) and (4.29) would read

$$a(p, q) b(p, q) \longleftrightarrow \frac{1}{2}(\mathbf{ab} + \mathbf{ba}), \quad (4.30)$$

$$a(p, q) \frac{\hbar}{2} \left(\frac{\delta}{\delta p} \frac{\partial}{\partial q} - \frac{\delta}{\delta q} \frac{\partial}{\partial p} \right) b(p, q) \longleftrightarrow \frac{i}{2}(\mathbf{ab} - \mathbf{ba}). \quad (4.31)$$

(4.30) would lead to I, (4.31) is equivalent to V'.

We examine which functions $f(a)$ satisfy I. From (4.28) we see that the correspondence

$$\text{if } a \longleftrightarrow \mathbf{a}, \text{ then } a^n \longleftrightarrow \mathbf{a}^n \text{ (for every integer } n) \quad (4.32)$$

only holds if

$$a^k \cos \frac{\hbar}{2} \left(\frac{\delta}{\delta p} \frac{\partial}{\partial q} - \frac{\delta}{\delta q} \frac{\partial}{\partial p} \right) a^l = a^{k+l} \text{ (for all integers } k \text{ and } l). \quad (4.33)$$

First take for a a homogeneous polynomial in p and q of degree n . An elementary calculation shows that the condition

$$a \cos \frac{\hbar}{2} \left(\frac{\delta}{\delta p} \frac{\partial}{\partial q} - \frac{\delta}{\delta q} \frac{\partial}{\partial p} \right) a = a^2 \quad (4.34)$$

or

$$a \left(\frac{\delta}{\delta p} \frac{\partial}{\partial q} - \frac{\delta}{\delta q} \frac{\partial}{\partial p} \right)^{2k} a = a^2 \text{ (for } 0 < 2k \leq n) \quad (4.35)$$

is only satisfied if a is of the form $(xp + yq)^n$. Then it follows that any polynomial in p and q can only satisfy (4.33) if it is a polynomial in $xp + yq$. This finally means that I can only be satisfied if a is a function of a certain linear combination $xp + yq$ of p and q . With the help of the Fourier expansion (4.22) it is easily seen that every (normalizable) function of $xp + yq$ does satisfy I. Therefore the least restricted form of I, which is consistent with the correspondence (4.21) is

$$f(xp + yq) \longleftrightarrow f(x\mathbf{p} + y\mathbf{q}). \quad (4.36)$$

As to V' , we see from (4.31) that for the correspondence (4.21) the bracket expression $((a(\phi, q), b(\phi, q)))$ (cf. 1.14) defined by

$$\text{if } a(\phi, q) \longleftrightarrow \mathbf{a} \text{ and } b(\phi, q) \longleftrightarrow \mathbf{b}, \text{ then } ((a(\phi, q), b(\phi, q))) \longleftrightarrow [\mathbf{a}, \mathbf{b}] \quad (4.37)$$

is given by

$$((a(\phi, q), b(\phi, q))) = a(\phi, q) \frac{2}{\hbar} \sin \left(\frac{\hbar}{2} \frac{\delta}{\delta \phi} \frac{\partial}{\partial q} - \frac{\delta}{\delta q} \frac{\partial}{\partial \phi} \right) b(\phi, q). \quad (4.38)$$

If $a(\phi, q)$ or $b(\phi, q)$ is a polynomial in ϕ and q of at most 2nd degree, we have a special case for which the bracket expressions $((a, b))$ and (a, b) coincide.

The correspondence (4.21) is a solution of III and IV. We have not investigated the possibility of other solutions with the same parameters ϕ and q .

5. Quasi-distributions.

5.01 *Proper and improper representations.* With Weyl's correspondence (4.22) as a special solution of

$$\mathbf{1} \longleftrightarrow 1 \quad \text{III}$$

$$\text{if } \mathbf{k} \longleftrightarrow k(\phi, q) \text{ and } \mathbf{a} \longleftrightarrow a(\phi, q),$$

$$\text{then } Tr(\mathbf{k}\mathbf{a}) = \frac{1}{\hbar} \iint d\phi dq k(\phi, q) a(\phi, q) \quad \text{IV}$$

(with parameters ϕ and q and density function $\rho(\phi, q) = 1/\hbar$), we obtain a special case of a transformation between a representation in terms of operators \mathbf{k} and \mathbf{a} and a representation in terms of functions $k(\phi, q)$ and $a(\phi, q)$. Quantum statistics are usually represented in terms of operators, classical statistics in terms of functions. We assert that the usual description is also the proper one. The statistical operator \mathbf{k} of the quantum representation and the statistical distribution function $k(\phi, q)$ of the classical representation are non-negative definite, but in general the quantum $k(\phi, q)$ and the classical \mathbf{k} are not. This makes that for orthogonal states, for which

$$Tr(\mathbf{k}_1 \mathbf{k}_2) = \frac{1}{\hbar} \iint d\phi dq k_1(\phi, q) k_2(\phi, q) = 0, \quad (5.01)$$

the product $\mathbf{k}_1 \mathbf{k}_2$ or $k_1(\phi, q) k_2(\phi, q)$ vanishes in the proper representation, but in the improper representation it need not. The equations

of motion of the quantum \mathbf{k} are described by infinitesimal unitary transformations, those of the classical $k(p, q)$ by infinitesimal canonical transformations (contact transformations), but the equations of motion of the classical \mathbf{k} and the quantum $k(p, q)$ are in general not of these types. Because the improper representation is formally equivalent to the proper one, it is (provided it is not misinterpreted) a correct description, though it is in general a rather impracticable one.

In spite of its deficiencies, or rather because of them, we discuss some aspects of the improper representation of quantum mechanics in terms of $k(p, q)$ and $a(p, q)$, i.e. the quasi-statistical description of the 1st kind Q^1 (cf. 1.19). It more or less illustrates the ways along which some opponents might hope to escape Bohr's reasonings and von Neumann's proof and the places where they are dangerously near breaking their necks.

5.02 *Transition functions.* For the transition functions $k_{\mu\nu}(p, q)$ corresponding to the transition operators (1.03) according to (4.21) we find with the help of the q -representation (occasionally expressing the inner product explicitly by an integral) similar to (3.16)

$$\begin{aligned} k_{\mu\nu}(p, q) &= \frac{1}{\hbar} \iint dx dy e^{\frac{i}{\hbar}(xp+ya)} \int dq' \varphi_{\mu}^{\dagger}(q') e^{\frac{x}{2} \frac{\delta}{\delta q'}} e^{\frac{i}{\hbar} yq} e^{-\frac{x}{2} \frac{\partial}{\partial q'}} \varphi_{\nu}(q') \\ &= \int dx \varphi_{\mu}^{\dagger}(q) e^{\frac{x}{2} \frac{\delta}{\delta q}} e^{\frac{i}{\hbar} xp} e^{-\frac{x}{2} \frac{\partial}{\partial q}} \varphi_{\nu}(q) \\ &= \int dx \varphi_{\mu}^{\dagger}\left(q + \frac{x}{2}\right) e^{\frac{i}{\hbar} xp} \varphi_{\nu}\left(q - \frac{x}{2}\right). \end{aligned} \quad (5.02)$$

Because the wave functions φ_{μ} are only determined but for a factor $e^{i/\hbar \gamma_{\mu}}$ (γ real), the $k_{\mu\nu}(p, q)$ are only determined but for a factor $e^{i/\hbar (\gamma_{\mu} - \gamma_{\nu})}$. The distribution functions, which are thus obtained with Weyl's correspondence ²⁾ become identical to those given by Wigner ¹⁰⁾.

5.03 *Proper value.* In a distribution \mathbf{k} or $k(p, q)$ a quantity \mathbf{a} or $a(p, q)$ can be regarded to have a proper value if the condition (2.10)

$$\text{Tr}(\mathbf{kf}(a)) = f(\text{Tr}(\mathbf{ka})) \quad (5.03)$$

or

$$\frac{1}{\hbar} \iint dp dq k(p, q) f(a(p, q)) = f\left(\frac{1}{\hbar} \iint dp dq k(p, q) a(p, q)\right) \quad (5.04)$$

is satisfied for every f . Whereas the validity of (5.04) is for a proper (non-negative definite) $k(p, q)$ already guaranteed by the validity of the special case $f(a) = a^2$, it is not for a proper \mathbf{k} or an improper $k(p, q)$. For a proper \mathbf{k} the validity of (5.03) or (2.11) requires that \mathbf{a} is of the form

$$a(x\mathbf{p} + y\mathbf{q}) \quad (5.05)$$

and \mathbf{k} an eigenstate of \mathbf{a} . For any $k(p, q)$ the validity of (5.04) requires that $k(p, q)$ is of the form

$$\delta(a(p, q) - a_\mu), \quad (5.06)$$

which is a proper (i.e. non-negative definite) one. Because (5.03) and (5.04) are identical, the conditions (5.05) and (5.06) are equivalent. This means that the eigenstates of the operators $a(x\mathbf{p} + y\mathbf{q})$ and of no other operators correspond with proper (and orthonormal and therefore non-overlapping) distributions of the form (5.06), in which a_μ is the corresponding eigenvalue. This case would be rather encouraging for a statistical description of the 1st kind S^1 , if it were not just an exceptional case.

The eigenfunctions of $a(x\mathbf{p} + y\mathbf{q})$ are in q -representation

$$\begin{aligned} \varphi_\rho(q) &= \frac{1}{\sqrt{x\hbar}} e^{\frac{i}{\hbar} \left(-\frac{1}{2xy} (yq - \rho)^2 + \gamma(\rho)\right)} \quad \text{for } x \neq 0, \\ \varphi_\rho(q) &= \sqrt{y} \delta(yq - \rho) e^{\frac{i}{\hbar} \gamma(\rho)} \quad \text{for } x = 0. \end{aligned} \quad (5.07)$$

($\gamma(\rho)$ real arbitrary). The corresponding eigenvalues are $a(\rho)$

$$a(x\mathbf{p} + y\mathbf{q})\varphi_\rho = a(\rho)\varphi_\rho. \quad (5.08)$$

ρ , which is the eigenvalue of $x\mathbf{p} + y\mathbf{q}$ (for arbitrary fixed x and y), runs between $-\infty$ and $+\infty$. The domain of eigenvalues of $a(x\mathbf{p} + y\mathbf{q})$ is therefore the same as that of the functions $a(z)$ ($-\infty \leq z \leq \infty$). This means that the domain of the proper values of observables, which have such, are unrestricted by quantum conditions.

Inserting the eigenfunctions (5.07) in (5.02) we get

$$k_{\mu\nu}(p, q) = \delta\left(xp + yq - \frac{\rho_\mu + \rho_\nu}{2}\right) e^{-\frac{i}{\hbar} \left(\left(\frac{p}{y} - \frac{q}{x}\right) \frac{\rho_\mu - \rho_\nu}{2} + \gamma(\rho_\mu) - \gamma(\rho_\nu)\right)}. \quad (5.09)$$

(The expression in brackets in the exponent in (5.09) is a canonical

conjugate of $xp + yq$). The $k_{\mu\mu}(\phi, q)$ are actually of the form (5.06).

5.04 *The harmonic oscillator.* After we have treated in 5.03 a special case for which the $k(\phi, q)$ are of proper type themselves, we now deal with a case for which their equations of motion are of proper type. According to (1.43) and condition V' they are if $((H(\phi, q), k(\phi, q)))$ coincides with $(H(\phi, q), k(\phi, q))$ and according to (4.38) this is the case for every $k(\phi, q)$ if $H(\phi, q)$ is a polynomial in ϕ and q of at most 2nd degree. This condition is satisfied for the harmonic oscillator, for which $H(\phi, q)$ coincides with the classical Hamiltonian

$$H(\phi, q) = \frac{\phi^2}{2m} + \frac{m\omega^2}{2} q^2 = \frac{\omega}{2} (\phi'^2 + q'^2); \quad \phi' = \frac{\phi}{\sqrt{m\omega}}, \quad q' = q\sqrt{m\omega}. \quad (5.10)$$

m is the mass, ω the classical circular frequency of the binding. We consider ϕ' and q' as new canonical coordinates and omit the dash.

In q -representation the normalized stationary solutions of the wave equation

$$-\frac{\hbar}{i} \frac{\partial}{\partial t} \varphi_n(q) = \frac{\omega}{2} \left(-\hbar^2 \frac{\partial^2}{\partial q^2} + q^2 \right) \varphi_n(q) \quad (5.11)$$

are

$$\varphi_n(q) = \frac{1}{\sqrt{2^n n!} \sqrt{\pi \hbar}} e^{-\frac{1}{2\hbar} q^2} H_n \left(\frac{q}{\sqrt{\hbar}} \right) e^{-in\omega} \quad (n=0, 1, 2, \dots). \quad (5.12)$$

The Hermitian polynomials $H_n \left(\frac{q}{\sqrt{\hbar}} \right)$ have the generating function

$$e^{\frac{-\xi^2 + 2\xi q}{\hbar}} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{\xi}{\sqrt{\hbar}} \right)^n H_n \left(\frac{q}{\sqrt{\hbar}} \right). \quad (5.13)$$

(5.02) becomes with (5.12)

$$k_{mn}(\phi, q) = \frac{1}{\sqrt{2^{m+n} n! m! \pi \hbar}} \int dx e^{-\frac{1}{2\hbar} (q + \frac{x}{2})^2} H_m \left(\frac{q + \frac{x}{2}}{\sqrt{\hbar}} \right) \cdot e^{\frac{i}{\hbar} xp} e^{\frac{1}{2\hbar} (q - \frac{x}{2})^2} H_n \left(\frac{q - \frac{x}{2}}{\sqrt{\hbar}} \right) e^{-i(m-n)\omega t}. \quad (5.14)$$

0.7 José Enrique Moyal

1 October 1910 – 22 May 1998⁹



J Moyal

Joe Moyal was born in Jerusalem and spent much of his youth in Palestine. He studied electrical engineering in France, at Grenoble and Paris, in the early 1930s. He then worked as an engineer, later continuing his studies in mathematics at Cambridge, statistics at the Institut de Statistique, Paris, and theoretical physics at the Institut Henri Poincaré, Paris.

After a period of research on turbulence and diffusion of gases at the French Ministry of Aviation in Paris, he escaped to London at the time of the German invasion in 1940. The eminent writer C.P. Snow, then adviser to the British Civil Service, arranged for him to be allocated to de Havilland's at Hatfield, where he was involved in aircraft research into vibration and electronic instrumentation.

During the war, hoping for a career in theoretical physics, Moyal developed his ideas on the statistical nature of quantum mechanics, initially trying to get Dirac interested in them, in December 1940, but without success. After substantial progress on his own, his poignant and intense scholarly correspondence with Dirac (Feb 1944 to Jan 1946, reproduced in ^{Moy06}) indicates he was not aware, at first, that his phase-space statistics-based formulation was actually equivalent to standard QM. Nevertheless, he soon appreciated its alternate beauty and power. In their spirited correspondence, Dirac patiently but insistently recorded his reservations, with mathematically trenchant arguments, although lacking essential appreciation of Moyal's novel point of view: A radical departure from the conventional Hilbert space picture ^{Moy49}. The correspondence ended in anticipation of a Moyal colloquium at Cambridge in early 1946.

⁹The material presented here contains statements taken from a previously published obituary, J Gani, "Obituary: José Enrique Moyal", *J Appl Probab* 35 (1998) 1012–1017.

That same year, Moyal's first academic appointment was in Mathematical Physics at Queen's University Belfast. He was later a lecturer and senior lecturer with M.S. Bartlett in the Statistical Laboratory at the University of Manchester, where he honed and applied his version of quantum mechanics ^{BM49}.

In 1958, he became a Reader in the Department of Statistics, Institute of Advanced Studies, Australian National University, for a period of 6 years. There he trained several graduate students, now eminent professors in Australia and the USA. In 1964, he returned to his earlier interest in mathematical physics at the Argonne National Laboratory near Chicago, coming back to Macquarie University as Professor of Mathematics before retiring in 1978.

Joe's interests were broad: He was an engineer who contributed to the understanding of rubber-like materials; a statistician responsible for the early development of the mathematical theory of stochastic processes; a theoretical physicist who discovered the "Moyal bracket" in quantum mechanics; and a mathematician who researched the foundations of quantum field theory. He was one of a rare breed of mathematical scientists working in several fields, to each of which he made fundamental contributions.

9-1-46

Dear Moyal,

I heard from Bartlett that you would be willing to talk about your quantum theory work at our colloquium, and I think it would be a good idea to have it discussed if you do not mind possible heavy criticism. Would Friday the 25th Jan at 3 pm suit you? If this does not leave you sufficient time we could make it a week later. If you cannot conveniently deal with it all in one afternoon there is no objection to your carrying on the following week.

Yours sincerely,

P. A. M. Dirac.

QUANTUM MECHANICS AS A STATISTICAL THEORY

BY J. E. MOYAL

Communicated by M. S. BARTLETT

Received 12 November 1947

I. 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. Whitaker(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)

\mathbf{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 \mathbf{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 \mathbf{s} are not sufficient in themselves to specify the system completely; we need, in addition, another complementary set, say \mathbf{r} , which does not in general commute with \mathbf{s} ; for example, a complete representation is given by either the Cartesian coordinates \mathbf{q} or their conjugate momenta \mathbf{p} , but the complete dynamical specification of the system requires both \mathbf{q} 's and \mathbf{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 \mathbf{G} corresponding to functions $G(r, s)$ of non-commuting observables; the expectation value of \mathbf{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

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

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

experim
to a dist
transfor
concentr
of the sa
case aris
patible
The st
for settir
variables
one ano

We deno
sentatio
and \mathbf{s} to
 \mathbf{r} and \mathbf{s}
muting
 $F(r, s)$ is
statistic
operator

the char

From we
is then

for conti

for discr
The of
and mor

(cf. Kern

† Whe
‡ The
continuo

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 $\exp\{i(\tau\mathbf{r} + \theta\mathbf{s})\}$ (known in statistical terminology as the *characteristic function*). On forming the corresponding operator

$$\mathbf{M}(\tau, \theta) = \exp\{i(\tau\mathbf{r} + \theta\mathbf{s})\} = \sum_n \frac{i^n}{n!} (\tau\mathbf{r} + \theta\mathbf{s})^n, \quad (3.1)$$

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

$$M(\tau, \theta) = (\psi, e^{i(\tau\mathbf{r} + \theta\mathbf{s})} \psi). \quad (3.2)$$

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

$$F(r, s) = \frac{1}{4\pi^2} \iint (\psi, e^{i(\tau\mathbf{r} + \theta\mathbf{s})} \psi) e^{-i(\tau r + \theta s)} d\tau d\theta \quad (3.3)$$

for continuous eigenvalues†, and

$$F(r_i, s_k) = \lim_{T \rightarrow \infty} \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T (\psi, e^{i(\tau\mathbf{r} + \theta\mathbf{s})} \psi) e^{-i(\tau r_i + \theta s_k)} d\tau d\theta \quad (3.4)$$

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} ($\mathbf{pq} - \mathbf{qp} = \hbar/i$),

$$\mathbf{M}(\tau, \theta) = e^{\frac{1}{2}i\hbar\tau\theta} e^{i\theta\mathbf{q}} e^{i\tau\mathbf{p}} = e^{-\frac{1}{2}i\tau\theta} e^{i\theta\mathbf{q}} e^{\frac{1}{2}i\tau\theta} \quad (3.5)$$

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

$$M(\tau, \theta) = \int \psi^*(q - \frac{1}{2}\hbar\tau) e^{i\theta q} \psi(q + \frac{1}{2}\hbar\tau) dq, \quad (3.6)$$

† 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

$$F(p, q) = \frac{1}{2\pi} \int \psi^*(q - \frac{1}{2}\hbar\tau) e^{-i\tau p} \psi(q + \frac{1}{2}\hbar\tau) d\tau, \tag{3.7}$$

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 $\Delta p \Delta q \geq \frac{1}{2}\hbar$ 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

$$m_{lk}(\tau, \theta) = (\psi_l, e^{i(\tau r + \theta s)} \psi_k) = m_{kl}^*(-\tau, -\theta) \tag{4.3}$$

of the operator (3.1) in the representation of the ψ_l . Explicitly, we have

$$f_{lk}(r, s) = \frac{1}{4\pi^2} \iint (\psi_l, e^{i(\tau r + \theta s)} \psi_k) e^{-i(\tau r + \theta s)} d\tau d\theta, \tag{4.4}$$

$$f_{lk}(r_\alpha, s_\beta) = \lim_{T \rightarrow \infty} \frac{1}{4T^2} \int_{-T}^T \int_{-T}^T (\psi_l, e^{i(\tau r + \theta s)} \psi_k) e^{-i(\tau r_\alpha + \theta s_\beta)} d\tau d\theta, \tag{4.5}$$

where (4.4) eigenvalue space of t

and also 1

In the general transform the vector (and self-formation Hermitian

We shall characterize the eigen

In the can be p explicit e

Substitu

The exp

is the eq

† Integ are discre

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†

(3.7)

$$\iint f_{lk}(r, s) f_{l'k'}^*(r, s) dr ds = h^{-1} \delta_{ll'} \delta_{kk'}, \quad (4.6)$$

† in (3.5),

(3.8)

$$\sum_{l, k} f_{lk}(r, s) f_{lk}^*(r', s') = h^{-1} \delta(r - r') \delta(s - s'), \quad (4.7)$$

and also the 'self-orthogonality' relations

$$\iint f_{lk}(r, s) dr ds = \delta_{lk}, \quad (4.8)$$

(3.9)

$$\sum_l f_{ll}(r, s) = h^{-1}. \quad (4.9)$$

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

(3.10)

$$f_{lk}(r, s) = f_{kl}^*(r, s). \quad (4.10)$$

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),

$$f_{lk}(p, q) = \frac{1}{2\pi} \int \psi_l^*(q - \frac{1}{2}\hbar\tau) e^{-i\tau p} \psi_k(q + \frac{1}{2}\hbar\tau) d\tau, \quad (4.11)$$

$$f_{lk}(p, q) = h^{-\frac{1}{2}} e^{\frac{i}{2}(p/q)\hbar} \partial^2/\partial p \partial q [\psi_l^*(q) \phi_k(p) e^{i p q/\hbar}]. \quad (4.12)$$

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

(4.1)

$$f_{p'p''}(p, q) = h^{-1} \delta\left(p - \frac{p' + p''}{2}\right) e^{i q(p'' - p')/\hbar}. \quad (4.13)$$

(4.2)

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

(4.3)

$$\begin{aligned} F(p, q) &= h^{-1} \iint \phi^*(p') \phi(p'') \delta\left(p - \frac{p' + p''}{2}\right) e^{i q(p'' - p')/\hbar} dp'' dp' \\ &= \frac{1}{2\pi} \int \phi^*(p + \frac{1}{2}\hbar\theta) e^{-i\theta q} \phi(p - \frac{1}{2}\hbar\theta) d\theta, \end{aligned} \quad (4.14)$$

(4.4)

is the equivalent of (3.7) in terms of the momentum wave functions $\phi(p)$.

(4.5)

† Integration must be replaced by summation in what follows when the eigenvalues of \mathbf{r}, \mathbf{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

$$\begin{aligned}\bar{G} &= \iint G(r, s) F(r, s) dr ds \\ &= \iiint G(r, s) (\psi, e^{i(\tau r + \theta s)} \psi) e^{-i(\tau r + \theta s)} dr ds d\tau d\theta \\ &= \left(\psi, \left\{ \iint \gamma(\tau, \theta) e^{i(\tau r + \theta s)} d\tau d\theta \right\} \psi \right),\end{aligned}\quad (5.1)$$

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

$$\gamma(\tau, \theta) = \iint G(r, s) e^{-i(\tau r + \theta s)} dr ds. \quad (5.2)$$

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

$$\mathbf{G} = \iint \gamma(\tau, \theta) e^{i(\tau r + \theta s)} d\tau d\theta, \quad (5.3)$$

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 \mathbf{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 $f_{lk}(r, s)$

$$\begin{aligned}G_{lk} &= \iint G(r, s) f_{lk}(r, s) dr ds = \iiint G(r, s) (\psi_l, e^{i(\tau r + \theta s)} \psi_k) dr ds d\tau d\theta \\ &= (\psi_l, \mathbf{G} \psi_k).\end{aligned}\quad (5.4)$$

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

$$\mathbf{G} = e^{\frac{i}{\hbar}(p_0 q_0)} \partial^2 / \partial p \partial q \mathbf{G}_0(\mathbf{q}, \mathbf{p}), \quad (5.5)$$

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 $e^{\frac{i}{\hbar}(p_0 q_0)} \partial^2 / \partial p \partial q$ 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 .

6. T

We now con-
do for this
distributor
As mention
mechanics,
case in the
a brief surv
can be exp
coordinates
The fund
and $F_0(p_0, q$

where K is
temporal *tr*
While F_0 and
dent of the
Hence the a
origin in t , a
time-depend
 K gives t
infinitesimal
 $p-\eta$ conditi

We make th
bability of a
 $t-t_0$. For t
tends to a fi

We shall cal
function at t

then the cha

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.

(5.1)

(5.2)

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

$$F(p, q; t) = \iint K(p, q | p_0, q_0; t - t_0) F_0(p_0, q_0; t_0) dp_0 dq_0, \tag{6.1}$$

(5.3)

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).

(5.4)

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

$$\Lambda(\tau, \theta | \eta, \xi; t - t_0) = \iint e^{i[\theta(q - \xi) + \tau(p - \eta)]} K(p, q | \eta, \xi; t - t_0) dp dq. \tag{6.2}$$

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 $t \rightarrow t_0$

$$\lim_{t \rightarrow t_0} \frac{\Lambda - 1}{t - t_0} = L(\tau, \theta | \eta, \xi). \tag{6.3}$$

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

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

phase-space

our theory. ψ_i can be responding

θ

(5.4) that

c variables words, our ns of non-phase-space

inates and theoretical η is

(5.5)

writing all pplying the . The form momenta, responding

Hence
$$\frac{\partial M}{\partial t} = \lim_{t_0 \rightarrow t} \int \int \frac{\Lambda - 1}{t - t_0} e^{i(\tau\eta + \theta\xi)} F_0(\eta, \xi; t_0) d\eta d\xi$$

$$= \int \int L(\tau, \theta | \eta, \xi) e^{i(\tau\eta + \theta\xi)} F(\eta, \xi; t) d\eta d\xi. \tag{6.5}$$

This can be expressed in the operational form

$$\frac{\partial M}{\partial t} = L\left(\tau, \theta \left| \frac{1}{i} \frac{\partial}{\partial \tau}, \frac{1}{i} \frac{\partial}{\partial \theta} \right.\right) M(\tau, \theta; t) \tag{6.6}$$

(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

$$S(p, q | \eta, \xi) = \int \int L(\tau, \theta | \eta, \xi) e^{i(\tau\eta - p) + i(\theta\xi - q)} d\tau d\theta, \tag{6.7}$$

we obtain for F the integro-differential equation

$$\frac{\partial}{\partial t} F(p, q; t) = \int \int S(p, q | \eta, \xi) F(\eta, \xi; t) d\eta d\xi. \tag{6.8}$$

If, on the other hand, it is possible to expand L in the form

$$L(\tau, \theta | \eta, \xi) = \lim_{t \rightarrow t_0} \int \int \sum_{n=0}^{\infty} \sum_{r=0}^n \frac{(i\tau)^{n-r} (i\theta)^r (p - \eta)^{n-r} (q - \xi)^r}{(n-r)! r!} \frac{K(p, q | \eta, \xi; t - t_0)}{t - t_0} dp dq$$

$$= \sum_{n=0}^{\infty} \sum_{r=0}^n \frac{(i\tau)^{n-r} (i\theta)^r}{(n-r)! r!} \alpha_{nr}(\eta, \xi) \tag{6.9}$$

(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

$$\frac{\partial M}{\partial t} = \int \psi^*(q) [\mathbf{M}, \mathbf{H}] \psi(q) dq = \frac{i}{\hbar} \int \psi^*(q) [\mathbf{M}\mathbf{H} - \mathbf{H}\mathbf{M}] \psi(q) dq, \tag{7.1}$$

where $\mathbf{M}(\tau, \theta)$ is the characteristic function operator (3.5), and \mathbf{H} the Hamiltonian operator, expressed from (5.3) by

$$\mathbf{H} = \int \int W(\sigma, \mu) e^{i(\sigma p + \mu q)} d\sigma d\mu, \tag{7.2}$$

$W(\sigma, \mu)$ is
Hence, $\frac{\partial M}{\partial t}$

Using ex
 $\frac{\partial M}{\partial t} =$

where $\partial/\partial t$
on F . Th

If L poss

$S(p, q | \eta$

then $F(p$

with the

which is

Invers
equation

There is

Finally,

classical

shown by

of two o

$R(p, q)$,

In other

of quant

$W(\sigma, \mu)$ being the Fourier inverse of the corresponding classical Hamiltonian $H(p, q)$. Hence, using expression (3.5), we obtain

$$(6.5) \quad \frac{\partial M}{\partial t} = \frac{i}{\hbar} \iint \int e^{i\hbar(\tau\theta + \sigma\mu)} [e^{i\hbar\sigma\theta} - e^{i\hbar\tau\mu}] W(\sigma, \mu) \psi^*(q) e^{i(\theta + \mu)q} e^{i(\tau + \sigma)p} \psi(q) dq d\sigma d\mu$$

$$(6.6) \quad = \frac{2}{\hbar} \iiint \int \sin \frac{1}{2}\hbar(\tau\mu - \sigma\theta) e^{i[(\tau + \sigma)p + (\theta + \mu)q]} W(\sigma, \mu) \times e^{i(\hbar/2)\partial^2/\partial p\partial q} [\hbar^{-1/2}\psi^*(q)\phi(p) e^{ipq/\hbar}] dp dq d\sigma d\mu.$$

express the functions; distribution

Using expression (3.10) for $F(p, q; t)$, we obtain the two equivalent expressions

$$\frac{\partial M}{\partial t} = \frac{i}{\hbar} \iint [H(p + \frac{1}{2}\hbar\theta, q - \frac{1}{2}\hbar\tau) - H(p - \frac{1}{2}\hbar\theta, q + \frac{1}{2}\hbar\tau)] F(p, q; t) e^{i(\tau p + \theta q)} dp dq, \quad (7.3)$$

(6.7)

$$\frac{\partial M}{\partial t} = \iint e^{i(\tau p + \theta q)} \left\{ \frac{2}{\hbar} \sin \frac{\hbar}{2} \left[\frac{\partial}{\partial p_F} \frac{\partial}{\partial q_H} - \frac{\partial}{\partial p_H} \frac{\partial}{\partial q_F} \right] H(p, q) F(p, q; t) \right\} dp dq, \quad (7.4)$$

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

$$(6.8) \quad L(\tau, \theta | p, q) = \frac{i}{\hbar} [H(p + \frac{1}{2}\hbar\theta, q - \frac{1}{2}\hbar\tau) - H(p - \frac{1}{2}\hbar\theta, q + \frac{1}{2}\hbar\tau)]. \quad (7.5)$$

If L possesses a Fourier transform

$t_0) dp dq$

$$S(p, q | \eta, \xi) = \frac{i}{\hbar} \iint [H(\eta + \frac{1}{2}\hbar\theta, \xi - \frac{1}{2}\hbar\tau) - H(\eta - \frac{1}{2}\hbar\theta, \xi + \frac{1}{2}\hbar\tau)] e^{i(\tau(\eta - p) + \theta(\xi - q))} d\tau d\theta, \quad (7.6)$$

then $F(p, q; t)$ satisfies an integro-differential equation of form (6.8)

(6.9)

$$\frac{\partial}{\partial t} F(p, q; t) = \iint S(p, q | \eta, \xi) F(\eta, \xi; t) d\eta d\xi, \quad (7.7)$$

F satisfies

with the kernel S given by (7.6). Similarly, we find from (7.4)

(6.10)

$$\frac{\partial}{\partial t} F(p, q; t) = \frac{2}{\hbar} \sin \frac{\hbar}{2} \left[\frac{\partial}{\partial p_F} \frac{\partial}{\partial q_H} - \frac{\partial}{\partial p_H} \frac{\partial}{\partial q_F} \right] H(p, q) F(p, q; t), \quad (7.8)$$

terminates,

which is easily shown equivalent to (6.10) with derivate moments

$$\alpha_{2n+1, r}(p, q) = (-1)^{n+r} (\frac{1}{2}\hbar)^{2n} \left(\frac{\partial}{\partial p} \right)^r \left(\frac{\partial}{\partial q} \right)^{2n+1-r} H(p, q), \quad \alpha_{2n, r}(p, q) \equiv 0. \quad (7.9)$$

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.

distributions, and from the

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 $i\hbar[\mathbf{R}\mathbf{G} - \mathbf{G}\mathbf{R}]$ 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

(7.1)

$$\frac{2}{\hbar} \sin \frac{\hbar}{2} \left[\frac{\partial}{\partial p_G} \frac{\partial}{\partial q_R} - \frac{\partial}{\partial p_R} \frac{\partial}{\partial q_G} \right] R(p, q) G(p, q). \quad (7.10)$$

Hamiltonian

(7.2)

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 G corresponding to the classical function $G(p, q)$. Let γ_i be the eigenvalues of G

$$G u_i(q) = \gamma_i u_i(q). \tag{8.4}$$

Calculating the mean of the commutator $[G, M]$ from the two sides of (8.4)

$$\begin{aligned} \int u_i^*(q) [GM - MG] u_k(q) dq &= (\gamma_i^* - \gamma_k) \iint e^{i(\tau p + \theta q)} g_{ik}(p, q) dp dq \\ &= \iint [G(p + \frac{1}{2}\hbar\theta, q - \frac{1}{2}\hbar\tau) - G(p - \frac{1}{2}\hbar\theta, q + \frac{1}{2}\hbar\tau)] g_{ik}(p, q) e^{i(\tau p + \theta q)} dp dq \\ &= \frac{2}{i} \iint e^{i(\tau p + \theta q)} \sin \frac{\hbar}{2} \left[\frac{\partial}{\partial p_g} \frac{\partial}{\partial q_g} - \frac{\partial}{\partial p_G} \frac{\partial}{\partial q_G} \right] G(p, q) g_{ik}(p, q) dp dq, \end{aligned} \tag{8.5}$$

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

$$\begin{aligned} S_G(p, q | \eta, \xi) &= \frac{i}{\hbar} \iint [G(p + \frac{1}{2}\hbar\theta, q - \frac{1}{2}\hbar\tau) - G(p - \frac{1}{2}\hbar\theta, q + \frac{1}{2}\hbar\tau)] e^{i\tau(\eta - p) + \theta(\xi - q)} d\tau d\theta \\ &= 2\pi i \sum_{i, k} (\gamma_i^* - \gamma_k) g_{ik}(p, q) g_{ik}^*(\eta, \xi). \end{aligned} \tag{8.7}$$

† This question was raised by the referee.

Having de-
the transfo

We introdu

for an arbi
functions a

where

On substit
 g_{ik} and γ_{ik} ,

The expans

λ_i

and similar

We have th
 g_{ik} and γ_{ik} ;

$K(p_2, q_2 | p_1,$

The transfo

cesses satis

Hostinsky (

The ener

(9.4) and (9

J

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

$$c_{kn} = \alpha_{kn}^* \alpha_{kn} = \iiint K(p, q | p_0, q_0; t) f_{kk}(p_0, q_0) f_{nn}(p, q) dp_0 dq_0 dp dq. \quad (10.4)$$

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

$$\frac{\partial F}{\partial t} + \frac{2}{\hbar} \sin \frac{\hbar}{2} \left\{ \frac{\partial}{\partial p}, \frac{\partial}{\partial q} \right\} H(p, q) F(p, q; t) = 0, \quad (11.1)$$

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

$$\frac{\partial F}{\partial t} + \left\{ \frac{\partial}{\partial p}, \frac{\partial}{\partial q} \right\} H(p, q) F(p, q; t) = 0. \quad (11.2)$$

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

$$\int_{S_0} F(p_0, q_0) dp_0 dq_0 = \int_{S_t} F(p, q; t) dp dq. \quad (11.3)$$

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

$$\int_{S_0} F(p_0, q_0) dp_0 dq_0 \neq \int_{S_t} F(p, q; t) dp dq. \quad (11.4)$$

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 $\hbar \rightarrow 0$, 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).

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

$$\left. \begin{aligned} \mu(q_1 q_2) &= \overline{q_1 q_2} - \overline{q_1} \overline{q_2} = \gamma \sum_{i,k} n_i n_k |Q_{nk}|^2, \\ \mu(p_1 p_2) &= \overline{p_1 p_2} - \overline{p_1} \overline{p_2} = \gamma \sum_{i,k} n_i n_k |P_{nk}|^2, \end{aligned} \right\} \quad (14.8)$$

where Q_{nk} , P_{nk} are the matrices of the individual q 's and 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 n th 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)

distribution
distribution

It is usual
 g_n when the
W corresponds
operator G
and hence

Now it is essential
condition is
variables r
distribution
of G , then
 $G = G(ar +$

In order
 ψ admits a
non-negative
one cyclic
transforma
coordinates

The transform

Separating

Comparing
of the form

Hence, if
suggested

Finally,
classical
observable
tions repre
sets, and
phase-space
momenta,

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 G is exactly equal to its eigenvalue g_n when the system is in the corresponding eigenstate. This means that the operator W corresponding to the function $W(G)$ should be equal to the function W of the operator G , $W = W(G)$, since if G is exactly equal to g_n the mean of W is $\bar{W} = W(g_n)$, and hence
$$\bar{W} = (\psi_n, W\psi_n) = (\psi_n, W(G)\psi_n) = (\psi_n, W(g_n)\psi_n) = W(g_n). \quad (15.1)$$

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

$$\frac{\partial H}{\partial \theta} = 0, \quad \frac{\partial H}{\partial g} = \text{constant} = \omega. \quad (15.2)$$

The transformed equation of the motion (7.8) can be written

$$\frac{\partial F}{\partial t} + \omega \frac{\partial F}{\partial \theta} + \frac{2}{\hbar} \sin \frac{\hbar}{2} \left\{ \frac{\partial}{\partial P_i}, \frac{\partial}{\partial Q_i} \right\} H F = 0. \quad (15.3)$$

Separating the variables, we have

$$\left. \begin{aligned} F(g, \theta, P_i, Q_i; t) &= F_1(\theta, t) F_r(g, P_i, Q_i), \\ \frac{1}{F_1} \left(\frac{\partial F_1}{\partial t} + \omega \frac{\partial F_1}{\partial \theta} \right) &= 2i\mu \quad (\mu \text{ constant}), \\ F_1 &= e^{i\mu(t+\theta/\omega)}. \end{aligned} \right\} \quad (15.4)$$

Comparing with the expansion of F in energy eigenfunctions, we see that it must be of the form
$$F(g, \theta, P_i, Q_i; t) = \sum_{i, k} a_i^* a_{ik} Q_{ik}(g, P_i, Q_i) e^{i(E_i - E_k)(t+\theta/\omega)/\hbar}. \quad (15.5)$$

Hence, if $F \geq 0$ 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

probability
absence of
no particles

(14.8)

ie particles

ORY
g quantum
urn to the
utations of

egative as
(generally
dynamical
uch as the
ordinates
nth energy
nfunction,
This is not
re those of
ntal equa-
d it is also
lves. Pro-
s of these

istributions
sical inter-
of § 3, can
ent of the
hase-space
example,
, q) corre-
id is hence
able (p, q)

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

$$\frac{\partial M}{\partial t} = \left(\psi^*, \frac{i}{\hbar} [M, H] \psi \right), \quad (15.6)$$

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

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
discussion,
never be in

It is nat
this subject
to those of
of coordinat
measureme
some hope
the phase
hope to ve
if experim
of q and p

Appendix

The spa
be obtaine

where $\rho(q$
 p conditio

On writin
the logar

$K(\tau | q) =$

leads to
Taylor e

† Note
giving an
‡ The
the distri

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

(A 3.3)

According to (5.2) and (5.3), the operator corresponding to $G(ar+bs)$, where a and b are constants, is

$$G = \iint e^{i(\tau r + \theta s)} d\tau d\theta \iint G(ar+bs) e^{-i(\tau r + \theta s)} dr ds. \tag{A 5.1}$$

Changing to the variables

(A 3.4)

$$\xi = ar+bs, \quad \eta = ar-bs, \quad \lambda = \frac{\tau}{2a} + \frac{\theta}{2b}, \quad \mu = \frac{\tau}{2a} - \frac{\theta}{2b}, \tag{A 5.2}$$

(A 3.5)

we find

$$G = \iint e^{i[(\lambda+\mu)ar + (\lambda-\mu)bs]} d\lambda d\mu \iint G(\xi) e^{-i(\lambda\xi + \mu\eta)} d\xi d\eta$$

$$= \int e^{i\lambda(ar+bs)} d\lambda \int G(\xi) e^{-i\lambda\xi} d\xi = G(ar+bs). \tag{A 5.3}$$

e equation
conditional
and making
the case of
miltonian is

(A 4.1)

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).

8) and sub-

SUMMARY

(A 4.2)

(7.8) by p_k

(A 4.3)

4.2), we find

(A 4.4)

bi equation

(A 4.5)

;(A 4.2) and

(A 4.6)

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) KERMAK, 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

1. When taking plane values as is finally reached. The stress is ill-defined the time there is a steady undergoes expression of its structure.

2. The studied, a broken up parallel to one sheet time the u. As the mechanical elementar a crystal broken at distorted process of of length the unit of the amount. Glide of of s. The sends the continuity here by t. the dimension.

3. The Taylor cor plane. A d above the

THE EXACT TRANSITION PROBABILITIES OF QUANTUM-MECHANICAL OSCILLATORS CALCULATED BY THE PHASE-SPACE METHOD

BY M. S. BARTLETT AND J. E. MOYAL

Received 14 September 1948

1. INTRODUCTION

The calculation by the usual perturbation methods of transition probabilities between the unperturbed states of a quantum-mechanical system yields approximate results, valid only for small perturbations. The object of this paper is to calculate the *exact* transition probabilities between the unperturbed states of quantum oscillators, valid for *large* as well as small perturbations, by using the 'phase-space' method developed by one of the authors (Moyal (1), referred to henceforth as (I)).

We give first the main results of (I) required in this paper. The probability distribution in phase-space of a system in a state described by the wave-function $\psi(q)$ in q -space is†

$$F(p, q) = \frac{1}{2\pi} \int \psi^*(q - \frac{1}{2}\hbar\tau) e^{-i\tau p} \psi(q + \frac{1}{2}\hbar\tau) d\tau. \quad (1.1)$$

Corresponding to the expansion of $\psi(q; t)$ in terms of energy eigenfunctions $u_n(q)$

$$\psi(q; t) = \sum_n a_n u_n(q) e^{-iE_n t/\hbar}, \quad (1.2)$$

we have an expansion for $F(p, q; t)$

$$F(p, q; t) = \sum_{k, n} a_k^* a_n f_{kn}(p, q) e^{i(E_k - E_n)t/\hbar} \quad (1.3)$$

in terms of the energy phase-space eigenfunctions

$$f_{kn}(p, q) = \frac{1}{2\pi} \int u_k^*(q - \frac{1}{2}\hbar\tau) e^{-i\tau p} u_n(q + \frac{1}{2}\hbar\tau) d\tau. \quad (1.4)$$

These functions form a complete orthogonal system in phase-space which is also 'self-orthogonal' and hermitian with respect to the indices k, n , i.e.

$$\iint f_{kn} f_{k'n'}^* dp dq = h^{-1} \delta_{kk'} \delta_{nn'}, \quad \iint f_{kn} dp dq = \delta_{kn}, \quad f_{kn} = f_{nk}^*. \quad (1.5)$$

Furthermore, the matrix G_{kn} corresponding to an ordinary function $G(p, q)$ is given by

$$G_{kn} = \iint G(p, q) f_{kn}(p, q) dp dq. \quad (1.6)$$

The transformation with time of $F(p, q; t)$ corresponding to the quantum equations of the motion is given symbolically by

$$\frac{\partial}{\partial t} F(p, q; t) = \frac{2}{\hbar} \sin \left\{ \frac{\hbar}{2} \left(\frac{\partial}{\partial p_F} \frac{\partial}{\partial q_H} - \frac{\partial}{\partial p_H} \frac{\partial}{\partial q_F} \right) \right\} H(p, q) F(p, q; t), \quad (1.7)$$

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

where $\partial/\partial p_H, \partial/\partial q_H$ operate only on the classical Hamiltonian $H(p, q)$ of the system; $\partial/\partial p_F, \partial/\partial q_F$ on $F(p, q, t)$. This is seen to be an extension of Liouville's theorem

$$\frac{\partial F}{\partial t} = \frac{\partial H}{\partial q} \frac{\partial F}{\partial p} - \frac{\partial H}{\partial p} \frac{\partial F}{\partial q}, \tag{1.8}$$

and reduces to the latter in the correspondence principle limit ($\hbar \rightarrow 0$) and for systems whose Hamiltonian is a polynomial of the 2nd degree or less in p and q . For such systems (and they include the free and the uniformly accelerated particle, and the oscillator) the transformation with time of $F(p, q; t)$ is of the 'deterministic' type of classical kinetic theory, each element of the distribution transforming in phase-space according to the laws of classical mechanics. A direct verification of the deterministic character of these systems is given in Appendix 1 and in § 2.†

Equation (1.7) specifies the infinitesimal transformation with time of $F(p, q; t)$. The transformation over a finite interval $t - t_0$ can be given in terms of a 'transformation function' $K(p, q | p_0, q_0; t - t_0)$;

$$F(p, q; t) = \iint K(p, q | p_0, q_0; t - t_0) F(p_0, q_0; t - t_0) dp_0 dq_0. \tag{1.9}$$

K is interpreted as the probability of p, q at t conditional in p_0, q_0 at t_0 , and may be expressed either in terms of the phase-space eigenfunctions by an expansion similar to (1.3)

$$K(p, q | p_0, q_0; t - t_0) = \hbar \sum_{k, n} f_{kn}(p_0, q_0) f_{kn}^*(p, q) e^{i(E_k - E_n)(t - t_0)/\hbar}, \tag{1.10}$$

or in terms of the wave transformation-function

$$\psi(q | q_0; t - t_0) = \sum_n u_n(q_0) u_n^*(q) e^{-iE_n(t - t_0)/\hbar}, \tag{1.11}$$

by an integral similar to (1.1)

$$K(p, q | p_0, q_0; t - t_0) = \frac{\hbar}{2\pi} \iint \psi^*(q - \frac{1}{2}\hbar\tau | q_0 - \frac{1}{2}\hbar\tau_0) e^{i(\tau_0 p_0 - \tau p)} \psi(q + \frac{1}{2}\hbar\tau | q_0 + \frac{1}{2}\hbar\tau_0) d\tau d\tau_0. \tag{1.12}$$

The phase-space theory of quantum mechanics may be applied to calculate the transition probabilities of a perturbed system. If K is known for the perturbed Hamiltonian H , and we wish to calculate the transition probabilities from the k th unperturbed state in an interval t , we take as initial distribution the k th diagonal phase-space eigenfunction corresponding to the unperturbed Hamiltonian $H^{(0)}$: $F_0(p_0, q_0) = f_{kk}^{(0)}(p_0, q_0)$. The transformed distribution after an interval t is then from (1.9) and (1.3)

$$F_{(k)}(p, q; t) = \iint K(p, q | p_0, q_0; t) f_{kk}^{(0)}(p_0, q_0) dp_0 dq_0 = \sum_{l, n} a_l^* a_n f_{ln}^{(0)}(p, q). \tag{1.13}$$

The transition probability from state k to state n is then given exactly by‡

$$p_{kn}(t) = a_n^*(t) a_n(t) = \hbar \iint F_{(k)}(p, q; t) f_{nn}^{(0)}(p, q) dp dq. \tag{1.14}$$

† Cf., in this connexion, Coulson and Rushbrooke (7).

‡ It may be shown that the transition probabilities obtained in this way are in fact identical with those of ordinary quantum theory; the proof of this statement is given in Appendix 2.

In the j
reduce to
classical r
at t follow
 $F_{(k)}(p, q; t$

We now c
angular fr
 $q = (m\omega/\hbar$

The energ

The phase

$f_{kn}(p, q) =$

$= (-1)^r$

$= (-1)^r$

$= (-1)^r$

or, on intr

A more co

The phase

Laguerre f

$f_{kn}(u$

The diagon

† An exp
wold (2).

In the particular case of deterministic systems the transformation function K must reduce to a product of delta-functions expressing the contact transformation of classical mechanics (this is verified in § 2 for the harmonic oscillator). The distribution at t follows simply by substituting the classical solutions in the initial distribution: $F_{(w)}(p, q; t) = f_{kk}^{(0)}[p_0(p, q, t), q_0(p, q, t)]$.

2. PHASE-SPACE THEORY OF THE HARMONIC OSCILLATOR

We now develop the phase-space theory of the one-dimensional oscillator of mass m , angular frequency ω , coordinate Q and momentum P . In terms of the reduced variables $q = (m\omega/\hbar)^{1/2} Q$, $p = (m\omega\hbar)^{-1/2} P$, its Hamiltonian is

$$H = \frac{1}{2}(P^2/m + m\omega^2 Q^2) = \frac{1}{2}(p^2 + q^2)\hbar\omega. \tag{2.1}$$

The energy eigenfunctions in p - and q -space are $u_n(q)$, $u_n(p)$, where

$$u_n(q) = (-1)^n (2^n \pi^{1/2} n!)^{-1/2} e^{1/2 q^2} \left(\frac{\partial}{\partial q}\right)^n (e^{-q^2}) = (2^n \pi^{1/2} n!)^{-1/2} e^{-1/2 q^2} H_n(q), \tag{2.2}$$

$$u_n(p) = \frac{1}{\sqrt{(2\pi)}} \int u_n(q) e^{-ipq} dq. \tag{2.3}$$

The phase-space energy eigenfunctions are then, from (1.4),†

$$\begin{aligned} f_{kn}(p, q) &= \frac{1}{2\pi} \int u_n^*(q - \frac{1}{2}\tau) e^{-i\tau p} u_n(q + \frac{1}{2}\tau) d\tau \\ &= (-1)^{n+k} (2\pi)^{-1} (2^{n+k} \pi n! k!)^{-1/2} \int e^{-i\tau p + \frac{1}{2}(a-\frac{1}{2}\tau)^2 + \frac{1}{2}(a+\frac{1}{2}\tau)^2} \left(\frac{\partial}{\partial q_1}\right)^k (e^{-(a_1-\frac{1}{2}\tau)^2}) \left(\frac{\partial}{\partial q_2}\right)^n (e^{-(a_2+\frac{1}{2}\tau)^2}) d\tau \end{aligned} \tag{1.10}$$

$$\begin{aligned} &= (-1)^{n+k} (2\pi)^{-1} (2^{n+k} \pi n! k!)^{-1/2} \left(\frac{\partial}{\partial q_1}\right)^k \left(\frac{\partial}{\partial q_2}\right)^n e^{a^2 - q_1^2 - a_2^2} \int e^{-i\tau p - \frac{1}{2}\tau^2 + \tau(a_1 - a_2)} d\tau \\ &= (-1)^{n+k} (2^{n+k} \pi^2 n! k!)^{-1/2} e^{p^2 + a^2} \left(\frac{\partial}{\partial q_1}\right)^k \left(\frac{\partial}{\partial q_2}\right)^n (e^{-2(a_1 - ip)(a_2 + ip)}), \quad (q_1 = q_2 = q), \end{aligned} \tag{1.11}$$

or, on introducing the variables $z = q + ip$, $z^* = q - ip$

$$f_{kn}(z, z^*) = (-\sqrt{2})^{n+k} (\pi^2 n! k!)^{-1/2} e^{zz^*} \left(\frac{\partial}{\partial z^*}\right)^k \left(\frac{\partial}{\partial z}\right)^n (e^{-2zz^*}). \tag{2.4}$$

A more convenient expression is obtained by substituting the variables

$$w = 2zz^* = 2(p^2 + q^2) = 4H/\hbar\omega, \quad \theta = \tan^{-1}(p/q).$$

The phase-space eigenfunctions then break up into the products of an associated Laguerre function of w and a trigonometric function of θ

$$\begin{aligned} f_{kn}(w, \theta) dw d\theta &= \frac{1}{2} (-1)^k (k! n!)^{-1/2} w^{1/2(k-n)} e^{1/2 w} \left(\frac{\partial}{\partial w}\right)^k (w^n e^{-w}) dw (2\pi)^{-1} e^{i(k-n)\theta} d\theta \\ &= \frac{1}{2} (-1)^k (k! / n!)^{1/2} w^{-1/2(k-n)} e^{-1/2 w} L_k^{n-k}(w) dw (2\pi)^{-1} e^{i(k-n)\theta} d\theta. \end{aligned} \tag{2.5}$$

The diagonal eigenfunctions are simple Laguerre functions

$$\begin{aligned} f_{nn}(w, \theta) dw d\theta &= \frac{1}{2} (-1)^n (n!)^{-1} e^{1/2 w} \left(\frac{\partial}{\partial w}\right)^n (w^n e^{-w}) dw d\theta / 2\pi \\ &= \frac{1}{2} (-1)^n e^{-1/2 w} L_n(w) dw d\theta / 2\pi. \end{aligned} \tag{2.6}$$

† An expression for these eigenfunctions has been found independently by Dr H. I. Groenewold (2).

the system;
rem
(1.8)
for systems
ich systems
e oscillator)
of classical
e according
ic character
f $F(p, q; t)$.
transforma-
(1.9)
and may be
sion similar
(1.10)
(1.11)
 $\frac{1}{2}\hbar\tau_0) d\tau d\tau_0$.
(1.12)
calculate the
rbed Hamil-
mperturbed
space eigen-
= $f_{kk}^{(0)}(p_0, q_0)$.
3)
, q). (1.13)
by†
(1.14)
fact identical
pendix 2.

The determinism of the oscillator may be verified by calculating the phase-space transformation function over a finite interval of time t , on using equation (1.12). The wave transformation function over this interval is given by the well-known Mehler expansion for Hermite polynomials (3)

$$\begin{aligned} \psi(q | q_0; t) &= \sum_n u_n(q_0) u_n^*(q) e^{-i(n+\frac{1}{2})\omega t} \\ &= (2\pi i \sin \omega t)^{-\frac{1}{2}} \exp \{i(2 \sin \omega t)^{-1} [(q^2 + q_0^2) \cos \omega t - 2qq_0]\}. \end{aligned} \tag{2.7}$$

The phase-space transformation function is then

$$\begin{aligned} K(p, q | p_0, q_0; t) &= (4\pi^2 \sin \omega t)^{-1} \iint \exp [i\{\tau_0 p_0 - \tau p + (\sin \omega t)^{-1} [(\tau q + \tau_0 q_0) \cos \omega t - \tau q_0 - \tau_0 q]\}] d\tau d\tau_0 \\ &= (4\pi^2 \sin \omega t)^{-1} \iint \exp \{i(\sin \omega t)^{-1} [\tau(q \cos \omega t - p \sin \omega t - q_0) \\ &\quad + \tau_0(q_0 \cos \omega t + p_0 \sin \omega t - q)]\} d\tau d\tau_0 \\ &= \sin \omega t \delta(q \cos \omega t - p \sin \omega t - q_0) \delta(q_0 \cos \omega t + p_0 \sin \omega t - q), \end{aligned} \tag{2.8}$$

a product of delta functions expressing the contact transformation of classical mechanics for the oscillator. The expansion (1.10) of K in terms of the phase-space eigenfunctions gives the following interesting formula for the associated Laguerre functions

$$\begin{aligned} (32\pi)^{-1} \sum_{k,n} (k! / n!) w_0^{-\frac{1}{2}(k-n)} e^{-\frac{1}{2}w_0} L_k^{n-k}(w_0) w^{-\frac{1}{2}(k-n)} e^{-\frac{1}{2}w} L_k^{n-k}(w) e^{i(n-k)(\theta - \theta_0 - \omega t)} \\ = \delta(w - w_0) \delta(\theta - \theta_0 - \omega t + 2r\pi), \end{aligned} \tag{2.9}$$

where the δ -functions are normalized over the ranges of w and θ , r being an integer such that $0 \leq \theta_0 + \omega t - 2r\pi \leq 2\pi$.

In Appendix 3, the above results are applied to derive the equilibrium phase-space distribution of members of a statistical assembly of oscillators.

3. TRANSITION PROBABILITIES OF A PERTURBED OSCILLATOR

We shall now apply the method outlined in § 1 to calculate the exact transition probabilities of a perturbed oscillator for a perturbing potential of the form $V = q\mathcal{E}(t)$, where $\mathcal{E}(t)$ is an arbitrary function of the time. On using equation (1.14), the transition probability from state k to state n is given in terms of the variables $w = 2(p^2 + q^2)$ and $\theta = \tan^{-1}(p/q)$ used in § 2 by

$$p_{kn}(t) = 8\pi \int_0^\infty \int_0^{2\pi} f_{kk}(w_0, \theta_0) f_{nn}^*(w, \theta) dw_0 d\theta_0, \tag{3.1}$$

where w_0, θ_0 are the initial values at $t = 0$, w and θ those at time t obtained from the classical solution for the perturbed oscillator

$$q = q_0 \cos \omega t + p_0 \sin \omega t - \omega \int_0^t \mathcal{E}(\tau) \sin \omega(t - \tau) d\tau,$$

$$p = p_0 \cos \omega t - q_0 \sin \omega t - \omega \int_0^t \mathcal{E}(\tau) \cos \omega(t - \tau) d\tau.$$

Hence

$$w = w_0 + 2(w_0 \alpha)^{\frac{1}{2}} \cos(\theta_0 - \phi) + \alpha, \tag{3.2}$$

where

$$(\frac{1}{2}\alpha)^{\frac{1}{2}} e^{i\phi} = i\omega \int_0^t \mathcal{E}(\tau) e^{i\omega\tau} d\tau.$$

It is easily seen that the perturbation function $\mathcal{E}(t)$ is expanded in powers of α . The expansion of the transition probability $p_{kn}(t)$ in powers of α yields a closed form for $p_{kn}(t)$.

$$\begin{aligned} p_{kn}(t) &= \frac{(-1)^k}{k!n!} \\ &= \frac{(-1)^k}{k!n!2} \\ &= \frac{(-1)^k}{k!n!} \end{aligned}$$

where $I_0(x)$ is the Bessel function of the first kind. The integrand in the expansion can be written as

$$p_{kn}(t) = \frac{(-1)^k}{k!n!} I_0(x)$$

The integrand in the expansion can be written as

$$p_{kn}(t) = \frac{(-1)^k}{k!n!} I_0(x)$$

or symbolically

$$p_{kn}(t) = \frac{(-1)^k}{k!n!} I_0(x)$$

With the aid of the generating function

$$\begin{aligned} G(\tau, \theta) &= \sum_{k,n} p_{kn}(t) e^{i(k\theta - n\theta_0 - \omega t)} \\ &= \frac{1}{1 - \dots} \end{aligned}$$

$$= \frac{1}{1 - \dots}$$

The coefficient of $e^{i(k\theta - n\theta_0 - \omega t)}$ in the p.g.f. $G_k(\tau)$ is the transition probability $p_{kn}(t)$ of both θ and θ_0 .

$$p_{kn}(t) = \dots$$

It is easily seen that $\epsilon = \frac{1}{2}\hbar\omega\alpha$ is the non-fluctuating part of the work done by the perturbing force, while ϕ is the phase change of the oscillator. We now substitute these solutions in (3.1) and also introduce two auxiliary variables γ, ξ which we shall equate to 1 after all the indicated integrations and differentiations have been carried out. This yields a closed expression for the transition probability

$$\begin{aligned}
 (2.7) \quad p_{kn}(t) &= \frac{(-1)^{k+n}}{k!n!2\pi} \int_0^\infty \int_0^{2\pi} e^{\frac{1}{2}(w_0+w)} \left(\frac{\partial}{\partial w_0}\right)^k (w_0^k e^{-\gamma w_0}) \left(\frac{\partial}{\partial w}\right)^n (w^n e^{-\xi w}) dw_0 d\theta_0 \\
 &= \frac{(-1)^{k+n}}{k!n!2\pi} \left(\frac{\partial}{\partial \gamma}\right)^k \left(\frac{\partial}{\partial \xi}\right)^n \gamma^k \xi^n \int_0^\infty \int_0^{2\pi} \exp\{w_0(1-\gamma-\xi) \\
 &\quad + (\frac{1}{2}-\xi)[\alpha+2(w_0\alpha)^{\frac{1}{2}}\cos(\theta_0-\phi)]\} dw_0 d\theta_0 \\
 &= \frac{(-1)^{k+n}}{k!n!} \left(\frac{\partial}{\partial \gamma}\right)^k \left(\frac{\partial}{\partial \xi}\right)^n \gamma^k \xi^n e^{\frac{1}{2}(\gamma-\xi)\alpha} \int_0^\infty e^{-w_0(\gamma+\xi-1)} I_0[2(w_0\alpha)^{\frac{1}{2}}(\frac{1}{2}-\xi)] dw_0 \quad (\gamma = \xi = 1),
 \end{aligned}$$

where $I_0(x)$ is the modified zero-order Bessel function of the first kind ($I_0(x) = J_0(ix)$). The integration is easily carried out, for example, by substituting for $I_0(x)$ its Taylor expansion and integrating term by term, whence

$$(2.8) \quad p_{kn}(t) = \frac{(-1)^{k+n}}{k!n!} \left(\frac{\partial}{\partial \gamma}\right)^k \left(\frac{\partial}{\partial \xi}\right)^n \left(\frac{\gamma^k \xi^n}{\gamma + \xi - 1} \exp\left[-\alpha \frac{(\gamma - \frac{1}{2})(\xi - \frac{1}{2})}{\gamma + \xi - 1}\right]\right) \quad (\gamma = \xi = 1). \quad (3.3)$$

Carrying out the indicated differentiation on γ^k, ξ^n , one finds that

$$(2.9) \quad p_{kn}(t) = (-1)^{k+n} \sum_{\nu=0}^k \binom{k}{\nu} \frac{1}{\nu!} \left(\frac{\partial}{\partial \gamma}\right)^\nu \sum_{\mu=0}^n \binom{n}{\mu} \frac{1}{\mu!} \left(\frac{\partial}{\partial \xi}\right)^\mu \left(\frac{1}{\gamma + \xi - 1} \exp\left[-\alpha \frac{(\gamma - \frac{1}{2})(\xi - \frac{1}{2})}{\gamma + \xi - 1}\right]\right) \quad (\gamma = \xi = 1)$$

or symbolically

$$(3.4) \quad p_{kn}(t) = (-1)^{k+n} L_k\left(-\frac{\partial}{\partial \gamma}\right) L_n\left(-\frac{\partial}{\partial \xi}\right) \left(\frac{1}{\gamma + \xi - 1} \exp\left[-\alpha \frac{(\gamma - \frac{1}{2})(\xi - \frac{1}{2})}{\gamma + \xi - 1}\right]\right) \quad (\gamma = \xi = 1).$$

With the use of the expression for the Laguerre polynomials generating function (3)

$$(3.5) \quad \frac{1}{1-\tau} \exp\left[-\frac{\tau x}{1-\tau}\right] = \sum_{n=0}^\infty L_n(x) \tau^n,$$

a 'probability generating function' (p.g.f.) may be calculated for the p

$$G(\tau, \theta) = \sum_{k,n} p_{kn} \theta^k \tau^n$$

$$\begin{aligned}
 (3.1) \quad &= \frac{1}{(1+\tau)(1+\theta)} \exp\left(-\frac{\theta}{1+\theta} \frac{\partial}{\partial \gamma}\right) \exp\left(-\frac{\tau}{1+\tau} \frac{\partial}{\partial \xi}\right) \\
 &\quad \left\{ \frac{1}{\xi + \gamma - 1} \exp\left(-\alpha \frac{(\gamma - \frac{1}{2})(\xi - \frac{1}{2})}{\gamma + \xi - 1}\right) \right\} \quad (\gamma = \xi = 1) \\
 (3.2) \quad &= \frac{1}{1-\tau\theta} \exp\left\{-\frac{\epsilon}{\hbar\omega} \frac{(1-\tau)(1-\theta)}{1-\tau\theta}\right\}. \quad (3.6)
 \end{aligned}$$

The coefficient of θ^k in the Taylor expansion of $G(\tau, \theta)$ in powers of θ only, will be the p.g.f. $G_k(\tau)$ for transitions from the k th state; that of $\theta^k \tau^n$ in the expansion in powers of both θ and τ will be $p_{kn}(t)$. For $n \geq k$

$$(3.7) \quad p_{kn}(t) = \frac{1}{(n-k)!} \left(\frac{\epsilon}{\hbar\omega}\right)^{n-k} e^{-\epsilon/\hbar\omega} \sum_{\mu=0}^k \sum_{\nu=0}^{k-\mu} \left(\frac{\epsilon}{\hbar\omega}\right)^{\mu+2\nu} \times \left(\frac{(n-k)!(n+\nu)!(-2)^\mu}{\mu! \nu! (n-k+\nu)! (k-\mu-\nu)! (n-k+\mu+2\nu)!}\right). \quad (3.7)$$

Since, as is obvious from (3.3) or (3.4), the p_{kn} satisfy the principle of detailed balance, i.e. $p_{kn} = p_{nk}$, the probability of transitions to states $n < k$ is simply obtained by interchanging the indices n and k in the right-hand side of (3.7).

The p.g.f. $G_0(\tau)$ and the probabilities $p_{0n}(t)$ for transitions from the ground state are simply those of a Poisson distribution

$$G_0(\tau) = e^{(\epsilon/\hbar\omega)(\tau-1)}, \quad p_{0n}(t) = \frac{1}{n!} \left(\frac{\epsilon}{\hbar\omega}\right)^n e^{-\epsilon/\hbar\omega}. \tag{3.8}$$

These exact results may now be compared with the approximate ones obtained by the standard perturbation method. The Taylor expansion of the approximate expression for the p.g.f.

$$G(\tau, \theta) \sim \frac{1}{1-\tau\theta} \left[1 - \frac{\epsilon}{\hbar\omega} \frac{(1-\tau)(1-\theta)}{1-\tau\theta} \right] \tag{3.9}$$

gives approximate values for the

$$p_{kn}(t) = \delta_{k,n} + \frac{\epsilon}{\hbar\omega} [(k+1)\delta_{k+1,n} + k\delta_{k-1,n} - (2k+1)\delta_{k,n}] \tag{3.10}$$

correct to the first power in ϵ , whose expression is identical with the first approximation of the perturbation method. Expression (3.7) shows that in general the probability of a 2nth-pole transition for small ϵ is of the order of $(\epsilon/\hbar\omega)^n$, i.e. only dipole transitions have an appreciable probability. The perturbation method equates to 1 the exponential factor $\exp\{-\epsilon/\hbar\omega\}$ in the exact expression (3.7). This procedure is justified only for small ϵ ; as the perturbation energy increases, however, multipole transitions become progressively more probable. In order to find the most probable ones from the ground state, let us substitute the continuous variable x for n in (3.8)

$$p_{0x} = \frac{1}{x!} \left(\frac{\epsilon}{\hbar\omega}\right)^x e^{-\epsilon/\hbar\omega},$$

$$\frac{dp_{0x}}{dx} = \frac{1}{x!} \left(\frac{\epsilon}{\hbar\omega}\right)^x e^{-\epsilon/\hbar\omega} \left[\log\left(\frac{\epsilon}{\hbar\omega}\right) - \frac{d}{dx}(\log x!) \right].$$

The most probable transition is therefore to the state n given by

$$\frac{d}{dx}(\log x!)_{(x=n)} = \psi(n) \sim \log(\epsilon/\hbar\omega),$$

where the logarithmic derivative $\psi(x)$ of the factorial function $x!$ is an increasing function of x (Jahnke-Emde (4)). For large x , $\psi(x) \sim \log x$, and hence $n\hbar\omega = E_n - E_0 \sim \epsilon$. Similar considerations apply to the general case: it will be seen from (3.7) that the most probable transitions are those from states k to states n such that $(n-k)\hbar\omega = E_n - E_k \sim \epsilon$. Hence the physically plausible result that for large perturbations the most probable transitions will be those for which the change in energy is approximately equal to the work done by the perturbing forces.

APPENDIX

(1) Free particle and particle under constant force

The way in which the phase-space method may be used to solve wave-packet problems is easily exemplified in the deterministic cases of the free particle and the particle under

constant force
momentum s

correspondin
(cf. Kennard
and taking t
distribution is l

The solution
for p and q in
distribution at :

It is easily sl

which is the

with $\phi_0(p_0)$ a
Similarly,
tribution $F_0(\phi_0)$

which again

of the Schröd

(2

It is conv
 $\langle s_1 | q_1 \rangle$ for tl
for state s_2 at
for the trans
bability in t
calculated by

P_{s_1}

constant force. In both these cases, the Schrödinger equation is most easily solved in momentum space. We describe the particle at $t = 0$ by a Gaussian wave-packet

$$\phi_0(p_0) = (2\pi s_0^2)^{-1/2} \exp[-p_0^2/4s_0^2] \tag{A.1.1}$$

corresponding to the minimum uncertainty $\sigma_0 s_0 = \frac{1}{2}\hbar$ on its position and momentum (cf. Kennard (5)), s_0 and σ_0 being respectively the mean square deviations of p_0 and q_0 and taking the origin at the mean of p_0 and q_0 . The corresponding phase-space distribution is by (1.1)

$$F_0(p_0, q_0) = \frac{1}{2\pi} \int \phi_0^*(p_0 - \frac{1}{2}\hbar\tau) e^{i\tau q_0} \phi_0(p_0 + \frac{1}{2}\hbar\tau) d\tau$$

$$= (\pi\hbar)^{-1} \exp[-\frac{1}{2}(p_0^2/s_0^2 + q_0^2/\sigma_0^2)]. \tag{A.1.2}$$

The solution is then obtained simply by substituting in (A.1.2) the classical solutions for p and q in terms of p_0, q_0, t . In the case of the free particle, this gives for the distribution at t

$$F(p, q; t) = (\pi\hbar)^{-1} \exp[-\frac{1}{2}(p/s_0)^2 - \frac{1}{2}(q/\sigma_0 - pt/m\sigma_0)^2]. \tag{A.1.3}$$

It is easily shown that this corresponds by (1.1) to the wave-function

$$\phi(p; t) = (2\pi s_0^2)^{-1/2} \exp[-(p/2s_0)^2 + ip^2t/2m\hbar], \tag{A.1.4}$$

which is the solution of the Schrödinger equation

$$\frac{p^2}{2m}\phi = -\frac{\hbar}{i}\frac{\partial\phi}{\partial t} \tag{A.1.5}$$

with $\phi_0(p_0)$ as initial wave-function at $t = 0$ (cf. Darwin (6)).

Similarly, for the particle under constant force mg we find for the same initial distribution $F_0(p_0, q_0)$

$$F(p, q; t) = \frac{1}{\pi\hbar} \exp\left[-\frac{1}{2}\left\{\left(\frac{p-mgt}{s_0}\right)^2 + \left(\frac{q-pt/m-\frac{1}{2}gt^2}{\sigma_0}\right)^2\right\}\right] \tag{A.1.6}$$

which again corresponds to the solution

$$\phi(p; t) = (2\pi s_0^2)^{-1/2} \exp\left[-\left\{\left(\frac{p-mgt}{2s_0}\right)^2 + \frac{i}{\hbar}\left(\frac{p^2t}{2m} - \frac{gpt^2}{2} + \frac{mg^2t^3}{6}\right)\right\}\right] \tag{A.1.7}$$

of the Schrödinger equation
$$\frac{p^2}{2m}\phi + \frac{mg\hbar}{i}\frac{\partial\phi}{\partial p} = -\frac{\hbar}{i}\frac{\partial\phi}{\partial t}. \tag{A.1.8}$$

(2) *Equivalence between the transition probabilities calculated by the phase-space method and those of standard quantum theory*

It is convenient in the calculations that follow to introduce Dirac's notation: $\langle s_1 | q_1 \rangle$ for the coordinate eigenfunction corresponding to state s_1 , at time t_1 , $\langle s_2 | q_2 \rangle$ for state s_2 at t_2 , $\langle q_1 | q_2 \rangle$ for the wave transition function from q_1 at t_1 to q_2 at t_2 , $\langle s_1 | s_2 \rangle$ for the transition matrix from s_1 at t_1 to s_2 at t_2 ; the corresponding transition probability in the standard theory is then $|\langle s_1 | s_2 \rangle|^2$. The transition probability $p_{s_1 s_2}$, calculated by the phase-space method, is from (1.13) and (1.14)

$$p_{s_1 s_2} = \int \dots \int f_{s_1 s_2}^*(p_2, q_2) K(p_2, q_2 | p_1, q_1) f_{s_1 s_2}(p_1, q_1) dp_2 dq_2 dp_1 dq_1. \tag{A.2.1}$$

Substituting from expressions (1.4) for $f_{ss}(p, q)$ in terms of $\langle s | q \rangle$ and (1.12) for K in terms of $\langle q_1 | q_2 \rangle$, we have

$$\begin{aligned} p_{s_1 s_2} &= \frac{\hbar^2}{4\pi^2} \int \dots \int \left\{ \langle s_1 | q_1 + \frac{1}{2}\hbar\tau_1 \rangle e^{-i\tau_1 p_1} \langle q_1 - \frac{1}{2}\hbar\tau_1 | s_1 \rangle d\tau_1 \right. \\ &\quad \times \int \langle q_1 + \frac{1}{2}\hbar\theta_1 | q_2 + \frac{1}{2}\hbar\theta_2 \rangle e^{i(\theta_1 p_1 - \theta_2 p_2)} \langle q_2 - \frac{1}{2}\hbar\theta_2 | q_1 - \frac{1}{2}\hbar\theta_1 \rangle d\theta_1 d\theta_2 \\ &\quad \left. \times \int \langle s_2 | q_2 - \frac{1}{2}\hbar\tau_2 \rangle e^{i\tau_2 p_2} \langle q_2 + \frac{1}{2}\hbar\tau_2 | s_2 \rangle d\tau_2 \right\} dp_1 dq_1 dp_2 dq_2 \\ &= \iint \langle s_1 | x_1 \rangle \langle x_1 | x_2 \rangle \langle x_2 | s_2 \rangle dx_1 dx_2 \times \iint \langle s_2 | x'_2 \rangle \langle x'_2 | x'_1 \rangle \langle x'_1 | s_1 \rangle dx'_1 dx'_2, \end{aligned}$$

where the change of variables $x_1 = q_1 + \frac{1}{2}\hbar\tau_1$, $x'_1 = q_1 - \frac{1}{2}\hbar\tau_1$, etc., has been made. Hence

$$p_{s_1 s_2} = |\langle s_1 | s_2 \rangle|^2, \quad (\text{A. 2.2})$$

i.e. the two expressions for the transition probabilities are formally equivalent.

(3) Phase-space distribution of a member of a statistical assembly of oscillators

The equilibrium distribution of the coordinates and momenta of one member of a statistical assembly of similar particles is expressed in I, p. 114, as a sum of diagonal phase-space energy eigenfunctions relative to the individual particles

$$f(p, q) = C \sum_k n_k f_{kk}(p, q), \quad (\text{A. 3.1})$$

where C is a normalizing constant and the $u_n = e^{-\epsilon_n/kT}$ in the case of a Maxwell-Boltzmann assembly. On substituting from (2.6) for an assembly of oscillators, this becomes

$$\begin{aligned} f(w, \theta) &= (4\pi)^{-1} (1 - e^{-\hbar\omega/kT}) \sum_n (-1)^n e^{-\frac{1}{2}w - n\hbar\omega/kT} L_n(w) \\ &= (4\pi)^{-1} \tanh(\hbar\omega/2kT) \exp\left\{-\frac{1}{2}[w \tanh(\hbar\omega/2kT)]\right\}, \end{aligned} \quad (\text{A. 3.2})$$

where the last line follows from (3.5). Transforming back from w and θ to p and q , we finally find a Gaussian distribution for the coordinates and momenta of a member of an assembly of oscillators

$$f(p, q) = (2\pi)^{-1} (\hbar\omega/E) \exp\left\{-\frac{1}{2}[(p^2 + q^2)(\hbar\omega/E)]\right\}, \quad (\text{A. 3.3})$$

where E is the mean energy

$$E = \frac{1}{2}\hbar\omega \coth(\hbar\omega/2kT) = \hbar\omega(e^{\hbar\omega/kT} - 1)^{-1} + \frac{1}{2}\hbar\omega. \quad (\text{A. 3.4})$$

We may note that the mean-square deviation of the energy calculated from the above distribution is $\sigma_E^2 = E^2$ instead of the usual expression $\sigma_E^2 = E^2 - (\frac{1}{2}\hbar\omega)^2$. This follows from the fact that the phase-space theory of quantum mechanics yields different distributions according to the basic system of variables chosen (see I, p. 100 for a fuller discussion). If the energy is one of these variables, then it is quantized, with possible values $\epsilon_k = (k + \frac{1}{2})\hbar\omega$, and its distribution is simply

$$\rho(\epsilon_k) = 2 \sinh(\hbar\omega/2kT) e^{-\epsilon_k/kT}, \quad (\text{A. 3.5})$$

yielding the second value $\sigma_E^2 = E^2 - (\frac{1}{2}\hbar\omega)^2$ for the m.s. deviation of the energy. If p and q are chosen as the basic system, then the energy distribution becomes continuous, with a m.s. deviation of $(\frac{1}{2}\hbar\omega)^2$ for each of the energy eigenfunctions, leading thus to the first value $\sigma_E^2 = E^2$.

The 'phase-
the transitic
results obta
by an expo
done by the
transitions
however, th
probable, th
mately equa

- (1) MOYAL,
- (2) GROENJ
- (3) SZEGÖ,
- (4) JAHNKI
- (5) KENNA
- (6) DARWI
- (7) COULSO

DEPARTMEN
UNIVERSIT

SUMMARY

The 'phase-space' method in quantum theory is used to derive *exact* expressions for the transition probabilities of a perturbed oscillator. Comparison with the approximate results obtained by perturbation methods shows that the latter must be multiplied by an exponential factor $\exp(-\epsilon/\hbar\omega)$, where ϵ is the non-fluctuating part of the work done by the perturbing forces; as long as ϵ is small, $\exp(-\epsilon/\hbar\omega) \sim 1$ and only dipole transitions have an appreciable probability. As the perturbation energy increases, however, this is no longer true, and multipole transitions become progressively more probable, the most probable ones being those for which the change in energy is approximately equal to the work done by the perturbing forces.

REFERENCES

- (1) MOYAL, J. E. *Proc. Cambridge Phil. Soc.* 45 (1949), 99-124.
- (2) GROENEWOLD, H. J. *Physica*, 12 (1946), 405-60.
- (3) SZEGÖ, G. *Orthogonal polynomials* (New York, 1939), pp. 97, 102.
- (4) JAHNKE, E. and EMDE, F. *Tables of functions* (3rd ed.) (Leipzig, 1938), p. 19.
- (5) KENNARD, E. H. *Z. Phys.* 44 (1927), 326.
- (6) DARWIN, S. G. *Proc. Roy. Soc. A*, 117 (1928), 258-93.
- (7) COULSON, C. A. and RUSHBROOKE, G. S. *Proc. Cambridge Phil. Soc.* 42 (1946), 286-91.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MANCHESTER

for K in

$d\theta_1 d\theta_2$

$\frac{1}{2}$,

Hence

(A. 2.2)

lent.

ators

ember of a
of diagonal

(A. 3.1)

well-Boltz-
is becomes

(A. 3.2)

p and q , we
member of

(A. 3.3)

(A. 3.4)

and from the
 $(\frac{1}{2}\hbar\omega)^2$. This
leads different
for a fuller
with possible

(A. 3.5)

energy. If
continuous,
leading thus to

MAVERICK MATHEMATICIAN

The Life and Science
of J.E. Moyal

Ann Moyal

Appendix 2. P.A.M. Dirac – J. E. Moyal: Correspondence, 1944-1946. Basser Library, Australian Academy of Science, Canberra, MS 45/3/[1]

Professor P.A.M. Dirac,
St. John's College,
CAMBRIDGE

3, Sandy Rise,
Wigston,
Leics,
February 18, 1944.

Dear Professor Dirac,

Professor Fowler has sent me a copy of his letter to Dr. Bartlett, in which he writes of his discussion with you and Dr. Jeffreys regarding the possibilities of a statistical basis for quantum mechanics.

He suggests I should discuss the matter with you sometime, and I should be glad to do so if you can spare the time.

You will remember no doubt we talked about this in December 1940, when I was beginning to consider these ideas.

Yours sincerely,
[J.E. Moyal]

7 Cavendish Avenue,
CAMBRIDGE,
21.2.44

Dear Moyal,

I should be glad to meet you any week-end.

On Saturdays I have a lecture from 12–1 and a fire watching in the evening, but apart from that I could meet you any time on Saturday or Sunday. So choose any week-end you like. The most convenient time for me would be Saturday morning at about 10.30 or 11, when I am in the Arts School, but if this is too early for you, would you come round to my house Saturday afternoon or Sunday?

Yours sincerely,
P.A.M. Dirac

7 Cavendish Avenue,
CAMBRIDGE,
6.3.44

Dear Moyal,

I should be glad to see you Saturday afternoon the 11th. If you come around 2.30 it would do very well.

I have enclosed a reprint I have just received from Whittaker, which you may care to read as it deals with the point at issue.

Yours sincerely,
P.A.M. Dirac

3 Sandy Rise,
WIGSTON.
Leics.
June 26th, 1944.

Dear Professor Dirac,

On thinking over the objection you raised when I last saw you, to my statistical treatment of quantum Mechanics, it has occurred to me that the difficulties are chiefly a question of interpretation. I think the theory can be rendered acceptable by abandoning the idea, taken over from the original (Bohr) quantum mechanics, that eigenstates have an objective reality.

One of the difficulties of the theory is that the probability distributions obtained for p and q from single eigenfunctions, can take negative values except perhaps for the ground state. Only linear superpositions of eigenfunctions lead to defined positive probability distributions in phase-space. Now, as I explained in my paper, I consider the form I obtained for the phase-space distribution $F(p,q)$ as in a way an extension, or rather, an exact form of Heisenberg's principle of uncertainty, in the sense that it imposes not only the well-known inequality for the dispersions of p and q , but a special form for their whole probability distribution. Perhaps then, the fact that phase-space distributions corresponding to single eigenstates can take negative values may be interpreted as meaning that an isolated conservative atomic or molecular system in a single eigenstate is a thing that cannot generally be observed without contradicting this generalised principle of uncertainty. I think this can be conceded, and no doubt physical arguments could be brought forward to support such a view. Only statistical assemblies and distributions corresponding to linear superpositions of eigenfunctions such that $F(p,q,t)$ is always greater than zero would be observable, and would have an objective reality.

If this is accepted, it then ceases to have a meaning — to talk about a system having exact values of energy, momentum etc. in a given eigenstate, so that the second difficulty, i.e. the fact that the theory does not necessarily lead to such values, also disappears. The only thing that has a physical meaning is the working out of the final statistical distributions over a number of states, representing the results of experiments. I think that in this way my theory may be reconciled with the usual form of quantum mechanics, and may possibly lead to new results capable of experimental verifications.

The interpretation of spectra, for example, would be obtained in the usual manner from the mean values of electric dipole moments, leading to the same results as the ordinary theory. The physical notation of quantum jumps must be abandoned. The possible frequencies of the spectral lines are exhibited in the expansion of phase-space distribution at time t in terms of the phase-space eigenfunctions for $f_{jk}(p, q)$

$$F(p, q, t) = \sum c_i^*(t) c_i(t) f_{jk}(p, q) e^{i h(E_i - E_k) t / \hbar}$$

The forbidden lines drop out, of course, on averaging of $F(p, q, t)$. A more refined interpretation would involve extending the theory to radiation and its inter-action with matter.

With regard to the Stern-Gerlach experiments, I should like to quote from C.G. Darwin's paper 'Free Motions in Wave Mechanics', *Proc. Roy. Soc. A.* 117 (1928) p. 260: 'in the Stern-Gerlach experiment, we do not say that the field splits the atom into two groups and then separates these. We say that a wave goes through the field, and when we calculate its intensity at the terminal plate, we find that it has two maxima which we then interpret as two patches of atoms.' This shows that the theory of the Stern-Gerlach experiment may be tackled by ordinary wave methods, without necessarily postulating exact eigenvalues for the angular momenta, and in fact, Darwin gives this theory in the same paper, on page 284. Actually, the treatment of such dynamical problems involving the evolution with time of wave packets may be simplified by the use of the methods developed in my paper, as I have shown for Darwin's treatment of the free and uniformly accelerated electron, where in addition to his results, I also obtained the joint distribution for p and q .

In fact, I regard such dynamical problems as one case where the methods outlined may have an advantage over the usual methods. Furthermore, as the theory leads to the distributions at phase-space, and also to correlations at two instants of time, there is a possibility it may lead to experimental verifications in the field of electron and molecular beams. Another field where I think the theory may be of some value is in the study of statistical assemblies, since it leads to phase-space distributions for p and q , (equivalent to the Maxwell-Boltzmann distribution) for Fermi-Dirac and Bose-Einstein assemblies. This may be of value in the kinetic theory of non-uniform fluids.

I should like now to submit to you a few ideas of a more speculative nature. In the theory as I have developed it in my paper, a combination of the transformation equations for $\psi(p)$, $\varphi(p)$ with Newtonian mechanics, leads to Schrödinger's equation and ordinary quantum mechanics. As I mentioned in the course of our conversation, substantially the same transformation equations combined with the mechanical equations of special relativity lead to the Klein-Gordon equation. One would expect new forms of quantum mechanics (such as your spinor equations for the electron) to appear from the combination of new transformation equations with the mechanical equations. As long, however, as these mechanical equations, whether classical or relativistic, are deterministic, the form of quantum mechanics obtained will be deterministic for isolated systems, and therefore unsatisfactory for nuclear theory. This is, I think, a further argument in favour of the idea that a satisfactory quantum theory of the nucleus must be based on some form of unitary theory involving the electro-magnetic field in a fundamental manner, since one would expect then the mechanical equations for a particle to be non-deterministic because it would never be isolated from the infinity of degrees of freedom of the radiation field.

May I take this opportunity of thanking you and Mrs. Dirac for your very kind hospitality on my last visit to Cambridge.

[J.E. Moyal]

19.3.45.

Dear Moyal,

Some work I have been doing lately is connected with your work on a joint probability distribution $F(p, q, t)$ and has led me to think that there may be a limited region of validity for the use of a joint probability distribution. However, I have rather forgotten the details of your work and would be glad if you could let me see again the part of it dealing with $F(p, q, t)$. I may get a more favourable opinion of it this time. Have you done any more work on it since our previous correspondence?

Yours sincerely,
P.A.M. Dirac

3, Sandy Rise,
WIGSTON.
Leics.
March 22nd, 1945.

Dear Professor Dirac,

Unfortunately, my paper is in the hands of Professor Chapman of the Imperial College, and I only have the one typescript. However, I have sent your letter to him with a request that he should forward you the paper as soon as he has finished with it. On the other hand, I have just heard from M.S. Bartlett, that he is back at Queen's; he is pretty familiar with my work, and I feel sure he will give you any explanation that you may require, if you care to get in touch with him, especially as he has worked out a new and improved method for obtaining the joint distributions.

I notice you have used Fock's operators in your paper on 'Quantum Electrodynamics'. I have been wondering whether the work to which you refer in your letter is connected with this, as these operators imply in a way eigenfunctions in phase-space. I thought I could see a way of tying it up with my work when I was reading your paper, but I did not get very far with it.

I am afraid I have not done very much since I last wrote to you, as my engineering work is keeping me pretty busy. However, I have worked out the relativistic extension for scalar wave-functions, which leads to the wave equation

$$\frac{1}{\hbar m_0} \sum (p_i - e A_i)^2 \psi = -\frac{\hbar}{i} \frac{\partial \psi}{\partial s}$$

where s is the local time of the particle. This is a 'time dependent' extension of the Klein-Gordon equation; I do not know whether it has been considered before. There is a difference in the interpretation, however. I take s as the independent variable and the space time co-ordinates, and the momentum energy vector as random functions of s . The ordinary probability distribution which is then a scalar in space-time, is given as in the non-relativistic theory by $\rho(q, s) = \psi \psi^*$. The joint-phase space distribution for co-ordinate-time and momentum-energy is obtained in terms of ψ as in the non-relativistic theory, and gives in the same way for the space-time conditional means of the momentum-energy vector

$$\rho \bar{p}_i = \frac{\hbar}{n_i} \left(\psi \frac{\partial \psi^*}{\partial q_i} - \psi^* \frac{\partial \psi}{\partial q_i} \right)$$

This is normal for the current density, but connects the charge density with the space-time conditional mean value of the energy, rather than with probability, giving thus an immediate interpretation of the negative values obtained when the energy eigenvalues are negative.

I think this interpretation of probability as a scalar in space-time is perhaps more satisfactory than as the time-component of a ψ -vector, though there is a conceptual difficulty, since ρ must then be considered as variable with the local time of the particle. Another difficulty is connected with the relation

$$\sum (p_i - e A_i)^2 = m_0^2$$

which would restrict the phase-space probability distribution to a 7-dimensional hyper surface. One way of turning this difficulty would be to consider m_0 itself as a random variable, perhaps capable of taking a number of eigenvalues — but all this is purely speculative. I am not really clear about the last part.

In collaboration with M.S. Bartlett, I have also carried further the treatment of the harmonic oscillator in phase-space. Some of the results are rather reminiscent of those you obtain with the ξ -operator. This work is fairly complete, and I should be able to let you have a typescript of it shortly, if you are interested.

I have also been considering applications to statistical mechanics, which, since they require distributions in phase-space, would seem to offer an obvious field to the theory. But apart from equilibrium distributions, I rather hope that the application of the theory of random functions, will also lead to methods generally suitable for non-uniform states and fluctuation problems.

3, Sandy Rise,
WIGSTON,
Leics.
March 23rd, 1945

Dear Professor Dirac, [\[2\]](#)

My letter to Professor Chapman yesterday, crossed with his, returning my typescript which I therefore enclose.

I also enclose the typescript of a note by M.S. Bartlett, which gives an improved method of obtaining the joint distribution.

ATTACHMENT:

Comments on Your Letter to Professor Dirac, 26.6.44 by M.S. Bartlett.

1. General Validity

The practical issue here seems to be simply this:-

Either

- i. Your theory is equivalent to the orthodox (non-relativistic) theory as regards all possible physical experiments (cf. the earlier equivalence of the Heisenberg and Schrödinger methods). The method used is then simply a matter of convenience, though it would be a great advantage to possess a firmer logical basis for the methods in current use.
- ii. Or your theory is not so equivalent. In that case acceptance or rejection is firstly a matter of experiment; but again since your theory is more rigorous than the standard, there should be better scope for modification of the particular physical postulates it contains.

2. Eigenstates

It also seems clear now that the analysis into eigenstates is a matter of mathematical technique. This is supported by:-

- a. The appearance of negative probabilities in the phase-space eigenfunctions. But apart from this appeal to your theory, we may note
- b. Equivalent expansions in different coordinates (e.g. the free electron in polar coordinates by Rejansky).
- c. The use of eigenfunctions as a general method of solving differential equations, the use of Fourier series being the best known example.
- d. The appearance of eigenfunctions in 'chain probability' problems. Re this point, Jeffreys' work is relevant, but I think the elementary algebra of wave vectors (of the kind often used in introductory textbooks on quantum theory) indicates rather more simply that in some respects (analogously to (c)) the technique is quite general and has nothing to do with quantum theory as such. The relevant algebra is developed in the attached notes.*

3. Discrete energy levels

The remarks under 1. are in sympathy with your view that here it is meaningless to ask whether the energy levels are really discrete, but to ask whether the theoretical spectra are correct. Incidentally one might note that while there is no objection to a conceptual discrete energy level existing over infinite time (as I pointed out in my reply to a previous comment of yours), it is true that in practice the observation of a spectra over a finite time implies a blurring of the lines. This is recognized and a theory has been worked out (see, for example the early chapters of Rosseland's Theoretical Astrophysics). This observational fact may tend to obscure any finer points on the energy level distributions.

Similarly with the Stern-Gerlach effect — it is a matter for agreement with experiment — though here I shall not try to comment since I believe this effect involves electron spin, with which your theory does not deal.

4. Interference and diffraction

Similarly also with these phenomena. There is a word of caution here. When I looked at this a little while ago in an attempt to determine as precisely as possible from observe[d] results the form of the uncertainty principle, I satisfied myself that the interference of protons and electrons after passing through two narrow slits will not arise if the latter are merely passively filtering a statistical assembly of particles with an initial distribution of position and momentum; it is essential to allow the uncertainty principle to imply an actual change in the momentum possibility distribution consequent on the positional probability distribution at the slits.

Compare the discussion by Whittaker (Proc. Phys. Soc. 55, p. 464, 1933) of polarisation of Nicol prisms. He asserted that this phenomenon was impossible to explain by any what he called 'crypto-deterministic' mechanism, citing an alleged proof by von Neumann of this. But it was clear that he was referring to a deterministic behaviour of the protons without interaction with the prism; and this point has been taken up by Pelzer (Proc. Phys. Soc. 56, p. 195, 1944), who shows that with such interaction Whittaker's assertion is not necessarily true.

This means, however, in connection with your suggestion of experimental verification with electron beams, that in successive measurements taken on a beam of photons or electrons, the effect of each measurement must be allowed for, and this will presumably affect the observed correlations at two instants of time.

5. Reversibility

The reference in the last paragraph of your letter to Dirac to nuclear theory was extremely interesting, though I think that a completely satisfactory extra-nuclear theory will not be possible either until radiation is satisfactorily incorporated. It is pointed out in the attached Notes* that irreversible changes appear excluded in the standard wave-vector technique (this is surprising in view of the common claim that the processes covered are non-deterministic). There is presumably the possibility, however, as apparently envisaged in your treatment of the electromagnetic field, of introducing irreversible changes in the well-known statistical way from reversible ones by averaging over a large number of irrelevant degrees of freedom after the complete equations have been set up.

20-4-45

Dear Moyal,

Thanks for sending me your manuscript again. The situation with regard to joint probability distributions is as follows, as I understand it.

A joint distribution function $F(p,q)$ should enable one to calculate the mean value of any function $f(p,q)$ in accordance with the formula

$$\text{mean}(f(p,q)) = \iint f(p,q)F(p,q)dpdq \quad (1)$$

I think it is obvious that there cannot be any distribution function $F(p,q)$ which would give correctly the mean value of any $f(p,q)$, since formula (1) would always give the same mean value for pq and for qp and we want their means to differ by $i\hbar$. However one can set up a d.f. $F(p,q)$ which gives the correct means values for a certain class of functions $f(p,q)$. The d.f. that you propose gives the correct mean

value for $e^{i(\tau A + \theta B)}$, for τ and θ any numbers, but would not give the correct mean value for other quantities, e.g. it would give the same mean value for $e^{i\tau A} e^{i\theta B}$, whereas we want this second quantity to be $e^{i\theta/2A}$ times the first. In some work of my own I was led to consider a d.f. which gives correctly the mean value of any quantities of the form $\sum_j f_j(p) g_j(q)$, i.e. all the p 's to the left of all the q 's in every product. My d.f. is not a real number in general, so it is worse than yours, which is real but not always positive, but mine is connected with a general theory of functions of non-commuting observables.

I am writing up my work for publication and I propose to refer to your work somewhat in these terms:-

'The possibility of setting up a probability for non-commuting observables in quantum mechanics to have specified values has been previously considered by J.E. Moyal, who obtained a probability for a coordinate q and a momentum p at any time to have specified values, which probability gives correctly the averages of any quantity of the form $e^{i(\tau A + \theta B)}$, where τ and θ are real numbers. Moyal's probability is always real, though not always positive, and in this respect is more physical than the probability of the present paper, but its region of applicability is rather restricted and it does not seem to be connected with a general theory of functions like the present one.'

Do you think this reference would correctly describe your work and do you have any objection to such a reference?

There may be other d.f.'s which are worth considering and there is a field of research open here. Will you be able to work on it?

Yours sincerely,
P.A.M. Dirac

18 Ambrose Avenue
London N.W. 11.
April 29th, 1945.^[3]

Dear Professor Dirac,

Many thanks for your letter. I was most interested by your remarks concerning your work on a general theory of functions of non-commuting observables, and should be very glad to see it. Are you acquainted with the work of Whittaker, and Kermack and McCrea on this subject? The references are: E.T. Whittaker, *Proc. Ed. Math. Soc. Ser. 2*, v. 2 (1931) 189–204; W.O. Kermack and W.H. McCrea, *ibid. ser. 2*, v. 2. (1931), 205–219 and 220–239.

If I understand correctly your remarks concerning joint probability distributions, you consider them as functions of the non-commuting variables P, Q , which will give correct averages for certain classes of functions of the latter. (I shall use hereafter P, Q for the non-commuting quantities, and p, q , for the corresponding commuting variables.) Such functions may of course prove extremely useful mathematically, but they can hardly be called probability distributions in any ordinary sense.

My approach to this problem has been entirely different. I have looked for a probability distribution in the ordinary sense, which will be a function of the ordinary, commuting variables p, q . Its connection with functions of the corresponding non-commuting operators P, Q of quantum mechanics, is that it should give correct means for such of these functions (i.e. Hermitian operators) as are formed to represent physical quantities. If a physical quantity is given in classical mechanics by a function

$M(p,q)$, (i.e. a Hamiltonian, or an angular momentum) a Hermitian operator $M(P,Q)$ is formed to represent it according to certain rules. I have looked for an $F(p,q)$ such that it will always give

$$(1) \bar{M} = \int \psi^*(q) M(P,Q) \cdot \psi(q) dq = \iint M(p,q) F(p,q) dp dq$$

It is obvious that such a function $F(p,q)$ should be connected with a unique method of forming the quantum mechanics operators from the corresponding classical mechanics functions if p and q (I am speaking of course, of the classical quantum mechanics for particles without spin). A first test for the correctness of such an $F(p,q)$, will therefore be that the corresponding method for forming operators should give correctly at least all the known Hermitian operators of the theory, (since a general method for forming these operators is not generally agreed upon in the standard theory).

The $F(p,q)$ which I propose in my paper fulfils these conditions. It can be expressed either as a series development in $\psi(p)$ and $\varphi(p)$ or as an integral expression in terms of the ψ 's alone or φ 's alone (the latter is due to M.S. Bartlett) as follows

$$(2) \quad F(p,q) = e^{\frac{\hbar}{2i} \frac{\partial^2}{\partial p \partial q}} \left\{ p^{-1} \psi^*(q) \varphi(p) e^{i p q / \hbar} \right\} \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i \tau p} \psi^*(q - \frac{\hbar \tau}{2}) \psi(q + \frac{\hbar \tau}{2}) d\tau.$$

I have shown that it corresponds univocally to the following method of forming operators (already proposed by McCrea). Let $M(p,q)$ be an ordinary function of p and q (e.g. some constant of the motion in classical mechanics). To form the corresponding operator $M(P,Q)$ we write first a function $M_p(P,Q)$ of the non-commuting operators P, Q , which is obtained from $M(p,q)$ by placing all the P 's to the right of the Q 's, i.e. by replacing all polynomial terms $q^m p^k$ in $M(p,q)$ by $Q^m P^k$. The correct operator $M(P,Q)$ is then obtained as

$$(3) M(P,Q) = e^{\frac{\hbar}{2i} \frac{\partial^2}{\partial P \partial Q}} M_p(P,Q)$$

Form (2) for $F(p,q)$ will give correct averages for all operators formed as in (3) by averaging in p - q space over the corresponding ordinary function $M(p,q)$, i.e.

$$\bar{M} = \int \psi^*(q) M(P,Q) \cdot \psi(q) dq = \iint M(p,q) F(p,q) dp dq$$

It is consequently incorrect in my view to say that the $F(p,q)$ in my paper will give correct averages only for functions of the form $\exp i(\alpha P + \beta Q)$. Actually, it will give the right averages for all operators formed as in (3), and in particular, for all the Hermitian operators considered in the classical quantum mechanics of particles without spin, e.g. Hamiltonian, angular momentum, total angular momentum, radial momentum, etc. It is easy to check that (3) does give the usual operator form for all these quantities. In the case of the example quoted in your letter, it will give correct average for $QP + \frac{\hbar}{2} = \frac{1}{2}(QP + PQ)$. (I may mention here that this form of $F(p,q)$ and method of forming operators is valid for rectilinear coordinates only.)

Furthermore, the $F(p,q)$ in my paper leads to certain forms for the space-conditional averages of the powers of p (i.e., averages of p^m for a given value of q), the first two being

$$(4) \rho(q) = \int F(p,q) dp = \psi \psi^*$$

$$(5) \quad \rho \bar{p} = \int p F(p, q) dp = \frac{\hbar}{2i} \left(\psi^* \frac{\partial \psi}{\partial q} - \psi \frac{\partial \psi^*}{\partial q} \right)$$

$$(6) \quad \rho \bar{p^2} = \int p^2 F(p, q) dp = \left(\frac{\hbar}{2i} \right)^2 \left(\psi^* \frac{\partial^2 \psi}{\partial q^2} - 2 \frac{\partial \psi^*}{\partial q} \frac{\partial \psi}{\partial q} + \psi \frac{\partial^2 \psi^*}{\partial q^2} \right)$$

Early in my work (Sect. II) I obtained a set of partial differential equations for probability distributions, which have the form of the hydrodynamic equations of continuity and motion and express conservation of probability. These are of quite general validity, and are not connected with any special form of $F(p, q)$ or any physical assumption. Substitution in these general equations of the expressions above for the space-conditional means of p , p^2 , taken in conjunction with the equations of classical mechanics, lead to the Schrödinger equation, as I have shown in my paper. The Schrödinger equation is thus shown to result from this special form for $F(p, q)$, the laws of classical mechanics, and the general properties of probability distributions for dynamical variables. I think this is the other essential condition for a correct $F(p, q)$: that it should be consistent both with the Schrödinger equation and the equations for conservation of probability.

Regarding the range of validity of form (2) for $F(p, q)$, and the fact that it leads to negative values for single eigenstates, I have already mentioned in my last letter that this may possibly mean, reverting to your view, that joint measurement of p and q is inconceivable in pure states, but only in a combination of states that leads to a defined positive $F(p, q)$. I think possibly this may be a general feature for any possible $F(p, q)$ in quantum mechanics, because of the necessary orthogonality properties of the phase-space eigenfunctions corresponding to pure states. Such (possibly) negative eigenfunctions, which must be compounded to give a positive probability function, occur in the classical calculus of probabilities in the theory of chain probabilities. However, as was pointed out by M.S. Bartlett, even the possibly negative $f(p, q)$ corresponding to a pure state will still lead to correct averages for operators of form (3), so that the theory retains its usefulness even in this connection. I pointed out in my last letter for example, how it could be used to calculate transition probabilities.

In conclusion, my view is that this form (2) of $F(p, q)$ has quite general validity, and that the theory it leads to, is entirely equivalent to the classical quantum mechanics of particles without spin.

I have considered the connection of this theory with the general theory of functions of non-commuting variables. From this point of view, the theory starts with $\exp i(\mathbf{p}P + \mathbf{q}Q)$, and leads to the general method (3) for forming observables. One might conceivably take another starting point, which would be connected with some other method for forming observables. However, apart from other considerations (cf. Hermann Weyl, "Theory of Groups and Quantum Mechanics" p. 275) all the other forms of $F(p, q)$ I tried, taken in conjunction with classical mechanics and the equations of conservation of probability, did not lead to the Schrödinger equation, but to some different wave equation. They correspond thus to some scheme different from the classical quantum mechanics. In particular I discarded for this reason the first $F(p, q)$ I tried, which was connected with the general operator form

$$(7) \quad M(P, Q) = \sum_{n,m} c_{nm} (P^n Q^m + Q^m P^n)$$

which gave the exponential form $\frac{1}{2} [e^{iQ} e^{iP} + e^{iP} e^{iQ}]$ and consequently had the form

$$(8) \quad F(p, q) = \frac{1}{2\pi} \int e^{-i\tau} \left[\psi^*(q) \psi(q + \hbar\tau) + \psi(q) \psi^*(q - \hbar\tau) \right] d\tau \\ = \frac{1}{2} \hbar^{-1} \left[\psi^*(q) \varphi(p) e^{ipq/\hbar} + \psi(q) \varphi^*(p) e^{-ipq/\hbar} \right]$$

I believe I showed you these attempts in 1940.

One of the problems in the theory of non-commuting variables, which I have not been able to solve is: what general transformation will leave form (2) for $F(p,q)$ and (3) for operators invariant? It is easy to see that this is the case for linear transformations from Cartesian coordinates, and also for the dynamical-contact transformation of classical mechanics; but it is not maintained e.g. for a transformation to polar coordinates and their conjugate momenta. An allied problem is to find a general form for $F(p,q)$ for any canonical coordinates corresponding to form (2) for rectilinear coordinates. I am hoping your work will give me a lead in this connection.

With regards to your query, I do not, for the reasons mentioned above, think that your reference to my work gives a correct description of it. It is certainly not correct in my view to say that form (2) for $F(p,q)$ is limited to giving correctly averages for quantities of the form $e^{i(E^*+RQ)}$; in fact, it will give averages for all observables formed as in (3), and this includes as far as I know, all the observables ordinarily considered in classical quantum theory. This would not perhaps matter a great deal, if my work was already published, since readers could then refer to the original. I have not however been able so far to arrange for its publication, due largely, as you will no doubt remember, to your veto, which made the late Professor Fowler hesitate about presenting it to the Royal Society. Your criticism is thus left without an answer. Your objection at the time, if I remember rightly, was chiefly that joint distributions for p and q had no physical meaning and consequently no validity or usefulness. I am glad to notice that you now think they open an interesting field of research.

Regarding your query to whether I shall be able to do further work on this subject, my main difficulty is again the fact that my existing work is not yet published. For one thing, I shall want to base future work, at least partly, on the papers now in your hands. It is also discouraging to accumulate for years unpublished results, as I have been doing. Finally, there are material difficulties: the papers you have seen, represent my first real effort at research in pure mathematics and theoretical physics; I was hoping that their publication would eventually enable me to transfer my activities entirely from the field of research in engineering and applied physics to that of pure science, and to do some serious work on theoretical physics. Failure to obtain publication has forced me to adjourn such plans sine die, and my present work is leaving me less and less time for pure research.

Yours sincerely
[J.E. Moyal]

c/o. Goscote Hotel,
Goscote Hall Road,
BIRSTALL.
Leics.
April 25th, 1945.

Dear Professor Dirac,

There are a few points in the paper I sent you which I should like to amplify.

First, regarding the range of validity of the $F(p,q)$ distributions, I have been considering the possibility of a modified interpretation of the mathematical formalism. You will have noticed that one of the difficulties of the theory is that the method of forming $F(p,q)$ does not lead to functions that are defined positive for all p and q when applied to a system in a single eigenstate. This might be interpreted,

reverting partly to the point of view expressed in your book, as indicating that simultaneous probability distributions for p and q have no precise meaning for a system in a single eigenstate, or again, that a classical particle picture is not valid for a system in a pure state, and that the hypothesis of pure state is incompatible with the simultaneous measurement of p and q . The classical particle amongst a number of states in such a manner is to make $F(p,q)$ positive.

This would limit the possibility of giving the probabilities of simultaneous values for p and q . However, as M.S. Bartlett points out in his paper, it does not necessarily upset the mathematical structure of the theory or its equivalence to classical wave mechanics. If, as I think, this equivalence is correct, then the theory should lead to correct results for the various quantities obtained by wave mechanics, such as frequencies and transition probabilities, even when dealing with negative functions $F(p,q)$. The appearance of the latter should then be taken to mean that the situation is such that simultaneous prediction of the values of p and q is impossible, but would not impair the calculation of other experimentally determinable quantities.

It would be possible to use the formalism of this theory to supplement in certain cases, the perturbation method in the calculation of transition coefficients. This can be done as follows: if the system is originally in the unperturbed eigenstate k , with the phase space eigenfunction $f_{kk}(p_0, q_0)$ corresponding to the q -space eigenfunction $u_k(p_0, q_0)$

$$(1) f_{kk}(p_0, q_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\tau p_0} u_k^*(q_0 - \frac{h\tau}{2}) u_k(q_0 + \frac{h\tau}{2}) d\tau$$

the phase space distribution $F(p,q,t)$ at time t would be obtained by substituting in $f_{kk}(p_0, q_0)$ the classical solution $p(t), q(t)$, in terms of the initial values p_0, q_0 , for the system under the action of the perturbing forces (when it is possible to find such solutions). In other words, one would apply to $f_{kk}(p_0, q_0)$ the contact transformation in time of classical mechanics to obtain $F(p,q,t)$ at time t ; one could then expand the latter in terms of unperturbed phase space eigenfunction $f_{in}(p,q)$

$$(2) F(p,q,t) = \sum_{in} a_{in}^* a_{in} f_{in}(p,q)$$

and obtain thus directly the transition coefficients $a_{kn}^*(t) a_{kn}(t) = A_{kn}$ from state k to state n .

Applied, for example, to the schematic case of an oscillator of charge e , following the application of a perturbing electric force of large wavelength, this method leads for the transition coefficient from the ground state to the k -th state to the exact expression

$$(3) A_{0k} = (\frac{1}{2} \Delta E t)^k e^{-\frac{1}{2} \Delta E t} / k!$$

(calling ΔE the increase in mean energy). The first term of the expansion of A_{0k} in power of ΔE coincides then with the first approximation by the perturbation method

$$(4) \begin{matrix} A_{0k} = \frac{1}{2} \Delta E t & \text{for } k=1 \\ = 0 & \text{" } k > 1 \end{matrix}$$

I have been considering the application of this method to radiation oscillators, in view of the possibility that some of the divergences may be due to a mathematical breakdown of the perturbation method.

Best regards,
[J.E. Moyal]

P.S. I have just received your letter but must defer answering for a few days, as I am moving to London. My new address will be 18. Ambrose Ave. N.W. 11.

11-5-45

Dear Moyal,

Thanks for your letter and your references to Whittaker and others. These papers are very interesting, though not directly connected with the subject under discussion.

I still do not agree that your d.f. gives correctly the average values of all Hermitian operators considered in classical mechanics. It is true that it works alright for $\frac{1}{2}(qp + pq)$, but it goes wrong as soon as one applies it to more complicated examples. For example your d.f. would give the same average for the two Hermitian operators QP^2Q and PQ^2P , whereas they ought to differ by $2\hbar^2$. You may answer that these two Hermitian operators do not correspond to classical quantities. To anticipate this answer I have worked out another example, which certainly is of practical importance. Take a harmonic oscillator of energy $\frac{1}{2}(Q^2 + P^2)$. Its average energy when it is at a temperature T is the average value of the Hermitian operator $e^{-\frac{1}{2}(Q^2 + P^2)/\hbar T}$. I have checked that this average value is not given correctly by your d.f. Your d.f. gives the correct average for quantities of the form $e^{i(aP + bQ)}$ and for quantities expressible linearly in terms of such quantities, e.g. $\iint f(a, b) e^{i(aP + bQ)} da db$ for any $f(a, b)$ or $\frac{1}{2\pi\hbar} e^{i(aP + bQ)}$, but is not more general than this. Do you not agree?

I have enclosed a copy of my paper. I should be glad if you would send it back in two or three weeks time, as I do not have another copy.

Do you want me to send you back your work now? I would be willing to help you publish it if you would change it so that it does not contain any general statements which I think to be wrong. I would suggest it would be better to publish the quantum theory part separately from the rest, because it is on rather a different footing (according to my view).

Yours sincerely,
P.A.M. Dirac

18-5-45

Dear Moyal,

Your theory gives correctly the average energy when the system is in a given state, (i.e. represented by a given wave function) but not when the system is at a given temperature. Take a harmonic oscillator with energy $\frac{1}{2}(q^2 + p^2)$. The probability of its being in the n -th state is proportional to the average value A_n of $e^{-\frac{1}{2}(q^2 + p^2)/\hbar T}$. According to your theory

$$A_n = \iint e^{-\frac{1}{2}(q^2 + p^2)/\hbar T} F_n(p, q) dp dq = \iint e^{-\frac{1}{2}(q^2 + p^2)/\hbar T} e^{-\frac{1}{2\pi\hbar} i} f_n(p, q) dp dq$$

with

$$f_n(p, q) = \hbar^{-1} \psi_n^*(q) \varphi_n(p) e^{i\pi p q / \hbar}$$

When the A_n 's have been calculated, we can get the average energy by

$$\bar{E} = \frac{\sum_n E_n A_n}{\sum_n A_n} = - \frac{\frac{d}{d(\frac{E}{h})} \sum_n A_n}{\sum_n A_n}$$

It is not very easy to calculate A_n , but is quite easy to calculate $\sum_n A_n$ from the known property of wave functions

$$\sum_n \psi_n^*(q) \phi_n(p) = h^{-1} e^{-iqp/h}$$

Thus

$$\sum_n f_n(p, q) = h^{-1}$$

and

$$\sum_n A_n = h^{-1} \iint e^{-\frac{h}{2}(q^2 + p^2)/hT} dp dq = \frac{\pi}{2} \frac{kT}{h}$$

We now get

$$\bar{E} = kT$$

which is the classical result and not the quantum one.

In Bartlett's paper which you just sent me, the quantum values for the energy of the harmonic oscillator are assumed and the correct value for \bar{E} was obtained because of this assumption. You can always get the right answer by borrowing sufficient results from the ordinary quantum theory. The true test of a theory is whether it always gives consistent results whichever way it is applied, and my way of evaluating given above shows that your theory does not always give consistent results. The discrepancy in this case arises because I use your d.f. for calculating the average of $e^{-\frac{h}{2}(p^2 + q^2)/hT}$, and this quantity is not expressible linearly in terms of $e^{i(qp + pq)}$.

You say your theory gives a different value for $\overline{E^2} - (\bar{E})^2$, and this can only mean that your theory is not consistent with the usual quantum values for the energy, otherwise there is no room for any uncertainty in the value of $\overline{E^2} - (\bar{E})^2$. Your theory gives a value for $\overline{E^2} - (\bar{E})^2$ greater than the usual one by an amount $\frac{1}{4}h^2$ (with $E = \frac{1}{2}(p^2 + q^2) - \frac{1}{2}h$). Thus for a harmonic oscillator in its state of lowest energy your theory will give fluctuations in energy corresponding to $\overline{E^2} - (\bar{E})^2 = \frac{1}{4}h^2$, instead of a constant energy. Surely you must agree that your theory is wrong in this case, and that therefore it has limitations.

The general statement in your work that I disagree with is the one (given in your last letter) that dynamical variables must be of the form $\iint \rho(\tau\sigma) e^{i(p\tau + q\sigma)} d\tau d\sigma$. The square of the energy of a harmonic oscillator, namely $\left[\frac{1}{2}(p^2 + q^2) - \frac{1}{2}h \right]^2$ is not of this form, and if you replace it by something that is of this form you get energy fluctuations in the state of lowest energy, which I think is a self-contradiction.

Yours sincerely,
P.A.M. Dirac

18 Ambrose Avenue
London N.W.11.
May 15th, 1945.

Dear Professor Dirac,

Many thanks for your letter and enclosed paper. I have not yet had time to read the latter, but I shall do so as soon as possible.

I am not quite clear as to how you worked out the average energy for an oscillator at temperature T . The theory in my paper gives correctly the average energy for a Maxwell-Boltzmann assembly of N oscillators. I enclose the draft of an unfinished paper by M.S. Bartlett and myself which gives the relevant calculations in §4 (you may also find §2 & §3 of some interest). A difference with the orthodox method is found not in the expression for the average energy \bar{E} , but in the standard deviation, which comes out as $\sigma^2(E) = \overline{E^2} - (\bar{E})^2 = (\overline{E^2})/N$ instead of $(\bar{E})^2/N - N(\frac{h\nu}{T})^2$ (not neglecting the ground state energy). I have always found so far that my treatment leads to the same average values as the usual methods, but shows difference in the fluctuations: this may lead to an experimental test of the theory.

I agree that my d.f. yields correct averages for quantities expressible linearly in terms of expressions $\exp i(\alpha P + \beta Q)$ such as

$$(1) G(P, Q) = \int \int_{-\infty}^{\infty} g(\tau, \theta) e^{i(\alpha P + \beta Q)} d\tau d\theta$$

but this includes quite a wide class of functions. In fact, it can be shown (c.f. McCoy, *Proc. Mat. Acad. Sc.*, 18 (1932) 634) that (1) is equivalent to the form for Hermitian operators mentioned in my last letter.

$$(2) G(P, Q) = e^{\frac{1}{2} \frac{\partial^2}{\partial P^2}} \cdot G_p(P, Q)$$

For a polynomial term $p^2 q^2$ the corresponding operator $(P^2 Q^2)_0$ obtained by (1) or (2) can be cast in a more symmetrical form

$$(3) (P^2 Q^2)_0 = \left(\frac{1}{2}\right)^2 \sum_{\ell=0}^2 \binom{2}{\ell} Q^{2-\ell} P^2 Q^{\ell}$$

In particular, for the term $(P^2 Q^2)_0$ mentioned in your letter, (2) and (3) lead to

$$(4) \begin{aligned} (P^2 Q^2)_0 &= Q^2 P^2 + \left(\frac{1}{2}\right) \cdot 4QP + \frac{1}{2} \left(\frac{1}{2}\right)^2 \cdot 4 \\ &= \frac{1}{4} \{ Q^2 P^2 + 2QP^2 Q + P^2 Q^2 \} = QP^2 Q - \frac{h^2}{T^2} \end{aligned}$$

(by the way, surely $QP^2 Q - PQ^2 P = 0$!).

The hypothesis on which I base my derivation of the d.f. (and therefore the rest of the theory) is equivalent to the assumption in the standard (matrix) theory that dynamical observables must be of the form (1) (non-dynamical operators might be construed in the statistical theory as symmetry, etc. conditions on the d.f.). Relation (1) is obviously more restrictive than Heisenberg's exchange relations alone: it might be considered as the basic postulate of a well-defined form of quantum kinematics. In this form, it has been given by H. Weyl, who bases his arguments in its favour on group-theoretical considerations: iP , iQ generate a unitary Abelian group in 'ray'-space; the hypothesis is then that dynamical observables are the matrices of the representation of this group's algebra, which are given by (1) if the group is supposed irreducible. My argument is, that it leads to a theory that is consistent both with the Schrödinger equation and the usual statistical interpretation. I think it should be possible to prove that it is the only form of quantum kinematics that does so, and that a different form would

necessitate revising either the statistical interpretation, or the wave-equation — but this is only a conjecture so far.

Summarizing, I think it would be fair to say that my paper gives a derivation of classical quantum mechanics on a purely statistical basis, (plus Newtonian mechanics) which is equivalent to the standard matrix theory with the addition of Weyl's postulate for a quantum kinetics and furthermore that it shows the consequences such a theory entails with regards to the problems of determinism, probability distributions, fluctuations, quantum statistics, etc. Would you agree to this character; and the controversial issue it raises? I am not clear, however, as to exactly what general statements you think are wrong.

I shall not need my typescript until there is a need of revising it for publication, so that you can return it whenever you have finished with the problems of determinism, fluctuations, quantum statistics, etc. Would you agree to this statement of the position?

I thank you for your (conditional) offer to help me publish my papers. I have no objection to publishing the quantum theory part separately; I agree, it is on a different footing from the rest, because of its more tentative character; and the controversial issue it raises. I am not clear, however, as to exactly what general statements you think are wrong.

I shall not need my typescript until there is a need of revising it for publication, so that you can return it whenever you have finished with it.

[J.E. Moyal]

18 Ambrose Avenue
London N.W.11.
May 26th, 1945.

Dear Professor Dirac,

I thank you for your letter of the 18th. With regards to your derivation of the average energy for an oscillator at fixed temperature, I don't know how this method works out in the standard theory, but the reason for the result you obtained on the basis of my theory is fairly obvious. You start with a Maxwell d.f. for p and q

$$(1) F(p, q) = e^{-\frac{1}{2}(p^2 + q^2)/kT}$$

You then work out the coefficient

$$(2) A_n = \iint e^{-\frac{1}{2}(p^2 + q^2)/kT} f_{nm}(p, q) dp dq$$

Since the $f_{nm}(p, q)$

$$(3) f_{nm}(p, q) = e^{-\frac{1}{2}\frac{p^2}{h^2}} \cdot h^{-\frac{1}{2}} \psi_n^*(q) \phi_n(p) e^{ipq/h}$$

form an orthogonal set in phase-space, the coefficient A_n is merely the Fourier coefficient a_{nn} in the expansion

$$(4) e^{-\frac{1}{2}(p^2+q^2)/kT} = \sum a_{nm} f_{nm}(p, q)$$

(It is possible to show that in (4) $a_{nn}=0$ for $n \neq m$). You then proceed to show through the A_n that for (1)

$$(5) \bar{E} = \frac{1}{2}(\overline{p^2} + \overline{q^2}) = kT$$

but this is of course obvious by a direct calculation

$$(6) \bar{E} = \frac{\iint \frac{1}{2}(p^2 + q^2) e^{-\frac{1}{2}(p^2+q^2)/kT} dp dq}{\iint e^{-\frac{1}{2}(p^2+q^2)/kT} dp dq}$$

The correct method for evaluating \bar{E} for an assembly of oscillators in my theory is the one given in my joint paper with Bartlett, and it leads to the usual result

$$(7) \bar{E} = \frac{hv}{e^{hv/kT} - 1} + \frac{hv}{2}$$

I don't think your remark on getting the right answer 'by borrowing sufficient results from the ordinary quantum theory' quite fair: in so far as my theory is equivalent to the ordinary theory, it leads to the same eigenvalues for the mean of the energy, as I have shown in my paper. In order to prove an inherent inconsistency in my theory one would have to show that the method you use follows necessarily from my basic postulates, but this is not the case. My method on the other hand is based on a theory for statistical assemblies resulting from these postulates (c.f. my paper, §10). As such, it is quite consistent with the rest of the theory, and also appears to lead to correct results.

The difficulty regarding the dispersion $\frac{1}{2}h^2$ for the energy of the oscillator in a single eigenstate is more serious. I think it is connected with the fact that $f(p, q)$ can be negative: if the conclusion is (in accordance with your views) that a joint d.f. for p and q is impossible in a single eigenstate, then the probability distribution for $\frac{1}{2}(p^2+q^2)$, and consequently the $\frac{1}{2}h^2$ dispersion, have no direct physical meaning. This could be interpreted through the fact that it is impossible to measure the energy in a single eigenstate in a finite time. Only a d.f. giving the band-width and intensity distribution of the spectrum lines would have a physical meaning, and could be compared with experiment. This would involve, however, extending the theory to include radiation.

I am prepared to mention your objections concerning the operator forms $\iint \rho(\tau, \sigma) e^{i(\sigma^2 + \omega\tau)} d\tau d\sigma$ in the body of my paper (do you agree that with the imposition of this restriction on operators for dynamical variables in the usual matrix theory, the latter becomes equivalent with my theory?)

I do not think there are any inherent inconsistencies in my theory, but I agree that this restriction leads to results that do not tally with certain hitherto accepted features of the usual theory, and may possibly clash with experimental results. Should the latter prove to be the case, then in my view the conclusion to be drawn from my work would be, that the usual statistical interpretation of classical quantum mechanics must be revised. Comparison with the experiment of such differences with the usual theory might perhaps be sought for in the fluctuations for statistical assemblies, the intensity distributions of spectral lines, or the calculation of transition probabilities.

If you agree to the above, then I should be glad to know if you are still prepared to help me in publishing my work and what form of publication you would suggest. I think I could condense the mathematical part into a paper in two parts of 15–20 pages each, and the quantum mechanics part into 20–25 pages.

I return your typescript, which I read with great interest, especially as I have treated the same subjects in my paper and arrived at different conclusions. For example, the operator form I use constitutes a general method for forming functions of observables which (as compared with yours) is unambiguous when the latter are non-commuting, and does not depend on their order. We have already discussed the d.f. for p and q at one instant of time, but I have also given an expression for their distribution at two instants of time, in terms of the phase-space eigenfunctions in my main paper (§14), and in terms of the transformation function $\langle q_1 | q_2 \rangle$ in §2 of the paper on the oscillator I sent you, which it is interesting to compare with your results on the same subject. My conclusion regarding trajectories in my theory is that for a conservative and unperturbed system they reduce to those of classical mechanics, I discussed the resulting implications with regards to the principle of uncertainty and the problem of determinism in §15, and showed in the succeeding paragraphs, that it leads to correct results in examples on the free and uniformly accelerated particle, and the oscillator. I have also worked out in collaboration with Bartlett an alternative method of calculating $\langle q_1 | q_2 \rangle$ from Hamilton's principal function in classical mechanics based on Whittaker's work.

[J.E. Moyal]

6-6-45

Dear Moyal,

I expect to be going abroad in a few days time and not to be back till the end of July, so I am returning your papers herewith in case you should need them in the meantime. Thanks for returning my paper.

It now appears that the dispersion of the energy in a stationary state is the simplest example which shows the limitations of your theory. This dispersion will be pretty general on your theory, and will probably occur with all stationary states and all dynamical systems. This is not a difficulty that can be got around in any way, because it contradicts the whole idea of sharp energy levels — it would imply a lack of sharpness in the energy levels much too great to be reconciled with experimental evidence. It shows therefore that the joint d.f. does not work in the case of E^2 . Also it does not work for higher powers of E .

If the limitations in the applicability of the joint d.f. are clear[ly] stated, which would mean partly rewriting it, I would be glad to help you publish your work. The quantum theory part of your work could form a paper which I could communicate to a scientific journal. With regard to the remainder, I do not know how much of it represents new research work and how much is an exposition of known results. Do you have any suggestion about where it should be published? What did Fowler say about it?

Yours sincerely,
P.A.M. Dirac

18 Ambrose Avenue,
LONDON N.W. 11,
17th June, 1945

Dear Professor Dirac,

I was sorry to see in the press that your visit to the U.S.S.R. was cancelled at the last moment: I

expect you must be very annoyed at the whole incident.

Your letter and my papers reached me only on Tuesday: the delay was apparently due to the fact that the envelope had broken open during transit; fortunately nothing seems to be missing.

I agree that the occurrence of non-zero dispersions in eigenstates is the main difficulty or limitation in my theory. I did point it out and discuss it at some length in the paper I sent you, and will of course do so again as clearly as I can when I redraft it for publication (which I intend to do in any case in order to produce a condensed version.)

My work on Random Functions is new. Professor Fowler's original suggestion was to present the whole work for publication in the Proc. Roy. Soc. (including the part on Quantum Mechanics) as three separate papers. My intention was then to rewrite it in a more condensed form, cutting out appendices, some of the examples, etc., so as to have three papers of 15 to 20 pages each. Would you consider this now as a suitable arrangement?

Bartlett has told me that you are holding colloquiums on Quantum Mechanics in Cambridge. Would it be possible for me to attend some of these? I shall be visiting Cambridge fairly regularly in connexion with my present duties, and it may prove possible to arrange for these visits to coincide with the date of your colloquium.

[J.E. Moyal]

7 Cavendish Avenue,
CAMBRIDGE
26.6.45

Dear Moyal,

The quantum theory part of your work could be written up as one paper, and the remainder as two more, provided it divides naturally into two parts. If it does not divide it might be better to keep it as one long paper. Probably the Proc. Roy. Soc. is the best journal for them.

We have been having Colloquiums, usually on Friday afternoons but sometimes on Monday afternoons. They will probably be resumed in October and we would be glad if you could come to any of them.

Yours sincerely,
P.A.M. Dirac

18 Ambrose Avenue,
London N.W. 11
10th July, 1945.

Dear Professor Dirac,

Many thanks for your letter of the 26th. As you suggest, I am now rewriting the part of my work on quantum mechanics as a separate paper. As regards the rest, I am rewriting it as a paper in two parts, which could then appear either separately or together, whichever is more convenient.

Thank you for your invitation to the colloquiums; I am looking forward to attending them.

I enclose some notes in which I have tried to develop a method which would overcome the difficulty about non-zero-dispersions for eigenvalues in my theory and also extend it to generalized canonical coordinates. This is still tentative in character, and there are several things I still want to clear up, but I should be glad in the meantime to have your opinion on this development. I also enclose some notes comparing the results in your paper with mine.*

Yours sincerely
[J.E. Moyal]

18 Ambrose Avenue,
London N.W. 11,
21st August [1945]

Dear Professor Dirac,

You may be interested in a paper by Wigner, Phys. Rev. 40 (1932), 749, which anticipates my derivation of the p-q distribution. I believe Bartlett has told you about this.

I understand from Bartlett, that you are leaving for the U.S. on the 30th. Would it be possible for me to send you the m.s. of my papers to you there, if and when I complete them?

With my best wishes for a pleasant journey.

Yours sincerely,
[J.E. Moyal]

17 Cavendish Avenue,
CAMBRIDGE,
31-10-45

Dear Moyal,

Your new version is more in accordance with the standard quantum mechanics, but it is considerably more complicated as you need a different joint prob. distr. for each system of coordinates. You are definitely departing from classical statistics when you make the joint prob. distr. depend on the system of coordinates, and if you depart so much from the usual classical ideas is there any point in trying to fit things into a classical framework? What advantages does your system have over the usual statistical interpretation of quantum mechanics? Any results that you get from your system must either conform to the usual quantum mechanics or else be incorrect. I think your kind of work would be valuable only if you can put it in a very neat form.

I am returning your paper herewith,

Yours sincerely,
P.A.M. Dirac

05,060 et 05,040⁰⁰. Une poussée leucocytaire avec polynucléose s'est développée et fut d'autant plus marquée que l'évolution fut plus rapide.

Certaines modifications de la moelle osseuse méritent d'être soulignées. Inconstantes, nous les avons observées dans les tout derniers jours de l'évolution de l'insuffisance parathyroïdienne, d'une part chez un chien mort de tétanie le 11^e jour, d'autre part, surtout, chez deux chiens dont la survie prolongée par l'absorption journalière de 2^e de lactate de calcium, a atteint 23, 28 jours. Chez ces animaux, atteints de cet état de cachexie progressive, décrite par E. Gley, s'est développée tardivement dans la moelle, une réaction réticulo-endothéliale intense, faite de noyaux nus, du type histiocytaire, baignant dans les mailles du collagène. Cette réaction tend à étouffer le tissu myéloïde normal qui persiste cependant. On constate 3 à 4 % de cellules indifférenciées.

Cet aspect n'est donc pas celui de la leucémie aiguë. Il s'oppose, jusqu'à un certain point, à l'aspect fibreux que prend la moelle osseuse dans l'hyperparathyroïdie. Mais si l'on se rappelle que le système réticulo-endothélial de la moelle est le plasmode dont les éléments donneront naissance, en se libérant, aux lignées myéloïde et peut être monocytaire du sang, il n'est pas illogique de penser que, dans certains cas, sous l'influence d'une insuffisance parathyroïdienne, la prolifération de ce tissu puisse aller jusqu'à la libération de formes jeunes indifférenciées, leucoblastiques. C'est là, pour l'instant du moins, une hypothèse de travail qui peut éclairer notre observation.

CORRESPONDANCE.

M. **DARIO ACEVEDO** adresse des remerciements pour la distinction accordée à ses travaux en 1940.

M. le **SECRETARE PERPETUEL** signale parmi les pièces imprimées de la Correspondance :

Bulletin de l'École polytechnique de Jassy. Tome 1, fascicule I, janvier à juin 1946.

MECANIQUE ONDULATOIRE. — *Sur les rapports entre la théorie des mélanges et la statistique classique*. Note (1) de M. **JACQUES YVON**.

Se limitant à des problèmes sans spin, Wigner (2) a montré la possibilité de calculer, en mécanique quantique, notamment à propos de questions de thermodynamique, les moyennes à l'aide d'un formalisme identique à celui de la statistique classique. Wigner remarque toutefois que la densité de proba-

(1) Séance du 1^{er} juillet 1946.

(2) *Phys. Rev.*, 40, 1932, p. 749.

bilité qu'il utilise, fonction des coordonnées et des impulsions, ne mérite pas ce nom à proprement parler parce qu'elle n'est pas nécessairement positive. Il lui semble de plus que cette densité n'est pas déterminée d'une manière univoque et qu'elle ne permet à coup sûr le calcul correct des moyennes que dans le cas assez limité où l'opérateur quantique correspondant dépend seulement soit des coordonnées, soit des impulsions, et jamais des deux à la fois.

Récemment, J. Bass ⁽³⁾, puis E. Arnous ⁽⁴⁾, ont repris cette question et, somme toute, tombent d'accord avec Wigner sur l'ensemble de ces conclusions.

Je me propose de montrer ici que les réserves faites sur l'étendue du domaine d'application de la densité de Wigner sont entièrement injustifiées, et d'indiquer de manière précise comment les calculs doivent être conduits. Je me servirai à cet effet des résultats acquis dans une Note antérieure ⁽⁵⁾ dont je conserverai les notations.

J'insiste sur le fait que pour donner à la question tout son intérêt, il ne faut pas se limiter, comme le font J. Bass et E. Arnous, à des cas purs, mais qu'il faut envisager le problème dans le cadre de la théorie des mélanges. La situation est représentée alors par un noyau $\mathcal{N}(x, y)$ self-adjoint auquel la théorie quantique impose d'être normé et d'avoir des valeurs propres positives. Dans la Note précédente, j'ai associé à tout opérateur différentiel self-adjoint G opérant sur les fonctions de x une fonction de l'espace et de l'impulsion $g(x, p)$ et à tout noyau self-adjoint \mathcal{N} une fonction $n(x, p)$ qui permettent de calculer la moyenne relative à G suivant la relation

$$(1) \quad \bar{G} = \iint g^*(x, p) n(x, p) dx^* dp.$$

Ce formalisme est déjà le formalisme de la statistique classique, mais utilise des quantités complexes. Mais, formons maintenant

$$(2) \quad \gamma(x, p) = \exp\left(\frac{1}{2} i \hbar \theta\right) g(x, p),$$

$$(3) \quad D(x, p) = \frac{1}{\pi \hbar} \iint \exp\left[2i \frac{(p-q)(x-y)}{\hbar}\right] n(y, q) dy dq.$$

A l'aide de ces fonctions, l'hermiticité de G et de \mathcal{N} s'exprime plus simplement qu'à l'aide de g et de n [formules (2) et (3) de la précédente Note]. Pour que G et \mathcal{N} soient self-adjoints, il faut et il suffit en effet que, respectivement, γ et D soient réels. g et n s'expriment en fonction de γ et D par les formules inverses de (2) et (3),

$$(4) \quad g(x, p) = \exp\left(-\frac{1}{2} i \hbar \theta\right) \gamma(x, p),$$

$$(5) \quad n(x, p) = \frac{1}{\pi \hbar} \iint \exp\left[2i \frac{(p-q)(y-x)}{\hbar}\right] D(y, q) dy dq.$$

⁽³⁾ *Comptes rendus*, 218, 1944, p. 141.

⁽⁴⁾ *Comptes rendus*, 221, 1945, p. 489.

⁽⁵⁾ *Comptes rendus*, 223, 1946, p. 311.

Introduisons γ et D dans (1) à la place de g et n . L'équation de la moyenne devient, après quelques transformations,

$$(6) \quad G = \iint \gamma(x, p) D(x, p) dx dp.$$

Maintenant le formalisme est beaucoup plus près du formalisme classique. La densité de probabilité D permet donc de calculer une moyenne quelconque. Ce progrès a été acquis en introduisant la fonction γ qui peut se déduire directement de l'opérateur G en effectuant le calcul suivant

$$(7) \quad \gamma(x, p) = \exp\left(\frac{\tau}{2} i \hbar \theta\right) \left[\exp\left(-i \frac{p \cdot x}{\hbar}\right) G \exp\left(i \frac{p \cdot x}{\hbar}\right) \right].$$

La différence fondamentale entre la théorie des mélanges et la statistique classique est que, en statistique classique, D est essentiellement positif, cependant qu'en statistique quantique D , essentiellement, correspond à un noyau \mathcal{N} dont les valeurs propres sont positives, ce qui est évidemment beaucoup moins simple. Cette dernière condition est l'expression du principe d'incertitude : elle empêche que D puisse présenter un maximum infiniment aigu à la fois par rapport aux deux variables x et p . Je rappellerai la formule de Wigner qui permet de calculer directement D à partir de \mathcal{N} . C'est :

$$(8) \quad D(x, p) = \int \exp\left(i \frac{p \cdot X}{\hbar}\right) \mathcal{N}\left(x - \frac{X}{2}, x + \frac{X}{2}\right) dX.$$

Ces nouveaux résultats me paraissent entraîner que, contrairement à ce que pouvait penser Wigner, il n'existe pas d'autre fonction analogue qui soit susceptible de jouir des mêmes propriétés. D est défini d'une manière unique.

Quant à la formule (7), elle fait correspondre univoquement et sans ambiguïté des expressions de forme classique aux grandeurs quantiques.

Une application immédiate des remarques précédentes est l'énoncé suivant (il n'est actuellement justifié naturellement que dans le domaine quantique très restreint où je me suis placé) :

Lorsque, dans la théorie des mélanges, on fait tendre la constante de Planck vers zéro, cette théorie se confond, à la limite, avec la mécanique statistique classique.

Le raisonnement relatif à ce passage à la limite fait naturellement intervenir l'équation évolutive de D , qui est une conséquence de l'équation de Schrödinger, que Wigner a formulée et qui tend à la limite vers l'équation de Liouville.

case for each and every k , the k th 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 $(0 \oplus M_b \oplus M_b \otimes' M_b \oplus \dots) \otimes (0 \oplus M_f \oplus M_f \otimes_1 M_f \oplus \dots)$, 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 $A_1 S A_1$, 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.

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

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f dq_1 \dots dq_n dp_1 \dots dp_n = 1, \tag{1}$$

(normalization)

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |f|^2 dq_1 \dots dq_n dp_1 \dots dp_n \text{ exists,} \tag{2}$$

(boundedness)

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

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

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

$$(A, B) = \left(\frac{2}{\hbar}\right)^{2n} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \cos \left\{ \frac{2}{\hbar} \sum_{j=1}^n \det \begin{vmatrix} 1 & q_j & p_j \\ 1 & \tau_j & \sigma_j \\ 1 & \xi_j & \eta_j \end{vmatrix} \right\}$$

$$\times A(\tau_k, \sigma_k) B(\xi_k, \eta_k) d\xi_1 \dots d\xi_n d\eta_1 \dots d\eta_n d\tau_1 \dots d\tau_n d\sigma_1 \dots d\sigma_n,$$

and

$$[A, B] = 2 \left(\frac{2}{\hbar}\right)^{2n} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \sin \left\{ \frac{2}{\hbar} \sum_{j=1}^n \det \begin{vmatrix} 1 & q_j & p_j \\ 1 & \tau_j & \sigma_j \\ 1 & \xi_j & \eta_j \end{vmatrix} \right\}$$

$$\times A(\tau_k, \sigma_k) B(\xi_k, \eta_k) d\xi_1 \dots d\xi_n d\eta_1 \dots d\eta_n d\tau_1 \dots d\tau_n d\sigma_1 \dots d\sigma_n.$$

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 anti-commutator 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

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

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

$$f = \frac{2^{2n}}{h^n} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \exp \left\{ \frac{2i}{\hbar} \sum_{j=1}^n [\rho_j(\tau_j - \xi_j) - \sigma_j(q_j - \xi_j) - \eta_j(\tau_j - q_j)] \right\} f(\tau_k, \sigma_k) f(\xi_k, \eta_k) d\xi_1 \dots$$

$$\times d\xi_n d\eta_1 \dots d\eta_n d\tau_1 \dots d\tau_n d\sigma_1 \dots d\sigma_n,$$

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

Let us make a change of variables of integration:

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

Then

$$f = \left(\frac{2}{n}\right)^n \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \exp \left\{ \frac{2i}{\hbar} \sum_{j=1}^n [p_j y_j - \frac{1}{2} \sigma_j (q_j + y_j - w_j) + \frac{1}{2} \eta_j (q_j - y_j - w_j)] \right\} f((w_k + y_k + q_k)/2, \sigma_k) \\ \times f((w_k + q_k - y_k)/2, \eta_k) dw_1 \cdots dw_n dy_1 \cdots dy_n \\ \times d\eta_1 \cdots d\eta_n d\sigma_1 \cdots d\sigma_n.$$

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

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

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

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} |G(q_k, w_k)|^2 dq_1 \cdots dq_n dw_1 \cdots dw_n \geq \sum_{i=1}^{\infty} \frac{1}{|\lambda_i|^2},$$

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

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(q_k, p_k) dq_1 \cdots dq_n dp_1 \cdots dp_n = 1,$$

by the normalization of f . Hence

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

⁴ E. C. Titchmarsh, *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

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

It is this form which Wigner⁶ chose "from all possible expressions, because it seems to be the simplest," although he knew only that it gave the correct marginal distributions. Moyal⁶ 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. Groenwold⁸ and Takabayasi⁹ 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 NEUMANN²

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

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

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

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

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).



The formulation of quantum mechanics in terms of phase space functions

By D. B. FAIRLIE

St Salvator's College, St Andrews

(Received 9 November 1963)

Abstract. A relationship between the Hamiltonian of a system and its distribution function in phase space is sought which will guarantee that the average energy is the weighted mean of the Hamiltonian over phase space. This relationship is shown to imply the existence of a wave function satisfying the Schrödinger equation, and dictates the possible forms of time-dependence of the distribution function.

The re-formulation of established theory in terms of new principles and alternative hypotheses sometimes facilitates the solution of specific problems, but chiefly illuminates the structure of the theory. Such is the case with the development of quantum mechanics in terms of the distribution function in phase space. We shall develop an eigenvalue equation for the distribution function, which may be solved for the case of a particle in a harmonic oscillator potential, and which has the same solution as that given by previous authors on the basis of Wigner's original introduction of the phase space function in terms of wave functions (7).

Baker (1) has shown that this wave-function expression may be deduced from certain results for the phase space function obtained by Moyal (5) and Takabayasi (6). However, his postulated condition for a stationary (time-independent) distribution function is by no means intuitively obvious. We shall deduce his results and the time-dependence of the distribution function from the hypothesis that this function is determined by the Hamiltonian so that the average energy is automatically constant. For reasons of notational simplicity the discussion will be confined to a two-dimensional phase space.

We postulate the existence of a bounded (i.e. square integrable) time-dependent distribution function $f(x, p, t)$ such that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, p) f(x, p, t) dx dp = E \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, p, t) dx dp, \quad (1)$$

where E , the energy of the system, is time-independent provided $H(x, p)$, the Hamiltonian, is. We seek an integral equation for f such that (1) will be automatically true.

Consider

$$I(x, p, t) = \frac{4}{\hbar^2} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \frac{2i}{\hbar} [p(\tau - \sigma) + x(\mu - \lambda) + \alpha(\sigma\lambda - \tau\mu)] \\ \times \phi(\alpha) H(\tau, \lambda) f(\sigma, \mu, t) d\tau d\lambda d\sigma d\mu d\alpha, \quad (2)$$

where $\int_{-\infty}^{\infty} \phi(\alpha) d\alpha = 1$ and \hbar is a constant, to be identified later with Planck's constant.

Then

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, p, t) dx dp &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x', p') f(x', p', t) dx' dp' \\ &= E \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', p', t) dx' dp'. \end{aligned} \quad (3)$$

We postulate that there are no intrinsically distinguished points in phase space; i.e. any local continuous mapping of x, p to x', p' should leave either all points or none fixed. These mappings are restricted to be pure translations by Brouwer's Translation Theorem (3). (I am indebted to Dr P. H. H. Fantham for this reference.) Thus we may have

$$I(x, p, t) = E f(x, p, t), \quad (4)$$

or more generally

$$I(x, p, t) = E f(x + a, p + b, t) \quad (4a)$$

as possible local relations between I and f . In other words the relationship is chosen to be translation invariant. We further require that if $H = \text{constant}$ there is no restriction on f . This will be so if $\phi(\alpha) = \delta(\alpha - 1)$, and the more general form of identification (4a) is rejected.

We shall now introduce the sine and cosine bracket notation of Baker and write

$$\begin{aligned} [f(x, p), g(x, p)] &= \frac{4}{h^2} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \sin \frac{2}{h} (p(\tau - \sigma) + x(\mu - \lambda) + \lambda\sigma - \tau\mu) \\ &\quad \times f(\tau, \lambda) g(\mu, \sigma) d\tau d\lambda d\mu d\sigma, \\ (f(x, p), g(x, p)) &= \frac{4}{h^2} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \cos \frac{2}{h} (p(\tau - \sigma) + x(\mu - \lambda) + \lambda\sigma - \tau\mu) \\ &\quad \times f(\tau, \lambda) g(\mu, \sigma) d\tau d\lambda d\mu d\sigma \end{aligned}$$

and employ the abbreviation

$$f \cdot g = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, p) g(x, p) dx dp.$$

The undernoted identities, which follow by formal calculation from the above definitions (1), are quoted for convenience.

$$\left. \begin{aligned} [A, B] \cdot C &= [C, A] \cdot B = [B, C] \cdot A, \\ [A, B] &= -[B, A], \\ (A, B) &= (B, A), \\ (A, B) \cdot C &= (C, A) \cdot B, \quad \text{etc.}, \\ [A, [B, C]] + [C, [A, B]] + [B, [C, A]] &= 0, \\ [A, (B, C)] &= ([A, B], C) + ([A, C], B). \end{aligned} \right\} \quad (5)$$

We may then write the condition (4) as

$$(H, f) + i[H, f] = 2Ef. \quad (6)$$

Consider two eigenfunctions f_i and f_j with real eigenvalues E_i and E_j , respectively, which satisfy (6). (The eigenvalues must be real since they represent the possible

energy values of the system.) Then with the aid of (5) we may deduce the following equations

$$H \cdot \{(f_i f_j^*) + i[f_i f_j^*]\} = 2E_i f_i \cdot f_j^*, \quad (7)$$

$$H \cdot \{(f_i^* f_j) - i[f_i^* f_j]\} = 2E_j f_i^* \cdot f_j. \quad (8)$$

Let us examine how far (7) and (8) are consistent with the assumption that equations of type (6) hold for both $(f_i f_j^*) + i[f_i f_j^*]$ and its complex conjugate, i.e.

$$(H\{(f_i f_j^*) + i[f_i f_j^*]\}) + i[H\{(f_i f_j^*) + i[f_i f_j^*]\}] = 2E_i\{(f_i f_j^*) + i[f_i f_j^*]\}, \quad (9)$$

$$(H\{(f_i^* f_j) - i[f_i^* f_j]\}) + i[H\{(f_i^* f_j) - i[f_i^* f_j]\}] = 2E_j\{(f_i^* f_j) - i[f_i^* f_j]\}. \quad (10)$$

Equations (9) and (10) are consistent with (7) and (8) upon integration.

We may then label the eigenfunctions of (6) according to their eigenvalues and those of their complex conjugates as

$$(f_i f_j^*) + i[f_i f_j^*] = k f_{ij}, \quad (11)$$

where k is a normalization constant.

By definition $f_{ij}^* = f_{ji}$. (12)

Rearrangement of (9) and (10) with their complex conjugate equations yields

$$(H, f_{ij}) = (E_i + E_j) f_{ij}, \quad (13)$$

$$[H, f_{ij}] = -i(E_i - E_j) f_{ij}. \quad (14)$$

Equation (14) has been given by Moyal (5) and is already implicit in previous work (7): equation (13) does not appear to have been stated before.

Comparison of (7) with the complex conjugate of (8) provides the further consistency requirement

$$f_i \cdot f_j^* = 0 \quad \text{unless} \quad E_i = E_j. \quad (15)$$

Similarly, $f_i \cdot f_j^* = 0$ unless f_i^* and f_j^* also have equal eigenvalues. In terms of the notation introduced above

$$f_{il} \cdot f_{mj} = 0 \quad \text{unless} \quad i = j \quad \text{and} \quad l = m. \quad (16)$$

This is the orthogonality condition for the eigenvalues of (6). (11) may be rewritten as

$$(f_{il}, f_{mj}) + i[f_{il}, f_{mj}] = K f_{ij}. \quad (17)$$

Then by (16)

$$f_{ji} \cdot (f_{il}, f_{mj}) + f_{ji} \cdot [f_{il}, f_{mj}] \neq 0,$$

or

$$\{(f_{ji}, f_{il}) + i[f_{ji}, f_{il}]\} \cdot f_{mj} \neq 0$$

by (5). Applying (16) again, we require

$$(f_{ji}, f_{il}) + i[f_{ji}, f_{il}] = f_{jm}.$$

This is possible only if $m = l$. Thus we see that for consistency, (17) must be modified to read

$$(f_{il}, f_{mj}) + i[f_{il}, f_{mj}] = K \delta_{im} f_{ij}. \quad (18)$$

In consequence of the orthogonality condition (16) $c_{ijl} = 0$ unless $l = j$. We may drop the third suffix and write the consequence of choosing $l = j$ as

$$c_{ij} = c_{ji}^* \tag{27}$$

and the expansion of $\partial f_{ij}/\partial t$ as
$$\frac{\partial f_{ij}}{\partial t} = c_{ij} f_{ij}. \tag{28}$$

The result of differentiating (18) and substituting (28) gives a further condition on c_{ij} , i.e.

$$c_{kl} + c_{lj} = c_{kj} \quad (\text{all } k, l \text{ and } j). \tag{29}$$

Put $k = l$ in the above relationship: then $c_{kk} = 0$, or $(\partial/\partial t)f_{kk} = 0$, i.e. the real eigenfunctions f_{ii} are stationary. This permits the expansion of H in terms of the f_{ii} as

$$H = \sum E_i f_{ii}, \tag{30}$$

since H is time-independent and real, and this resolution of H guarantees the consistency of (18). Put $k = j$ in (29); then

$$c_{kl} + c_{lk} = 0, \tag{31}$$

i.e.
$$c_{kl} + c_{kl}^* = 0. \tag{32}$$

Thus c_{kl} is pure imaginary, and must be written in the form

$$c_{kl} = \pm i(F(k) - F(l)) \tag{33}$$

in order to satisfy (29). This is as far as we can go in the determination of c_{kl} without involving some dynamical principle; we have not used the fact that t is a time variable except in supposing the independence of E and H of it. We appeal to the correspondence principle to verify that the simplest non-trivial choice of arbitrary function in (33), namely

$$c_{kl} = \frac{i}{\hbar} (E_k - E_l) \tag{34}$$

will do. The factor $1/\hbar$ is necessary to give c_{kl} the dimension of an inverse time.

Then equations (13) and (14) take the form

$$(H, f_{ij}) = (E_i + E_j) f_{ij}, \tag{35}$$

$$[H, f_{ij}] = \hbar \frac{\partial f_{ij}}{\partial t} \tag{36}$$

and the second equation in the limit $\hbar \rightarrow 0$ is just the classical Liouville equation, thus verifying (34).

Finally, as an example of the simultaneous solution of the above system we consider the harmonic oscillator Hamiltonian

$$H = \frac{1}{2}(p^2 + x^2) \hbar \omega. \tag{37}$$

If we write
$$f_{kl}(x, p, t) = u_{kl}(x, p) \exp i(k-l) \omega t \tag{38}$$

both (35) and (36) may be written as differential equations of finite order

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial p^2} \right) u_{kl} - 4(x^2 + p^2) u_{kl} + 2(k+l-1) u_{kl} = 0, \tag{39}$$

$$-p \frac{\partial u_{kl}}{\partial x} + x \frac{\partial u_{kl}}{\partial p} - (k-l) u_{kl} = 0, \tag{40}$$

Wigner Representation of Quantum Operators and Its Applications to Electrons in a Magnetic Field

Ryogo KUBO*

*Department of Physics, University of Pennsylvania
and Department of Physics, University of Tokyo*

(Received June 3, 1964)

The Wigner representation $F_s(\mathbf{p}, \mathbf{q})$ of a quantum operator $F(\mathbf{p}, \mathbf{q})$ is defined by $F_s(\mathbf{p}, \mathbf{q}) = \hbar \text{Tr} F(\mathbf{p}, \mathbf{q}) \Delta(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q})$ where Δ is a quantum analogue of a delta function in the phase space. This gives in particular the Wigner distribution function for a density operator. Basic theorems are summarized for the computation rules for quantummechanical operators in the Wigner representation. This is applied, in particular, to electrons in a magnetic field, for which a Wigner d. f. is introduced to describe the distribution of physical momenta $\vec{\pi} = m\vec{v}$, \vec{v} being the velocity, and the position \mathbf{x} . This description has the advantage to avoid the use of a vector potential and so to be gauge-independent. As examples of application, the diamagnetism and the Hall effect are briefly treated. Further applications of this treatment will be published later.

§ 1. Introduction

Many years ago, Wigner¹⁾ introduced into the quantum mechanics a phase-space distribution function which is an analogue to the distribution function of classical statistical mechanics and is called the Wigner distribution function (abbreviated in the following as Wigner d. f.). A Wigner d. f. has no definite sign unlike usual distribution functions, but nevertheless it provides a nice way of formulating quantum mechanics as a probabilistic theory. This point of view was extensively examined by Moyal²⁾. For a practical purpose of computation, Wigner distribution functions are particularly useful in order to obtain quantum corrections to classical formulae, because it gives a systematic method of expanding physical quantities in terms of \hbar , or the value of the non-commuting variables. Thus, Wigner d. f. has been extensively used in the theories of gaseous and liquid systems³⁾. Now, similar applications can be made for treating electrons in a magnetic field. If the magnetic field H lies in the z -direction and is uniform in space, the physical momenta,

$$\pi_x = mv_x, \quad \pi_y = mv_y$$

m being the mass and v_x and v_y , the velocity components, have the commutator

$$[\pi_x, \pi_y] = -\frac{\hbar e H}{c i} \equiv -\frac{\hbar'}{i}, \quad \hbar' \equiv \frac{\hbar e H}{c}.$$

If it is desired to obtain an expansion in terms of H or \hbar' , then the Wigner d. f. will be found to be a very useful tool. The purpose of the present paper is to show a few examples of such application of Wigner d. f.⁴⁾ However, considering the fact that only few literatures are available for the reference of basic theorems concerning the Wigner d. f., Part I of this paper is devoted to a brief summary of such theorems, some of which are presented there in somewhat new forms although most of them have been already known.

Part I. Wigner Representation

§ 2. Symmetrized Operators

For simplicity's sake, most of the following formulae are written for a system with only one degree of freedom, the canonical variables of which will be denoted by p and q , but they apply to many degrees of freedom with obvious modifications. Gothic letters, as \mathbf{p} or \mathbf{q} , mean quantum-mechanical operators and corresponding classical variables are denoted by p or q .

A function of p and q , $A_s(\mathbf{p}, \mathbf{q})$, is said to be a symmetrized operator if it has the form

* Present Address: Department of Physics, University of Tokyo.

$$A_s(\mathbf{p}, \mathbf{q}) = \int_{-\infty}^{\infty} d\xi d\eta G_A(\xi, \eta) \exp i(\xi\mathbf{p} + \eta\mathbf{q}) \quad (2.1)$$

because, as evident from the definition of the exponential function, any product of \mathbf{p} and \mathbf{q} in this function is totally symmetrized with respect to their order of multiplication. Generally speaking, a quantum mechanical operator is not necessarily symmetrized, but it can be expressed in terms of symmetrized operators. For instance in

$$\mathbf{p}\mathbf{q} = \frac{1}{2}(\mathbf{p}\mathbf{q} + \mathbf{q}\mathbf{p}) + \frac{\hbar}{2i} \equiv \{\mathbf{p}\mathbf{q}\}_s + \frac{\hbar}{2i},$$

the left hand side is not symmetrized, whereas the right hand side is so. A prescription how to rewrite an operator function, $F(\mathbf{p}, \mathbf{q})$ into a symmetrized form $F_s(\mathbf{p}, \mathbf{q})$ is given by the following theorem. Other equivalent methods will be given later.

Theorem 1. For a given operator function, $F(\mathbf{p}, \mathbf{q})$, define $G(\xi, \eta)$ by

$$\begin{aligned} G(\xi, \eta) &= \text{Tr } F(\mathbf{p}, \mathbf{q}) \exp \{-i(\xi\mathbf{p} + \eta\mathbf{q})\}, \quad (2.2) \\ &\equiv \int_{-\infty}^{\infty} \langle q' - \hbar\xi | F | q' \rangle dq' \exp \left\{ -i\eta \left(q' - \frac{\hbar\xi}{2} \right) \right\}, \end{aligned} \quad (2.3)$$

where (2.3) is an explicit form of (2.2) in the q -representation. Then the operator $F(\mathbf{p}, \mathbf{q})$ can be expressed by

$$\begin{aligned} F(\mathbf{p}, \mathbf{q}) &= F_s(\mathbf{p}, \mathbf{q}) \\ &= \frac{\hbar}{2\pi} \int_{-\infty}^{\infty} G(\xi, \eta) d\xi d\eta \exp i(\xi\mathbf{p} + \eta\mathbf{q}), \quad (2.4) \end{aligned}$$

which gives its symmetrized form.

This theorem may be stated in an alternative form:

Theorem 2. For a given operator function, $F(\mathbf{p}, \mathbf{q})$, define

$$\begin{aligned} F_s(\mathbf{p}, \mathbf{q}) &= \hbar \text{Tr } F(\mathbf{p}, \mathbf{q}) \mathcal{A}(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}) \quad (2.5) \\ &= \int \left\langle q - \frac{\hbar\xi}{2} \left| F \right| q + \frac{\hbar\xi}{2} \right\rangle e^{i\xi\mathbf{p}} \hbar d\xi. \quad (2.6) \end{aligned}$$

Then we have

$$\begin{aligned} F(\mathbf{p}, \mathbf{q}) &= F_s(\mathbf{p}, \mathbf{q}) \\ &= \iint d\mathbf{p} d\mathbf{q} F_s(\mathbf{p}, \mathbf{q}) \mathcal{A}(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}), \quad (2.7) \end{aligned}$$

where

$$\begin{aligned} \mathcal{A}(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}) \\ = \frac{1}{(2\pi)^2} \iint_{-\infty}^{\infty} \exp i(\xi(\mathbf{p} - \mathbf{p}) + \eta(\mathbf{q} - \mathbf{q})). \quad (2.8) \end{aligned}$$

is a quantum-analogue of a delta function in the phase space.

For the proof of the theorems, we refer to a lemma which is due to Moyal:

Lemma

$$\begin{aligned} \exp i(\xi\mathbf{p} + \eta\mathbf{q}) &= e^{i\xi\mathbf{p}} e^{i\eta\mathbf{q}} e^{-i\hbar\xi\eta/2} = e^{i\eta\mathbf{q}} e^{i\xi\mathbf{p}} e^{i\hbar\xi\eta/2} \\ &= e^{i\xi\mathbf{p}/2} e^{i\eta\mathbf{q}} e^{i\xi\mathbf{p}/2} = e^{i\eta\mathbf{q}/2} e^{i\xi\mathbf{p}} e^{i\eta\mathbf{q}/2}, \quad (2.9) \end{aligned}$$

or in the q -representation,

$$\begin{aligned} \langle q' | \exp i(\xi\mathbf{p} + \eta\mathbf{q}) | q'' \rangle \\ = \exp i\eta \left(q' + \frac{1}{2} \hbar\xi \right) \delta(q' + \hbar\xi - q''). \quad (2.10) \end{aligned}$$

These are easily obtained from the commutation rule of \mathbf{p} and \mathbf{q} .

By Eq. (2.10), (2.2) gives (2.3), which is inserted into (2.4) to give

$$\begin{aligned} \langle q' | F_s(\mathbf{p}, \mathbf{q}) | q'' \rangle \\ = \frac{\hbar}{2\pi} \iint_{-\infty}^{\infty} \exp \left\{ i\eta \left(q' + \frac{\hbar}{2} \xi \right) \right\} \delta(q' + \hbar\xi - q'') d\xi d\eta \\ \times \langle q'' - \hbar\xi | F | q'' \rangle dq'' \exp \left\{ -i\eta \left(q'' - \frac{\hbar}{2} \xi \right) \right\} \\ = \langle q' | F | q'' \rangle. \end{aligned}$$

This establishes the equality of $F(\mathbf{p}, \mathbf{q})$ and $F_s(\mathbf{p}, \mathbf{q})$, the latter being symmetrized by its definition (2.4).

By Eqs. (2.2) and (2.8), $F_s(\mathbf{p}, \mathbf{q})$ is in fact equal to

$$F_s(\mathbf{p}, \mathbf{q}) = \frac{\hbar}{2\pi} \iint_{-\infty}^{\infty} G(\xi, \eta) d\xi d\eta e^{i(\xi\mathbf{p} + \eta\mathbf{p})}. \quad (2.11)$$

Thus Eq. (2.7) is equivalent to Eq. (2.4). Equation (2.6) follows from (2.3) and (2.11).

Now, let ρ be a density operator to describe a statistical property of a quantum system which may be in a pure or mixed state.

Then we have:

Theorem 3. The Wigner d. f. for the density operator ρ is defined by

$$\begin{aligned} f(\mathbf{p}, \mathbf{q}) &= \hbar \langle \mathcal{A}(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}) \rangle \\ &\equiv \hbar \text{Tr } \rho \mathcal{A}(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}). \quad (2.12) \end{aligned}$$

Then ρ is represented by

$$\rho(\mathbf{p}, \mathbf{q}) = \iint d\mathbf{p} d\mathbf{q} f(\mathbf{p}, \mathbf{q}) \mathcal{A}(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}). \quad (2.13)$$

This is a direct consequence of the preceding theorems. Equation (2.12) is equivalent to the more familiar definition of Wigner d. f.,

$$f(p, q) = \int \left\langle q - \frac{r}{2} | \rho | q + \frac{r}{2} \right\rangle e^{irp/\hbar} dr, \quad (2.14)$$

as is seen at once from Eq. (2.6). This can be obtained by the matrix representation of A ,

$$\begin{aligned} &\langle q' | A(p-\mathbf{p}, q-\mathbf{q}) | q'' \rangle \\ &= \frac{1}{2\pi} \int d\xi \delta\left(q' - q + \frac{\hbar\xi}{2}\right) \delta(q' - q'' - \hbar\xi) e^{-i\xi p}, \end{aligned} \quad (2.15)$$

which follows from the lemma (2.10). Equation (2.15) also gives

$$\text{Tr } A(p-\mathbf{p}, q-\mathbf{q}) = \frac{1}{h}, \quad (2.16)$$

and

$$\iint dpdq A(p-\mathbf{p}, q-\mathbf{q}) = 1. \quad (2.17)$$

Accordingly the normalization

$$\text{Tr } \rho(\mathbf{p}, \mathbf{q}) = \frac{1}{h} \iint dpdq f(p, q), \quad (2.18)$$

is proved.

It should be remembered that a density operator $\rho(\mathbf{p}, \mathbf{q})$, for example,

$$\rho(\mathbf{p}, \mathbf{q}) = e^{-\beta \mathcal{H}(\mathbf{p}, \mathbf{q})},$$

is not symmetrized. Therefore the calculation of the corresponding Wigner d. f. is equivalent to rewriting it into a *symmetrized form*.

§ 3. Wigner Operators

We have seen that a phase function, $F_s(p, q)$, is found for a given operator $F(\mathbf{p}, \mathbf{q})$ by Eq. (2.5). This is generally a one-to-one correspondence, so that $F_s(p, q)$ may be regarded as a representation of the operator $F(\mathbf{p}, \mathbf{q})$, which will be called the *Wigner representation* of F . The functional form of $F_s(p, q)$ is identical with that of $F(\mathbf{p}, \mathbf{q})$ if and only if the latter is already symmetrized. Generally, a product of two operators, A and B , is not symmetrized even when they are symmetrized. Thus, we have to find the rule to construct the Wigner representation of such a product from the known Wigner representations of A and B . For this purpose, we introduce differential operators, which will be called *Wigner operators*, by the following definitions:

$$A_w \equiv A_s\left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q}, q - \frac{\hbar}{2i} \frac{\partial}{\partial p}\right)$$

$$\begin{aligned} &= \frac{\hbar}{2\pi} \iint d\xi d\eta G_A(\xi, \eta) \\ &\quad \times \exp i \left\{ \xi \left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q} \right) + \eta \left(q - \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) \right\} \end{aligned} \quad (3.1)$$

$$\begin{aligned} \bar{A}_w &\equiv A_s\left(p - \frac{\hbar}{2i} \frac{\partial}{\partial q}, q + \frac{\hbar}{2i} \frac{\partial}{\partial p}\right) \\ &= \frac{\hbar}{2\pi} \iint d\xi d\eta G_A(\xi, \eta) \\ &\quad \times \exp i \left\{ \xi \left(p - \frac{\hbar}{2i} \frac{\partial}{\partial q} \right) + \eta \left(q + \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) \right\}, \end{aligned} \quad (3.2)$$

where

$$\begin{aligned} G_A(\xi, \eta) &= \frac{1}{h} \iint A_s(p, q) e^{-i(\xi p + \eta q)} dpdq \\ &= \text{Tr } A(\mathbf{p}, \mathbf{q}) \exp \{-i(\xi \mathbf{q} + \eta \mathbf{p})\}. \end{aligned} \quad (3.3)$$

By Eq. (2.11), $A_s(p, q)$ is the Wigner representation of the operator, $A = A(\mathbf{p}, \mathbf{q})$. We then have:

Theorem 4. Let $A_s(p, q)$ and $B_s(p, q)$ be the Wigner representations of the operators A and B . Then

$$A_w \cdot 1 = \bar{A}_w \cdot 1 = A_s(p, q), \quad (3.4)$$

and

$$\begin{aligned} &\text{Wigner representation of } \mathbf{AB} \\ &= A_w B_w \cdot 1 = \bar{B}_w A_w \cdot 1. \end{aligned} \quad (3.5)$$

For the proof, we note the identities

$$\begin{aligned} &\exp i(\xi_1 \mathbf{p} + \eta_1 \mathbf{q}) \exp i(\xi_2 \mathbf{p} + \eta_2 \mathbf{q}) \\ &= \exp \left[i\{(\xi_1 + \xi_2)\mathbf{p} + (\eta_1 + \eta_2)\mathbf{q}\} + \frac{i\hbar}{2}(\xi_1 \eta_2 - \eta_1 \xi_2) \right], \end{aligned} \quad (3.6)$$

$$\begin{aligned} &\exp i \left\{ \xi \left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q} \right) + \eta \left(q - \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) \right\}, \\ &= \exp i(\xi p + \eta q) \exp \frac{\hbar}{2} \left(\xi \frac{\partial}{\partial q} - \eta \frac{\partial}{\partial p} \right), \end{aligned} \quad (3.7a)$$

$$= \exp \frac{\hbar}{2} \left(\xi \frac{\partial}{\partial q} - \eta \frac{\partial}{\partial p} \right) \exp i(\xi p + \eta q), \quad (3.7b)$$

and

$$\begin{aligned} &\exp i \left\{ \xi_1 \left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q} \right) + \eta_1 \left(q - \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) \right\} \\ &\times \exp i \left\{ \xi_2 \left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q} \right) + \eta_2 \left(q - \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) \right\} \\ &= \exp i \left\{ (\xi_1 + \xi_2) \left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q} \right) \right. \\ &\quad \left. + (\eta_1 + \eta_2) \left(q - \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) \right\} \exp \frac{i\hbar}{2}(\xi_1 \eta_2 - \eta_1 \xi_2), \end{aligned} \quad (3.8)$$

Equation (3.6) follows from the lemma (2.9), whereas (3.7) and (3.8) follow from a similar identity,

$$\begin{aligned} \exp\left(ax + b\frac{\partial}{\partial x}\right) &= e^{ax}e^{b(\partial/\partial x)}e^{ab/2} \\ &= e^{b(\partial/\partial x)}e^{ax}e^{-ab/2}. \end{aligned} \quad (3.9)$$

Equation (3.4) is easily seen from the definitions of A_w and \bar{A}_w , (3.1) and (3.2), by using Eq. (3.7a); namely,

$$\begin{aligned} A_w \cdot 1 &= \frac{\hbar}{2\pi} \int_{-\infty}^{\infty} d\xi d\eta G_A(\xi, \eta) \exp i(\xi p + \eta q) \\ &\quad \times \exp \frac{\hbar}{2} \left(\xi \frac{\partial}{\partial q} - \eta \frac{\partial}{\partial p} \right) \cdot 1 \\ &= \frac{\hbar}{2\pi} \int_{-\infty}^{\infty} d\xi d\eta G_A(\xi, \eta) \\ &\quad \times \exp i(\xi p + \eta q) = A_s(p, q), \end{aligned}$$

and similarly for $\bar{A}_w \cdot 1$. In order to show (3.5), we observe first that Eqs. (2.4) and (3.6) give

$$\begin{aligned} \mathbf{AB} &= \left(\frac{\hbar}{2\pi}\right)^2 \iiint G_A(\xi_1, \eta_1) G_B(\xi_2, \eta_2) d\xi_1 d\eta_1 d\xi_2 d\eta_2 \\ &\quad \times \exp \left[i(\xi_1 + \xi_2)p + (\eta_1 + \eta_2)q \right] \\ &\quad + \frac{i\hbar}{2} (\xi_1 \eta_2 - \eta_1 \xi_2) \Big], \end{aligned} \quad (3.10)$$

which shows that

Wigner representation of \mathbf{AB}

$$\begin{aligned} &= \left(\frac{\hbar}{2\pi}\right)^2 \iiint G_A(\xi_1, \eta_1) G_B(\xi_2, \eta_2) d\xi_1 d\eta_1 d\xi_2 d\eta_2 \\ &\quad \times \exp \left[i(\xi_1 + \xi_2)p + (\eta_1 + \eta_2)q \right] \\ &\quad + \frac{i\hbar}{2} (\xi_1 \eta_2 - \eta_1 \xi_2) \Big]. \end{aligned} \quad (3.11)$$

Now, use Eq. (3.1) for explicit expressions of A_w and B_w , Eq. (3.8) for the product of exponential operators and Eq. (3.7a) for operating the differential operator on 1; then one easily sees that $A_w B_w \cdot 1$ is in fact equal to (3.11). By a similar calculation the second equality in (3.5) can be established. This last point can, however, be stated by a more general theorem; i. e.,

Theorem 5. For Wigner operators defined by (3.1) and (3.2), we have the identity,

$$A_w \bar{B}_w = \bar{B}_w A_w, \text{ or } [A_w, \bar{B}_w] = 0. \quad (3.12)$$

This follows from the identity,

$$\begin{aligned} &\left[\exp i \left\{ \xi_1 \left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q} \right) + \eta_1 \left(q - \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) \right\}, \right. \\ &\left. \exp i \left\{ \xi_2 \left(p - \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) + \eta_2 \left(q + \frac{\hbar}{2i} \frac{\partial}{\partial p} \right) \right\} \right] = 0. \end{aligned} \quad (3.13)$$

This is proved by repeated use of the relations (3.7a, b) and (3.9). Therefore, Eq. (3.5) is obtained as

$$A_w B_w \cdot 1 = A^w \bar{B}_w \cdot 1 = \bar{B}_w A_w \cdot 1. \quad (3.14)$$

This means that the product of two operators, \mathbf{AB} , can be interpreted into the Wigner representation in two ways (which are of course equivalent to one another); namely, if \mathbf{A} multiplies \mathbf{B} from the left, A_w operates on B_w ; if \mathbf{B} multiplies \mathbf{A} from the right, \bar{B}_w operates on A_w . Therefore, we may call A_w , defined by (3.1), a *left Wigner operator*, and \bar{A}_w , defined by (3.2) a *right Wigner operator*. For example,

$$\begin{aligned} A_w B_w C_w D_w \cdot 1 &= A_w B_w \bar{D}_w C_w \cdot 1 = \dots \\ &= \bar{D}_w A_w B_w C_w \cdot 1 = \bar{D}_w A_w C_w \bar{B}_w \cdot 1 = \dots \\ &= \bar{D}_w C_w \bar{B}_w \bar{A}_w \cdot 1, \end{aligned} \quad (3.15)$$

is the Wigner representation of the operator product, \mathbf{ABCD} ; all of the expressions in Eq. (3.15) give the same result, because they merely correspond to different interpretation of multiplication from the left or from the right.

If an operator $F(\mathbf{p}, \mathbf{q})$ is given in the form (2.7), its trace is obtained by

$$\text{Tr } F(\mathbf{p}, \mathbf{q}) = \frac{1}{h} \iint dp dq F_s(p, q), \quad (3.16)$$

because of Eq. (2.16). Thus, taking the trace of a quantum operator is the phase-space integration of its Wigner representation. For this we have the basic theorem:

Theorem 6. The trace of \mathbf{AB} , if it exists, is given by

$$\text{Tr } \mathbf{AB} = \frac{1}{h} \iint dp dq A_s(p, q) B_s(p, q), \quad (3.17)$$

in the Wigner representation. This is proved as follows;

$$\begin{aligned} \text{Tr } \mathbf{AB} &= \frac{1}{h} \iint dp dq A_w B_w \cdot 1 \\ &= \frac{1}{h} \iint dp dq \frac{\hbar}{2\pi} \iint d\xi d\eta \\ &\quad \times \exp \frac{\hbar}{2} \left(\xi \frac{\partial}{\partial q} - \eta \frac{\partial}{\partial p} \right) \exp i(\xi p + \eta q) \\ &\quad \times G_A(\xi, \eta) B_s(p, q), \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{h} \iint dp dq \frac{\hbar}{2\pi} \iint d\xi d\eta e^{i(\xi p + \eta q)} G_A(\xi, \eta) B_s(p, q), \\
 &= \frac{1}{h} \iint dp dq A_s(p, q) B_s(p, q).
 \end{aligned}$$

The second line is obtained from the first line, which follows from (3.16) and (3.5), by writing A_w in the form (3.2) with the transformation (3.7b) and by using the relation (3.4) for $B_w \cdot 1$. The third line is the result of partial integration, the boundary values being assumed to vanish.

By the theorems 4, 5, 6, Eq. (3.15) and by the cyclic property of trace, the trace of a product, $ABCD$, for example, can be expressed in various forms, all of which give of course the identical value provided that the trace does converge. Thus, for example, we may write

$$\begin{aligned}
 \text{Tr } ABCD &= \frac{1}{h} \iint dp dq A_w B_w C_w D_s(p, q) \\
 &= \frac{1}{h} \iint dp dq A_s(p, q) B_w C_w D_s(p, q) \\
 &= \frac{1}{h} \iint dp dq (\bar{B}_w A_s(p, q)) (C_w D_s(p, q)) \\
 &= \frac{1}{h} \iint dp dq (A_w B_s(p, q)) (C_w D_s(p, q)) \\
 &= \frac{1}{h} \iint dp dq (D_w A_s(p, q)) (B_w C_s(p, q)), \text{ etc.}
 \end{aligned} \tag{3.18}$$

There exists another transformation which gives different appearance of the expression keeping, however, the value of the trace invariant. This is shown by

$$\begin{aligned}
 \text{Tr } AB \cdots D &= \frac{1}{h} \iint A_w B_w \cdots D_w \cdot 1 \\
 &= \frac{1}{h} \iint dp dq \exp\left(-\frac{\epsilon \hbar}{2i} \frac{\partial^2}{\partial p \partial q}\right) \\
 &\quad \times \exp\left(\frac{\epsilon \hbar}{2i} \frac{\partial^2}{\partial p \partial q}\right) A_w \cdots D_w \exp\left(-\frac{\epsilon \hbar}{2i} \frac{\partial^2}{\partial p \partial q}\right) \cdot 1 \\
 &= \frac{1}{h} \iint dp dq A_w^\epsilon \cdots D_w^\epsilon 1
 \end{aligned} \tag{3.19}$$

where

$$\begin{aligned}
 A_w^\epsilon &= \exp\left(\frac{\epsilon \hbar}{2i} \frac{\partial^2}{\partial p \partial q}\right) A_s\left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q}, q - \frac{\hbar}{2i} \frac{\partial}{\partial p}\right) \\
 &\quad \times \exp\left(-\frac{\epsilon \hbar}{2i} \frac{\partial^2}{\partial p \partial q}\right) \\
 &= A_s\left(p + \frac{\hbar(1+\epsilon)}{2i} \frac{\partial}{\partial q}, q - \frac{\hbar(1-\epsilon)}{2i} \frac{\partial}{\partial p}\right).
 \end{aligned} \tag{3.20}$$

The theorems and remarks mentioned above can be applied for calculations of statistical averages defined by a given density operator ρ . Thus we can write

$$\langle A \rangle = \text{Tr } A\rho = \frac{1}{h} \iint dp dq A_s(p, q) f(p, q), \tag{3.21}$$

or more generally

$$\begin{aligned}
 \langle AB \cdots D \rangle &= \text{Tr } AB \cdots D\rho \\
 &= \frac{1}{h} \iint dp dq A_w B_w \cdots D_w f(p, q),
 \end{aligned} \tag{3.22}$$

where A_w , etc. are Wigner operators operating on the Wigner d. f.. By such relations as (3.18) and (3.19), the expression (3.22) may also be written in a number of different ways.

Finally, we remark

Theorem 7. The Wigner operator A_w defined by (3.1) may be written as

$$A_w = \exp\left\{\frac{\hbar}{2i} \left(\frac{\partial}{\partial p_A} \frac{\partial}{\partial q} - \frac{\partial}{\partial q_A} \frac{\partial}{\partial p}\right)\right\} A_s(p, q), \tag{3.23}$$

where $\partial/\partial p_A$ or $\partial/\partial q_A$ means the differentiation of the function $A_s(p, q)$ while $\partial/\partial p$ or $\partial/\partial q$ operates on a function which follows after A_w ; namely

$$\begin{aligned}
 A_w f(p, q) &= \sum_{m, n} \frac{(-)^n \hbar^{n+m}}{m! n! (2i)^{n+m}} \frac{\partial^{n+m} A_s(p, q)}{\partial p^n \partial q^m} \frac{\partial^{n+m} f(p, q)}{\partial q^n \partial p^m}.
 \end{aligned} \tag{3.24}$$

The differentiations of A_s and f can be separated in this way by virtue of the identity (3.7a).

§ 4. Equation of Motion and Bloch Equation

By the basic theorems given in the preceding sections, it is easy to interpret quantum-mechanical expressions for operators into the Wigner representation or to interpret expressions in the Wigner representation back into more familiar quantum-operator expressions. In this way, one can choose either representation convenient of the purpose of calculation.

The equation of motion of the density matrix ρ ,

$$\frac{\partial \rho}{\partial t} = \frac{1}{it} [\mathcal{H}(\mathbf{p}, \mathbf{q}, t), \rho], \tag{4.1}$$

is now read, in the Wigner representation, as

$$\frac{\partial f}{\partial t} = \frac{1}{i\hbar} (\mathcal{H}_w - \overline{\mathcal{H}}_w) f \equiv i\mathcal{L}f, \quad (4.2)$$

where $\mathcal{H}(\mathbf{p}, \mathbf{q}, t)$ is the Hamiltonian, which may explicitly depend on the time t , and \mathcal{H}_w and $\overline{\mathcal{H}}_w$ are the corresponding Wigner operators. If $\mathcal{H}(\mathbf{p}, \mathbf{q}, t)$ is already a symmetrized operator, as it usually is, then the "Liouville operator" for the Wigner d. f. is given by

$$\begin{aligned} i\mathcal{L} &\equiv \frac{1}{i\hbar} \left\{ \mathcal{H} \left(p + \frac{\hbar}{2i} \frac{\partial}{\partial q}, q - \frac{\hbar}{2i} \frac{\partial}{\partial p}, t \right) \right. \\ &\quad \left. - \mathcal{H} \left(p - \frac{\hbar}{2i} \frac{\partial}{\partial q}, q + \frac{\hbar}{2i} \frac{\partial}{\partial p}, t \right) \right\} \\ &= \frac{1}{i\hbar} \{ \mathcal{H}(p_w, q_w, t) - \mathcal{H}(\overline{p}_w, \overline{q}_w, t) \}. \end{aligned} \quad (4.3)$$

If this is expanded in a power series of \hbar by the formula (3.24), the first term in $O(\hbar^0)$ is of course the classical Liouville operator, and the following terms appearing only in even powers of \hbar give its quantum corrections.

Correspondingly, the Heisenberg equation of motion of a dynamical variable $A = A(\mathbf{p}_t, \mathbf{q}_t)$,

$$\frac{d}{dt} A(\mathbf{p}_t, \mathbf{q}_t) = \frac{1}{i\hbar} [A(\mathbf{p}_t, \mathbf{q}_t), \mathcal{H}(\mathbf{p}_t, \mathbf{q}_t, t)] \quad (4.4)$$

is read as

$$\frac{d}{dt} A_s(\mathbf{p}_t, \mathbf{q}_t) = -i\mathcal{L}(\mathbf{p}_t, \mathbf{q}_t, t) A_s(\mathbf{p}_t, \mathbf{q}_t), \quad (4.5)$$

where $\mathcal{L}(\mathbf{p}_t, \mathbf{q}_t, t)$ has the same functional form as (4.3). Obviously, in the classical limit of $\hbar \rightarrow 0$, Eq. (4.5) is identical with the classical equation of motion

$$\begin{aligned} \frac{d}{dt} A(\mathbf{p}_t, \mathbf{q}_t) &= -i\mathcal{L}_{cl}(\mathbf{p}_t, \mathbf{q}_t, t) A(\mathbf{p}_t, \mathbf{q}_t) \\ &= \frac{\partial A}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i}, \end{aligned} \quad (4.6)$$

which is equivalent to the Hamilton equation,

$$\dot{p}_i = -\frac{\partial}{\partial q_i} \mathcal{H}(\mathbf{p}_t, \mathbf{q}_t, t), \quad \dot{q}_i = \frac{\partial}{\partial p_i} \mathcal{H}(\mathbf{p}_t, \mathbf{q}_t, t), \quad (4.7)$$

because the phase function $A(\mathbf{p}_t, \mathbf{q}_t)$ depends on t through the temporal change of \mathbf{p}_t and \mathbf{q}_t , which is determined by (4.7) for a given initial condition, e. g. $p_0 = p, q_0 = q$. In the same way, Eq. (4.5) determines the temporal change of the Wigner representation of a quantum operator $A(\mathbf{p}_t, \mathbf{q}_t)$ which obeys Eq. (4.4) with the corresponding initial condition

$$A(\mathbf{p}_0, \mathbf{q}_0) = A(\mathbf{p}, \mathbf{q}).$$

The explicit time dependence of the Wigner representation,

$$A_s(\mathbf{p}_t, \mathbf{q}_t) = A(t; \mathbf{p}, \mathbf{q}), \quad (4.8)$$

with the initial condition

$$A_s(0; \mathbf{p}, \mathbf{q}) = A_s(\mathbf{p}, \mathbf{q}), \quad (4.9)$$

may be defined by

$$A_s(t; \mathbf{p}, \mathbf{q}) = \hbar \text{Tr} A(\mathbf{p}_t, \mathbf{q}_t) \Delta(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}). \quad (4.10)$$

In particular,

$$\begin{aligned} p_t &= \hbar \text{Tr} \mathbf{p}_t \Delta(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}), \\ q_t &= \hbar \text{Tr} \mathbf{q}_t \Delta(\mathbf{p} - \mathbf{p}, \mathbf{q} - \mathbf{q}), \end{aligned} \quad (4.11)$$

give the Wigner representations of canonical variables with the initial condition, $p_0 = p, q_0 = q$. By differentiating Eq. (4.10) with respect to t and by using Eq. (4.4), (4.11) and the theorems 3 and 4, we derive

$$\frac{d}{dt} A_s(t; \mathbf{p}, \mathbf{q}) = -i\mathcal{L}(\mathbf{p}_t, \mathbf{q}_t, t) A_s(t; \mathbf{p}, \mathbf{q}). \quad (4.12)$$

In particular, p_t and q_t defined by (4.11) satisfy

$$\frac{dp_t}{dt} = -i\mathcal{L} p_t, \quad \frac{dq_t}{dt} = -i\mathcal{L} q_t. \quad (4.13)$$

This shows that the solution of (4.12) is given by (4.8). This can be directly seen by noticing that the formal solution of Eq. (4.12) can be written as

$$\begin{aligned} A_s(t; \mathbf{p}, \mathbf{q}) &= S(t) A_s(\mathbf{p}, \mathbf{q}) \\ &\equiv S(t) A_s(\mathbf{p}, \mathbf{q}) S(t)^{-1}, \end{aligned} \quad (4.14)$$

where

$$\begin{aligned} S(t) &= \exp \left(-i \int_0^t \mathcal{L}(\mathbf{p}, \mathbf{q}, t') dt' \right) \\ S(t)^{-1} &= \exp \left(i \int_0^t \mathcal{L}(\mathbf{p}, \mathbf{q}, t') dt' \right), \end{aligned} \quad (4.15)$$

are ordered exponential operators ordered chronologically as indicated by arrows. Since \mathcal{L} is a differential operator, the last factor $S^{-1}(t)$ of (4.4) results in 1. In the same way we have

$$\begin{aligned} p_t &= S(t) p \equiv S(t) p S^{-1}(t) \\ q_t &= S(t) q \equiv S(t) q S^{-1}(t) \\ \frac{\partial}{\partial p_i} &= S(t) \frac{\partial}{\partial p} S(t)^{-1}, \quad \frac{\partial}{\partial q_i} = S(t) \frac{\partial}{\partial q} S(t)^{-1}. \end{aligned} \quad (4.16)$$

Thus, by differentiating (4.14) with respect to t , we obtain

$$\begin{aligned} \frac{dA_s(t; p, q)}{dt} &= -iS(t)\mathcal{L}(p, q, t)A_s(p, q) \\ &= -iS(t)\mathcal{L}(p, q, t)S(t)^{-1}S(t)A_s(p, q) \\ &= -i\mathcal{L}(p_t, q_t, t)A_s(t; p, q), \end{aligned}$$

which proves that (4.14) is the solution of (4.12) and at the same time that it is equal to (4.8).

If the density operator represents the canonical distribution with the parameter

$$\beta = 1/kT,$$

it satisfies the Bloch equation,

$$\frac{\partial}{\partial \beta} \rho = -\mathcal{H} \rho, \quad (4.17a)$$

or

$$\frac{\partial}{\partial \beta} \rho = -\rho \mathcal{H}, \quad (4.17b)$$

with the initial condition,

$$\rho_{\beta=0} = 1.$$

The Bloch equation, (4.17a) or (4.17b), is interpreted into the Wigner representation as

$$\frac{\partial f}{\partial \beta} = -\mathcal{H}_w f, \quad (4.18a)$$

or

$$\frac{\partial f}{\partial \beta} = -f \mathcal{H}_w. \quad (4.18b)$$

Therefore the Wigner d. f. for the canonical distribution satisfies

$$\begin{aligned} \frac{\partial f}{\partial \beta} &= -\frac{1}{2}(\mathcal{H}_w + \bar{\mathcal{H}}_w), \\ &\equiv -\frac{1}{2}\{\mathcal{H}(p_w, q_w) + \mathcal{H}(\bar{p}_w, \bar{q}_w)\}f. \end{aligned} \quad (4.19)$$

Thus, it can be written as

$$\begin{aligned} f(p, q; \beta) &= e^{-\beta \mathcal{H}_w} \cdot 1 \\ &= e^{-\beta \bar{\mathcal{H}}_w} \cdot 1 \\ &= e^{-\beta(\mathcal{H}_w + \bar{\mathcal{H}}_w)/2} \cdot 1. \end{aligned} \quad (4.20)$$

Any of these expressions can be used for computation.

Part II. Electrons in a Magnetic Field

§ 5. Commutation Rules and Wigner Representation

Let us now consider an electron in a magnetic field \vec{H} which is derived from a vector potential $\vec{A}(\vec{x})$. The Hamiltonian is then given by

$$\mathcal{H} = \frac{1}{2m} \left(\vec{p} - \frac{e}{c} \vec{A} \right)^2 + V(\vec{x}), \quad (5.1)$$

where \vec{p} is the canonical momentum. The physical momentum $\vec{\pi}$ may be defined by

$$\vec{\pi} \equiv m \frac{d\vec{x}}{dt} = \vec{p} - \frac{e}{c} \vec{A}, \quad (5.2)$$

which satisfy the commutation rules

$$[\pi_x, \pi_y] = -\frac{\hbar e}{ic} \left\{ \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right\} = -\frac{\hbar e H_z}{ic}, \quad (5.3)$$

and

$$[\pi_x, x] = \frac{\hbar}{i}, \text{ etc.} \quad (5.4)$$

We shall use in the following the variables, $\vec{\pi}$ and \vec{x} , instead of \vec{p} and \vec{x} . These are not canonical since they have the extra non-vanishing commutator (5.3), but they are independent of the gauge of the vector potential and have the advantage of being more physical than the canonical variables \vec{p} and \vec{x} . In terms of these variables, the Hamiltonian (5.1) keeps the form,

$$\mathcal{H} = \frac{1}{2m} \vec{\pi}^2 + V(\vec{x}), \quad (5.5)$$

the magnetic field entering only into the commutator (5.3).

Now, a Wigner distribution function, originally defined as a function of \vec{p} and \vec{x} , can be regarded as a function of $\vec{\pi}$ and \vec{x} simply by the change of independent variables defined by (5.2), which has the Jacobian equal to 1; i. e.

$$\frac{\partial(\pi_x, \pi_y, \pi_z, x, y, z)}{\partial(p_x, p_y, p_z, x, y, z)} = 1.$$

Then any Wigner operator, introduced for the p - x representation of Wigner distribution functions can be transformed into a differential operator operating on functions of $\vec{\pi}$ and \vec{x} . By changing the independent variables in this way, we have the transformation,

$$\frac{\partial}{\partial p_x} = \frac{\partial}{\partial \pi_x}, \text{ etc.}, \quad (5.6)$$

$$\begin{aligned} \left(\frac{\partial}{\partial x} \right)_p &= \left(\frac{\partial}{\partial x} \right)_\pi \\ &- \frac{e}{c} \left(\frac{\partial A_x}{\partial x} \frac{\partial}{\partial \pi_x} + \frac{\partial A_y}{\partial x} \frac{\partial}{\partial \pi_y} + \frac{\partial A_z}{\partial x} \frac{\partial}{\partial \pi_z} \right), \text{ etc.} \end{aligned}$$

Thus we have the following transformation of the Wigner operators:

$$x_w \equiv x - \frac{\hbar}{2i} \frac{\partial}{\partial p_x} = x - \frac{\hbar}{2i} \frac{\partial}{\partial \pi_x}, \text{ etc.} \quad (5.7)$$

$$\begin{aligned}
\pi_{zw} &\equiv p_z + \frac{\hbar}{2i} \frac{\partial}{\partial x} \\
&- \frac{e}{c} A_x \left(x - \frac{\hbar}{2i} \frac{\partial}{\partial p_x}, y - \frac{\hbar}{2i} \frac{\partial}{\partial p_y}, z - \frac{\hbar}{2i} \frac{\partial}{\partial p_z} \right) \\
&= \pi_x + \frac{\hbar}{2i} \frac{\partial}{\partial x} \\
&- \frac{\hbar e}{2ic} \left(\frac{\partial A_x}{\partial x} \frac{\partial}{\partial \pi_x} + \frac{\partial A_y}{\partial x} \frac{\partial}{\partial \pi_y} + \frac{\partial A_z}{\partial x} \frac{\partial}{\partial \pi_z} \right) \\
&+ \frac{e}{c} \left\{ A_x(x, y, z) \right. \\
&- \left. A_x \left(x - \frac{\hbar}{2i} \frac{\partial}{\partial \pi_y}, y - \frac{\hbar}{2i} \frac{\partial}{\partial \pi_x}, z - \frac{\hbar}{2i} \frac{\partial}{\partial \pi_z} \right) \right\}, \\
&\text{etc. (5.8)}
\end{aligned}$$

If the magnetic field is varying in space only slowly, the last term of (5.8) can be expanded to result in

$$\begin{aligned}
\pi_{zw} &= \pi_x + \frac{\hbar}{2i} \frac{\partial}{\partial x} \\
&- \frac{\hbar e}{2ic} \left(H_z \frac{\partial}{\partial \pi_y} - H_y \frac{\partial}{\partial \pi_z} \right), \text{ etc. (5.9)}
\end{aligned}$$

which is exact if \vec{H} is uniform or is a good approximation if

$$\hbar |\text{grad } H| / H \ll |\pi|, \quad (5.10)$$

for relevant values of the momentum, i. e., if the magnetic field hardly changes in the characteristic wave length of the electron.

Thus we can define a Wigner representation for the quantum mechanics of electrons in a magnetic field, which may be called the $(\vec{\pi}, \vec{x})$ Wigner representation. Unlike the Wigner representation discussed in Part I, this is characterized by the use of non-canonical variables. It is rather important from a general methodological point of view to recognize that noncanonical variables can be introduced in the Wigner representation in such a simple manner.

Let us assume a uniform, constant magnetic field \vec{H} , for which (5.9) holds exactly. The equation of motion of the density matrix will be written in the $(\vec{\pi}, \vec{x})$ Wigner representation as

$$\frac{\partial}{\partial t} f(\vec{\pi}, \vec{x}, t) = i \mathcal{L}_H f(\vec{\pi}, \vec{x}, t), \quad (5.11)$$

where

$$\begin{aligned}
i \mathcal{L}_H &= \frac{1}{i\hbar} \{ \mathcal{H}(\vec{\pi}_w, \vec{x}_w) - \mathcal{H}(\vec{\pi}_w, \vec{x}_w) \} \\
&= -\frac{1}{m} \vec{\pi} \frac{\partial}{\partial \vec{x}}
\end{aligned}$$

$$\begin{aligned}
&+ \frac{1}{i\hbar} \left\{ V \left(\vec{x} - \frac{\hbar}{2i} \frac{\partial}{\partial \vec{\pi}} \right) - V \left(\vec{x} + \frac{\hbar}{2i} \frac{\partial}{\partial \vec{\pi}} \right) \right\} \\
&+ \frac{e}{c} \vec{H} \cdot \vec{\pi} \times \frac{\partial}{\partial \vec{\pi}}. \quad (5.12)
\end{aligned}$$

The quantum mechanical Liouville operator, $i \mathcal{L}_H$, in this representation takes a particularly simple form. The first two terms of (5.12), i. e.,

$$i \mathcal{L} = -\frac{1}{m} \vec{\pi} \frac{\partial}{\partial \vec{x}} + \frac{1}{i\hbar} \{ V_w - \bar{V}_w \}, \quad (5.13)$$

is exactly the same as that in the absence of a magnetic field, while the last term

$$i \mathcal{L}' = \frac{e}{mc} \vec{H} \cdot \vec{\pi} \times \frac{\partial}{\partial \vec{\pi}} = \omega_c \left(\pi_x \frac{\partial}{\partial \pi_y} - \pi_y \frac{\partial}{\partial \pi_x} \right), \quad (5.14)$$

represents the Lorentz force. The last expression is for the case where \vec{H} lies in the z -direction, ω_c being the cyclotron frequency

$$\omega_c = eH/mc. \quad (5.15)$$

Note that (5.14) is very simple because of the assumption of a constant mass m in (5.1). This is one of the advantages of using a Wigner representation in a magnetic problem.

§ 6. Landau Diamagnetism

As an illustration of the application of the Wigner representation in magnetic problems let us here consider the diamagnetism of electrons. For simplicity, we ignore the electron interactions so that the problem is essentially a one-particle problem. First we treat non-degenerate electrons obeying the Boltzmann statistics. Degenerate cases can easily be obtained from the results of a non-degenerate case.

Therefore, we now calculate the density matrix

$$\rho = e^{-\beta \mathcal{H}}, \quad (6.1)$$

in the $(\vec{\pi}, \vec{x})$ Wigner representation. The Wigner d. f. is then given by

$$f(\vec{\pi}, \vec{x}) = e^{-\beta \mathcal{H}_w} \cdot 1, \quad (6.2)$$

where the Wigner operator H_w is given by

$$\begin{aligned}
\mathcal{H}_w &= \frac{1}{2m} \vec{\pi}_w^2 + V_w = \mathcal{H}_w^0 + \mathcal{H}_w' + \mathcal{H}_w'' \\
\mathcal{H}_w^0 &= \frac{1}{2m} \left\{ \left(\pi_x + \frac{\hbar}{2i} \frac{\partial}{\partial x} \right)^2 + \left(\pi_y + \frac{\hbar}{2i} \frac{\partial}{\partial y} \right)^2 \right. \\
&\quad \left. + \left(\pi_z + \frac{\hbar}{2i} \frac{\partial}{\partial z} \right)^2 \right\}
\end{aligned}$$