QUANTUM MECHANICS AND HILBERT SPACE

GEORGE W. MACKEY, Harvard University

1. Introduction. This expository article is written in the spirit of the various "What is a ———?" articles that have appeared in this MONTHLY from time to time. It is an attempt to give mathematicians who have not studied quantum mechanics some idea of what it is about and to give those who have studied it only from the point of view of the physicists some idea of how it may be formulated† in a mathematically precise fashion. We shall avoid the difficulties that still plague the advanced portions of the subject by dealing only with non-relativistic quantum mechanics.

2. The nature of quantum mechanics. The basic difference between classical mechanics and quantum mechanics is that the $2n$-tuple of real numbers which characterizes the state of a system in classical mechanics is replaced in quantum mechanics by a family of probability measures on the real line. The $2n$ first-order differential equations in $2n$-space which describe how the state of the classical system changes in time are replaced by a differential equation in an infinite-dimensional space which describes how these probability measures change with time. Under certain circumstances the probability measures are so highly concentrated (*i.e.*, assign a probability so close to one to such a short interval) that they behave like numbers. The resulting system of numbers then changes approximately according to the laws of classical mechanics. In this sense classical mechanics is a limiting case of quantum mechanics.

The passage from classical mechanics to the more exact quantum mechanics becomes necessary when one can no longer ignore the way in which different measurements interfere with one another—for example, when dealing with the submicroscopic world. Even in dealing with familiar large-scale phenomena this passage would become necessary if one were to make measurements to a much higher degree of accuracy than is now possible.

The reason that the above-mentioned interference forces a retreat from numbers to probability measures may be explained as follows. The act of measuring the value of an observable A may change the system in such a way that a measurement of B immediately following the measurement of A will not agree with a measurement of B immediately preceding the measurement of A. Moreover the change in the result of measuring B will be of an unpredictable nature. As analysis of the situation shows, this circumstance precludes the possibility of assigning any meaning to the statement "A and B have the values a and b at the time t." On the other hand it does not preclude the possibility of assigning

† While the formulation given here is based essentially on that to be found in von Neumann's classic book, Mathematische Grundlagen der Quantenmechanik, it differs from it in a number of respects. It is the writer's personal way of looking at the foundations of quantum mechanics and cannot be claimed to be the standard mathematical view if indeed there is such. For another approach the reader is referred to Chapter I of some recent University of Chicago notes by Irving Segal.

a meaning to the statement "At the time t measurements of A and B are statistically distributed with probability measures† α and β. This is because a statistical distribution is determined by making measurements on a sample containing a large number of replicas of the system under consideration and one may use distinct samples in determining α and β.

3. General mathematical formulation. Let $\mathfrak{O}$ denote the set of all "observables" associated with a given physical system. By a *state* of our system we shall mean a *possible* simultaneous set of statistical distributions for the "observables" in $\mathfrak{O}$. We denote the set of all states by $\mathfrak{S}$. It will not be necessary for us to analyze the notions of state and observable. From a purely mathematical point of view $\mathfrak{O}$ is an abstract set and each member of $\mathfrak{S}$ is a function defined on $\mathfrak{O}$ and having values in the set $\mathfrak{M}$ of all probability measures on the line. The system $\mathfrak{O}$, $\mathfrak{S}$ determines completely what we may call the statics of the system. To determine the dynamics we must specify how members of $\mathfrak{S}$ change with time; that is we must determine for each real t the one-to-one transformation L_t of $\mathfrak{S}$ onto $\mathfrak{S}$ such that $L_t(\phi)$ is the state of the system a time t after it was in the state ϕ. Of course the family L_t need not be given directly in integrated form, but may be determined only by giving some sort of differential equation in t.

It should be noted that our formulation is still broad enough to include classical mechanics. In that case $\mathfrak{O}$ is the set of all sufficiently regular real-valued functions defined on classical phase space $\mathcal{P}$ and the general member of $\mathfrak{S}$ is defined by a probability measure on $\mathcal{P}$. The function from $\mathfrak{O}$ to $\mathfrak{M}$ defined by the probability measure α in $\mathcal{P}$ takes the function f on $\mathcal{P}$ into the probability measure $E \to \alpha(f^{-1}(E))$. The L_t are obtained by integrating the classical equations of motion on $\mathcal{P}$ and using the resulting one parameter group of one-to-one transformations of $\mathcal{P}$ onto itself to induce a corresponding group of one-to-one transformations of $\mathfrak{S}$ onto itself. Actually, of course, classical mechanics is usually concerned only with the special states defined by those probability measures which are concentrated in a single point of $\mathcal{P}$. Indeed the motion of these determines the motion of all of the others and it is for this reason that classical mechanics‡ need not concern itself with measures as such.

Given two states of a physical system one can consider a third state in which the system is in the first state with probability p and in the second with probability $1-p$. This leads us to make the following assumption about $\mathfrak{S}$. If ϕ_1 and ϕ_2 are in $\mathfrak{S}$ and $0 \leqq p \leqq 1$ then $p\phi_1 + (1-p)\phi_2$ is also in $\mathfrak{S}$. A state ϕ which can be so constructed from two other states is said to be a *mixed state*. A state which cannot be so constructed is said to be a *pure state*. When $\mathfrak{O}$ and $\mathfrak{S}$ are as in the preceding paragraph the pure states are obviously just those defined by probability measures concentrated in single points of $\mathcal{P}$. Physically speaking, a pure

† See the foundations of probability, by P. R. Halmos, this MONTHLY vol. 51, 1944, pp. 493-510.

‡ The word "need" here must be interpreted in a suitably restricted sense. Of course, it may still be quite important to discuss the motion of measures which are not concentrated in points—for example, in classical statistical mechanics.

state is one in which one cannot concentrate the probability measure for any observable without spreading out the probability measure for another. The characteristic features of quantum mechanics arise from the fact that the pure states are not trivially pure as they are in classical mechanics—for each pure state there are many observables with highly unconcentrated probability measures.

If f is a real-valued function of a real variable and A is an observable, there is an obvious physical sense in which one can construct an observable $f(A)$. A measurement of $f(A)$ is simply a measurement of A, giving the value a say, followed by a computation of $f(a)$. This leads us to make the following further assumption about the mathematical system $\mathfrak{O}$, $\mathfrak{S}$. For each $A\in\mathfrak{O}$ and each real Borel function f on the real line there exists $B\in\mathfrak{O}$ such that for every Borel set E and each $\phi\in\mathfrak{S}$ $\phi(B)(E)=\phi(A)(f^{-1}(E))$. If we assume, as we shall, that no two distinct members of $\mathfrak{O}$ are carried into the same probability measure by all members of $\mathfrak{S}$, then B is uniquely determined by f and A. We denote it by $f(A)$.

4. The transition from observables to questions. Let us call an observable A a *question* if for every state ϕ the probability measure $\phi(A)$ is concentrated in the two points 0 and 1. It is not difficult to prove that A is a question if and only if $A^2=A$. If A is any observable and E is any Borel subset of the real line let ψ_E be the characteristic function of E; that is the function which is one for $x\in E$ and zero for x not in E. Then $\psi_E(A)$ is always a question which we shall denote by Q_E^A. The family of questions Q_E^A obtained by holding A fixed and letting E vary through the Borel subsets of the real line clearly determines the observable A. When Q is a question and ϕ is a state the probability measure $\phi(Q)$ is completely determined by its value at the set consisting of one alone. We shall call this value $m_\phi(Q)$. Since $m_\phi(Q_E^A)=m_\phi(\psi_E(A))=\phi(A)(\psi_E^{-1}(\{1\}))=\phi(A)(E)$ we see that the state ϕ is completely determined by the values of m_ϕ at the questions in $\mathfrak{O}$.

Let us denote the set of all questions in $\mathfrak{O}$ by q. We partially order q by setting $Q_1\leqq Q_2$ whenever $m_\phi(Q_1)\leqq m_\varphi(Q_2)$ for all states ϕ. If Q_1 and Q_2 are questions such that $Q_1\leqq 1-Q_2$, we say that Q_1 and Q_2 are mutually exclusive and write $Q_1\perp Q_2$. If $Q_1, Q_2, \cdots$ are questions such that $Q_i\perp Q_j$ for $i\neq j$ there clearly exists at most one question Q such that $m_\phi(Q)=m_\phi(Q_1)+m_\phi(Q_2)+\cdots$ for all $\phi\in\mathfrak{S}$. We assume that Q always exists and denote it by $Q_1+Q_2+\cdots$. It is not difficult to show† that Q is the least question R such that $R\geqq Q_j$ for all j.

If A is an observable then the mapping $E\to Q_E^A$ clearly has the following properties: (1) $E\cap F=0$ implies $Q_E^A\perp Q_F^A$, (2) $Q_{\{0\}}^A=0$ and $Q_{[-\infty,\infty]}^A=1$,‡ (3) $E=E_1\cup E_2\cup\cdots$ where $E_i\cap E_j=0$ for $i\neq j$ implies that $Q_E^A=Q_{E_1}^A+Q_{E_2}^A+\cdots$. If $E\to Q_E$ is any mapping of Borel sets in the line into questions which has the three properties just listed we shall call it a *question-valued measure*. Let $\bar{\mathfrak{O}}$ denote the set of all question-valued measures. For each $\phi\in\mathfrak{S}$ and $Q\in\bar{\mathfrak{O}}$ let

† We are indebted to R. V. Kadison for this remark. In our original manuscript the fact in question appears as an added assumption.

‡ 1 is the unit constant observable, that is the observable A such that $\phi(A)$ is concentrated in the point 1 for all states ϕ.

$\phi(Q)(E)=m_\phi(Q_E)$. If we identify the members A of $\mathfrak{O}$ with their associated question-valued measures Q^A so that we may consider $\mathfrak{O}$ as a subset of $\bar{\mathfrak{O}}$, then the preceding definition amounts to an extension of each ϕ from $\mathfrak{O}$ to $\bar{\mathfrak{O}}$. The system $\bar{\mathfrak{O}}$, $\mathfrak{S}$ then satisfies all axioms already laid down for $\mathfrak{O}$, $\mathfrak{S}$, and in addition the following: Every question-valued measure is the question-valued measure associated with some observable. We shall assume that we have added the observables in $\bar{\mathfrak{O}}-\mathfrak{O}$ to our system, that is, that our system $\mathfrak{O}$, $\mathfrak{S}$ satisfies the axiom just enunciated. Thus $\mathfrak{O}$ and $\mathfrak{S}$ can be reconstructed as soon as we are given q and the m_ϕ. It is clear that we could have taken q and the m_ϕ as our starting-point instead of $\mathfrak{O}$ and $\mathfrak{S}$.

The m_ϕ clearly have the following properties: (1) If $Q_1, Q_2, \cdots$ are questions such that $Q_i \perp Q_j$ for $i \neq j$ then $m_\phi(Q_1+Q_2+\cdots)=m_\phi(Q_1)+m_\phi(Q_2)+\cdots$, (2) $m_\phi(1)=1$. We shall call any function from the questions to the positive real numbers which has these two properties a *measure on the questions*. It is tempting to add the assumption that every measure on the questions is the m_ϕ for some state ϕ. Like the corresponding assumption for question-valued measures it can be attained by enlarging the system. However, the physical significance of states and observables differ in such a way that one is much more willing to make the first enlargement than the second. There is a sense in which one can construct observables from questions, but we have to be content with the states provided by nature. If we did add this assumption (and for all one knows to the contrary it holds for the system $\mathfrak{O}$, $\mathfrak{S}$ actually in use for quantum mechanics) it would follow that the pair $\mathfrak{O}$, $\mathfrak{S}$ is completely determined by the partially ordered set q of all questions and the "complementation operation" $Q \to 1-Q$. In any event it is natural to call this complemented partially ordered set the *logic*† of our system. The statics of the system is determined by its logic and a certain convex set of measures on the questions.

5. The logic of quantum mechanics. We have not yet added enough axioms to exclude classical mechanics. If we interpret "sufficiently regular" in our description of the $\mathfrak{O}$, $\mathfrak{S}$ for classical mechanics to mean being a Borel function, then the questions are in a natural one-to-one correspondence with the Borel subsets of phase space. The logic of classical mechanics is a Boolean algebra—the Boolean algebra of all Borel subsets of phase space. We see in particular that in the case of classical mechanics our questionable axiom is indeed satisfied. Every measure on the questions corresponds to a state. It is interesting to note that any two classical mechanical systems have isomorphic logics. This follows from a known set-theoretical result according to which there exists a one-to-one Borel set-preserving map between any two separable complete metric spaces.

We single out the system of quantum mechanics not by adding more general axioms but by stating explicitly what its logic is and which measures on the

† *Cf.* The logic of quantum mechanics, by G. Birkhoff and J. von Neumann, Ann. Math., vol. 37, 1936, pp. 823–843.

questions are associated with states. As in the case of classical mechanics, all quantum mechanical systems have (to within isomorphism) the same logic. This logic is the partially ordered set of all closed subspaces of a separable infinite dimensional Hilbert space (see the next section for a definition of Hilbert space), complementation being passage from a subspace to its orthogonal complement.

It would, of course, be very interesting if this structure for q and $Q \to 1-Q$ could be deduced from a set of physically meaningful and plausible axioms. Such a deduction is not at present available. It has simply been found that physical systems behave as if q had the indicated structure. For us it is an axiom. To mathematicians familiar with lattice theory this axiom is not so artificial and arbitrary as it may seem at first sight to others. The lattice of all subspaces of a vector space is one of the principal examples of a partially ordered set which is not a Boolean algebra but nevertheless has many regularity properties. Moreover, one can show that if the lattice of all closed subspaces of a Banach space has an order inverting involution with the properties of $Q \to 1-Q$, then this Banach space has an equivalent norm under which it is a Hilbert space. We shall postpone our description of $\mathcal{S}$ until after the exposition of basic Hilbert space theory to which we propose to devote the next two sections.

6. Definition and elementary properties of Hilbert space. A Hilbert space $\mathfrak{H}$ is a vector space over the complex numbers in which there is given a complex-valued function of two variables (ϕ, ψ) such that: (1) For fixed ψ, (ϕ, ψ) is a linear function of ϕ, (2) $(\phi, \psi) = (\overline{\psi, \phi})$, (3) $(\phi, \phi) > 0$ unless $\phi = 0$. (4) Under $\rho(\phi, \psi) = \sqrt{(\phi-\psi, \phi-\psi)}$, $\mathfrak{H}$ is a complete metric space. The complex number (ϕ, ψ) is called the inner product of ϕ and ψ, and $\sqrt{(\phi, \phi)}$ is called the norm of ϕ and written $\|\phi\|$.

The two most familiar examples of Hilbert spaces are the following: (I) $\mathfrak{H}$ is the set of all sequences $c_1, c_2, \cdots$ of complex numbers such that $|c_1|^2 + |c_2|^2 + \cdots < \infty$, and the inner product of $c_1, c_2, \cdots$ with $d_1, d_2, \cdots$ is $c_1\bar{d}_1 + c_2\bar{d}_2 + \cdots$, (II) $\mathfrak{H}$ is the set of all complex-valued Lebesgue measurable square-integrable functions on the interval $[a, b]$ $(-\infty \leqq a < b \leqq \infty)$, functions equal almost everywhere being identified, with $(f, g) = \int_a^b f(x)\bar{g}(x)dx$. On modifying (II) by passing from $[a, b]$ to a general measure space one gets a class of examples including both (I) and (II). It is important to note that except for differences in "dimension" all Hilbert spaces are abstractly equivalent—the dimension of a Hilbert space being defined as the maximum number of elements in a set $\{\phi_\alpha\}$ such that $(\phi_\alpha, \phi_\beta) = 0$ for $\alpha \neq \beta$ and $(\phi_\alpha, \phi_\alpha) = 1$. Actually any two such sets which are maximal in the sense that they cannot be enlarged and which lie in the same Hilbert space have the same cardinal number. Such a maximal set is said to be an orthonormal basis for the space. If $\mathfrak{H}$ has dimension $\aleph_0$, as do examples (I) and (II) above, and $\phi_1, \phi_2, \cdots$ is an orthonormal basis, then every ϕ in $\mathfrak{H}$ is uniquely representable in the form $c_1\phi_1 + c_2\phi_2 + \cdots$, where $|c_1|^2 + |c_2|^2 + \cdots < \infty$. The mapping $\phi \to \{c_1, c_2 \cdots\}$ is then an inner product preserving linear map of $\mathfrak{H}$ on example (I). Practically the same considerations lead to a correspond-

ing proof for higher dimensional spaces. A separable Hilbert space always has dimension $\aleph_0$ or less.

A closed subspace of a Hilbert space is by definition a linear subspace which is closed in the metric. Every closed subspace is a Hilbert space in its own right. If M is a closed subspace then $M^\perp$, the orthogonal complement of M, is by definition the set of all ψ such that $(\phi, \psi)=0$ for all $\phi \in M$. It can be shown that $M^{\perp\perp}=M$ for all M, and that every $\phi \in \mathfrak{H}$ is uniquely of the form $\phi_1+\phi_2$, where $\phi_1 \in M$ and $\phi_2 \in M^\perp$.

7. Linear operators in Hilbert space. Let $\mathfrak{H}$ be a Hilbert space. Let T be a linear function mapping $\mathfrak{H}$ into $\mathfrak{H}$. T is continuous if and only if it is bounded, in the sense that $\|T(\phi)\|/\|\phi\|$ is bounded for $\phi \neq 0$. The least upper bound of these numbers $\|T(\phi)\|/\|\phi\|$ is called the norm $\|T\|$ of T. For each bounded linear operator T there is a unique bounded linear operator T^* such that $(T(\phi), \psi) = (\phi, T^*(\psi))$ for all ϕ and ψ in $\mathfrak{H}$. T^* is called the adjoint of T. If $T=T^*$, T is said to be *self-adjoint*. If $T^*T=TT^*=I$ (where I is the identity operator) then T is said to be *unitary*. T is unitary if and only if it is one-to-one, has all of $\mathfrak{H}$ for its range and preserves the inner product. The unitary operators are thus the automorphisms of $\mathfrak{H}$. Let M be a closed subspace of $\mathfrak{H}$. For each $\phi \in \mathfrak{H}$ let $P_M(\phi)$ denote the unique $\psi \in \mathfrak{H}$ such that $\phi-\psi$ is in $M^\perp$; that is, let P_M be the "projection" of ϕ on M. Then P_M is a bounded self-adjoint linear operator and $P_M^2=P_M$. Conversely if T is any self-adjoint bounded linear operator such that $T^2=T$, it is easy to show that $T=P_M$ where M is the set of all ϕ such that $T(\phi)=\phi$. For this reason the bounded self-adjoint linear operators T such that $T^2=T$ are called *projections*. The mapping $M \to P_M$ sets up a one-to-one correspondence between the closed subspaces and the projections such that for all M, $P_{M^\perp}=1-P_M$. Let T be any bounded self-adjoint linear operator. Then $(T(\phi), \psi)$ as a function of ϕ and ψ has properties (1) and (2) listed in the definition of an inner product. Moreover it also has the property: $|(T(\phi), \psi)| \leqq \|T\| \cdot \|\phi\| \, \|\psi\|$. We shall call any function $[\phi, \psi]$ which has properties (1) and (2), and is such that $|[\phi, \psi]| \leqq K\|\phi\| \, \|\psi\|$ for some positive real number K, a bounded Hermitian bilinear form. It is a theorem that every bounded Hermitian bilinear form is of the form $(T(\phi), \psi)$ for a uniquely determined self-adjoint bounded linear operator T.

8. Quantum statics. Let $\mathfrak{H}$ be the Hilbert space whose lattice of closed linear subspaces is the logic of our quantum-mechanical system $\mathfrak{O}$, $\mathfrak{S}$ so that we may identify questions in $\mathfrak{O}$ with closed subspaces of $\mathfrak{H}$. Let ϕ be any element of $\mathfrak{H}$ such that $\|\phi\|=1$. Then $M \to (P_M(\phi), \phi)$ is readily verified to be a measure on the questions. More generally if $\phi_1, \phi_2, \cdots$ is any sequence of members of $\mathfrak{H}$ all having norm one and $\gamma_1, \gamma_2, \cdots$ is any sequence of positive real numbers such that $\gamma_1+\gamma_2+\cdots=1$ then $\gamma_1(P_M(\phi_1)\cdot\phi_1)+\gamma_2(P_M(\phi_2), \phi_2)+\cdots$ converges for all M and as a function of M is a measure on the questions. We complete our axioms concerning quantum statics by assuming that the measures on the ques-

tions which define states are just those of the form† $M \rightarrow \gamma_1(P_M(\phi_1), \phi_1) + \gamma_2(P_M(\phi_2), \phi_2) + \cdots$. It is easy to show that a state is pure if and only if it is defined by a single element of $\mathfrak{K}$; that is if and only if the associated measure on the questions is of the form $M \rightarrow (P_M(\phi), \phi)$. It is also easy to show that ϕ_1 and ϕ_2 define the same pure state if and only if $\phi_1 = e^{i\theta}\phi_2$ for some real θ.

If A is any observable and $E \rightarrow Q_E^A$ is the corresponding question-valued measure we may construct what we shall call a "projection-valued measure" by replacing each closed subspace Q_E^A by the projection upon it. We leave the formulation of the definition of projection-valued measure to the reader. We will denote the projection-valued measure thus associated with the observable A by P^A and the projection which it assigns to the Borel set E by P_E^A. Obviously the mapping $A \rightarrow P^A$ sets up a one-to-one correspondence between all observables on the one hand and all projection-valued measures on the other. The probability distribution of the observable A in the pure state defined by the Hilbert-space element ϕ is of course just $E \rightarrow (P_E^A(\phi), \phi)$—a formula which sums up the whole of quantum statics. In the sequel we shall speak only of pure states and we shall find it convenient to identify them with the Hilbert-space elements which define them.

9. Observables and operators. Let A be an observable which is *bounded* in the sense that $P_E = I$ for some bounded subset E of the real line. Then for all ϕ and ψ in $\mathfrak{K}$, $E \rightarrow (P_E^A(\phi), \psi)$ is a countably additive complex-valued set function $\alpha_{\phi,\psi}^A$ with respect to which the Stieltjes integral $\int_{-\infty}^{\infty} x d\alpha_{\phi,\psi}^A(x)$ exists. The resulting function of ϕ and ψ is easily seen to be a bounded Hermitian bilinear form. Hence there exists a bounded self-adjoint linear operator T^A uniquely determined by A such that for all ϕ and ψ in $\mathfrak{K}$, $(T^A(\phi), \psi) = \int_{-\infty}^{\infty} x d\alpha_{\phi,\psi}^A(x)$. Thus every bounded observable A is associated with a unique bounded self-adjoint linear operator. It follows from the celebrated spectral theorem that the converse is true. Every bounded self-adjoint linear operator is derivable in the manner just described from a uniquely determined bounded projection-valued measure, and hence is of the form T^A for a uniquely determined bounded observable A. $A \rightarrow T^A$ thus sets up a one-to-one correspondence between the set of all bounded observables in $\mathfrak{O}$ and the set of all bounded self-adjoint linear operators. This correspondence may be extended to one between all observables in $\mathfrak{O}$ and all self-adjoint linear operators by suitably defining the notion of unbounded self-adjoint linear operator. Let T be linear but let its domain be (perhaps) only a dense linear subspace of $\mathfrak{K}$, and let us make no assumptions about its boundedness or continuity. The adjoint of T is then defined as follows: ψ is in the domain

† As indicated earlier, no other measures on the questions are known at present. Indeed, since these words were written, A. M. Gleason has made considerable progress toward showing that there are none. Using his results one can replace our assumption about the specific form of the allowable measures by one which asserts, simply, that these measures are not too badly "non-measurable."

of T^* if and only if $(T(\phi), \psi)$ is continuous as a function of ϕ for all ϕ in the domain of T. $T^*(\psi)$ is then the unique element of $\mathfrak{H}$ such that $(T(\phi), \psi) = (\phi, T^*(\psi))$ for all ϕ in the domain of T. T is said to be self-adjoint whenever $T = T^*$ (equality including having the same domain). When T is bounded and everywhere defined this definition coincides with the one already given. The construction of an unbounded self-adjoint operator from an unbounded projection-valued measure is analogous to the corresponding construction in the bounded case but is complicated by the fact that the relevant bilinear form is not everywhere defined. We shall not give details here. We exploit the one-to-one correspondence between observables and self-adjoint linear operators which we have just described by identifying every observable with its corresponding operator.

Let A be a self-adjoint linear operator in $\mathfrak{H}$ and let P^A be the corresponding projection valued measure. Let G denote the union of all open sets E for which $P^A_E = 0$. Then G itself is an open set and $P^A_G = 0$. The closed complement of the open set G is called the *spectrum* S_A of A. The significance of S_A for A regarded as an observable is obviously the following. In every state the probability measure of A is concentrated in S_A and for every open interval F on the real line which intersects S_A there exists a state in which the probability measure of A is concentrated in F. In this sense S_A is the set of possible values for the observable A. Every closed subset of the real line is the spectrum of some self-adjoint linear operator and hence the set of possible values of some observable.

The set of all $\lambda \in S_A$ such that $P^A_{\{\lambda\}} \neq 0$, where $\{\lambda\}$ denotes the set whose only element is λ, is called the *point spectrum* of A. If ϕ is an element of norm one in the range of $P^A_{\{\lambda\}}$ then it is a state in which the observable A has the value λ with probability one. Conversely if ϕ is a pure state in which A has the value λ with probability one then ϕ is in the range of $P^A_{\{\lambda\}}$. The range of $P^A_{\{\lambda\}}$ coincides with the set of all ϕ for which $A(\phi) = \lambda\phi$; that is with the set of all *proper vectors* of A belonging to the *proper value* λ. The point spectrum of A is always at most denumerable. A is said to have a *pure point spectrum* if $P^A_D = I$ where D is the point spectrum; that is if in every state the probability measure of A is concentrated in D. A has a pure point spectrum if and only if there exists a basis of proper vectors. One of the more spectacular ways in which quantum mechanics differs from classical mechanics is in the occurrence in quantum mechanics of observables with a sizable point spectrum whose classical analogues are continuous functions on phase space.

The identification of observables with self-adjoint operators together with the operation $a \to f(A)$ already defined for observables leads at once to a definition of $f(A)$ whenever f is a Borel function and A is a self-adjoint operator. This definition can of course be formulated without reference to quantum mechanics. When f is a polynomial it agrees with the usual algebraic definition of a polynomial function of an operator. If A_1 and A_2 are bounded self-adjoint linear operators it can be shown that $A_1A_2 = A_2A_1$ if and only if $P^{A_1}_E P^{A_2}_F = P^{A_2}_F P^{A_1}_E$ for all E and F. When A_1 and A_2 are unbounded and consequently not everywhere defined this condition on the projection-valued measures is taken as the defini-

tion of the commutativity of the operators. With this definition it can be shown that A_1 and A_2 commute if and only if there exists a third self-adjoint operator A_3 and Borel functions f_1 and f_2 such that $A_1=f_1(A_3)$ and $A_2=f_2(A_3)$. There is thus an obvious sense in which observables which commute when considered as operators are "simultaneously measurable". The circumstance which prevents our assigning an obvious physical sense to $f(A_1, A_2)$, where f is a Borel function of two variables and A_1 and A_2 are observables, is the possible lack of "simultaneous measurability" of A_1 and A_2. It can be shown that whenever A_1 and A_2 are self-adjoint linear operators for which there is a mapping $f\rightarrow f(A_1, A_2)$ having reasonable properties, then A_1 and A_2 commute. In this sense commutativity of A_1 and A_2 considered as operators is equivalent to simultaneous observability of A_1 and A_2 considered as observables.

10. Quantum dynamics. We have yet to study the nature of the one-parameter group of transformations $t\rightarrow L_t$ of $\mathcal{S}$ onto $\mathcal{S}$ which describes the way in which our system $\mathcal{O}$, $\mathcal{S}$ changes with time. Our basic dynamical assumption is that each L_t is defined by a unitary transformation U_t of $\mathcal{H}$ onto $\mathcal{H}$, and that L_t depends upon t in such a manner that $|(U_t(\phi), \psi)|$ is continuous in t for all ϕ and ψ in $\mathcal{H}$. Each U_t is determined by L_t up to multiplication by a complex number of modulus one. It turns out that one can choose the arbitrary constants so that $U_{t_1+t_2}=U_{t_1}U_{t_2}$ for all t_1 and t_2, and so that $(U_t(\phi), \psi)$ is itself continuous. Moreover, the fact that L_t is of the given form may be deduced from quite general hypotheses about the action of the L_t on $\mathcal{S}$. The continuous one-parameter group $t\rightarrow U_t$ is, of course, not quite uniquely determined by the L_t. If c is any real number then $t\rightarrow U_t$ and $t\rightarrow e^{-itc}U_t$ come from the same L_t's and define the same dynamics. This however is the extent of the ambiguity.

Let H be any self-adjoint linear operator. We define e^{-itH} for all real t as $\cos(tH)-i\sin(tH)$. Since $\sin x$ and $\cos x$ are bounded functions, e^{-itH} is a bounded linear operator even when H is not bounded. As a matter of fact, it is unitary, and $t\rightarrow e^{-itH}$ is a continuous one-parameter unitary group of operators. According to a fundamental theorem of M. H. Stone, every continuous one-parameter group of unitary operators is of this form, with a uniquely determined self-adjoint H. Thus $H\rightarrow U_t$, where $U_t=e^{-itH}$, sets up a one-to-one correspondence between the self-adjoint operators and the continuous one-parameter unitary groups. If $t\rightarrow U_t$ determines the dynamics of our system and $U_t=e^{-itH}$, then the self-adjoint operator H also determines the dynamics of our system. Since $e^{-itc}e^{-itH}=e^{-it(c+H)}$, the ambiguity in U produces a certain ambiguity in H as well. H is determined only up to an additive constant operator by the L_t. We shall ignore this ambiguity and suppose a particular H chosen once and for all. We shall call H the *dynamical operator* of the system.

If ϕ is in the domain of H and is of norm one then the variable pure state $e^{-itH}(\phi)=\phi_t$ satisfies the differential equation: $d\phi_t/dt=-iH(\phi_t)$. This is the abstract form of what is called Schrödinger's equation, and plays the role in quantum mechanics played by Hamilton's equations in classical mechanics. It

is a first-order differential equation whose solutions are the trajectories of the pure states.

Let ϕ be a proper vector of H having norm one and corresponding proper value λ. Then $e^{-itH}(\phi) = e^{-it\lambda}\phi$, and thus is the same state for all time t. In other words, the state defined by ϕ is constant in time. It is what is called a *stationary state*. Conversely every stationary state ϕ is a proper vector of H. The existence of pure states which are stationary is a feature of quantum mechanics which has no counterpart in classical mechanics. Of course, a given H need not have any proper vectors, but those which arise most commonly in practice have a good many.

In classical mechanics certain functions on phase space are constant on the trajectories and are called integrals. It is natural to call a quantum mechanical observable an *integral* if its probability-measure in every state ϕ is constant in time. An easy calculation shows that the observable A is an integral if and only if it commutes with H when regarded as an operator. Now H itself is a self-adjoint operator and as such is an observable; moreover this observable is an integral. It clearly plays a central role in the theory. As we shall see, it is a constant multiple of an observable which is the quantum-mechanical analogue of the energy integral of classical mechanics. Thus the stationary states of a quantum-mechanical system are just those in which the energy has a definite value with probability one and each stationary state is associated with a definite value of the energy.

11. The quantum mechanics of n "*interacting particles*". The discussion so far has been rather general and abstract. Except in the case of the energy we have given no indications whatever concerning the physical meaning of the observable identified with a particular operator. Moreover, we have not explained how one constructs the quantum-mechanical refinement of a single classical system. We now remedy these deficiencies by exhibiting and discussing a quantum-mechanical system which in the limiting case of highly concentrated probabilities behaves like a system of n classical mass particles moving under mutual central forces.

Consider the Hilbert space of all square-summable functions on Euclidean $3n$ space R^{3n}. Let D denote the formal differential operator

$$f \to \sum_{j=1}^{n} -\frac{1}{\mu_j}\left(\frac{\partial^2 f}{\partial x_j^2} + \frac{\partial^2 f}{\partial y_j^2} + \frac{\partial^2 f}{\partial z_j^2}\right) + Wf,$$

where W is a real function on R^{3n} and the μ_j are positive real numbers. For a wide class of choices of W there is a unique self-adjoint operator H_D whose restriction to the twice differentiable functions coincides with D. We consider the quantum-mechanical system whose Hilbert space is $\mathfrak{L}^2(R^{3n})$ and whose dynamical operator is H_D. For each real-valued Borel function g on R^{3n} let A_g be the self-adjoint operator $f \to fg$—the domain when g is unbounded being defined in the obvious way. The projection-valued measure associated with A_g can be shown to

take the set E into the projection A_ψ where ψ is the characteristic function of $g^{-1}(E)$. Thus the probability measure for A_g in the state f is

$$E \to \int_{g^{-1}(E)} \cdots \int |f|^2 dx_1 \cdots dz_n.$$

Let f be a state such that the mean value $\overline{A}_{x_j} = \int \cdots \int x_j |f|^2 dx_1 \cdots dz_n$ of A_{x_j} exists. A simple calculation shows that

$$\frac{d^2}{dt^2}(\overline{A}_{x_j}) = -\frac{2}{\mu_j}\int \cdots \int |f|^2 \left(\frac{\partial W}{\partial x_j}\right) dx_1 \cdots dz_n = -\frac{2}{\mu_j}\overline{A}_{\frac{\partial W}{\partial x_j}}$$

whenever the relevant derivatives exist. There are of course similar results for the rates of change of the mean values of the A_{y_j} and the A_{z_j}. When f is zero outside of a small neighborhood of a point in R^{3n} the mean values in these formulas are approximately the coordinates of this point and (assuming W to have continuous partial derivatives) the values of the $\partial W/\partial x_j$, $\partial W/\partial y_j$, $\partial W/\partial z_j$ at this point. In other words, when f varies in such a manner that the probability measures of the A_{x_j}, A_{y_j} and A_{z_j} in that state remain highly concentrated, so that the set of them may be represented approximately by a point in R^{3n}, then this point moves so as to satisfy the differential equations

$$\frac{\mu_j}{2}\frac{d^2x_j}{dt^2} = -\frac{\partial W}{\partial x_j}, \qquad \frac{\mu_j}{2}\frac{d^2y_j}{dt^2} = -\frac{\partial W}{\partial y_j}, \qquad \frac{\mu_j}{2}\frac{d^2z_j}{dt^2} = -\frac{\partial W}{\partial z_j}.$$

These are simply the classical equations of motion for a system of n mass-particles of masses $\frac{1}{2}\mu_j$ and potential energy function W. They are also the classical equations of motion for such a system where the masses are $\frac{1}{2}c\mu_j$, the potential energy function is cW, and c is any positive real number. In classical mechanics masses are determined by the equations of motion only up to a multiplicative constant which is then fixed by arbitrarily choosing a unit of mass. If quantum mechanics had been thought of at the outset one would presumably have made use of the natural unit of mass provided by Schrödinger's equation and defined $\frac{1}{2}\mu_j$ as the mass of the j-th particle. As it is, the mass m_j of the j-th particle is some constant multiple of $\frac{1}{2}\mu_j$. This constant is the same for all systems (for any given units of mass, length and time) and is generally denoted by $\hbar$. $2\pi\hbar$ is the fundamental constant introduced into the old quantum theory by Max Planck and called Planck's constant. Planck's constant is usually denoted by h. In terms of the m_j, the classical potential energy $V = W/\hbar$, and $\hbar$, Schrödinger's equation for our system takes the standard form:

$$-\frac{1}{i}\frac{\partial f}{\partial t} = -\frac{\hbar}{2}\sum_{j=1}^{n}\frac{1}{m_j}\left(\frac{\partial^2 f}{\partial x_j^2} + \frac{\partial^2 f}{\partial y_j^2} + \frac{\partial^2 f}{\partial z_j^2}\right) + \frac{Vf}{\hbar}.$$

Henceforth, we shall suppose V to be the potential energy arising from mutual central forces between the classical particles so that, in particular, there are linear and angular momentum integrals for the limiting classical system.

In $g \to A_g$ we have a natural one-to-one correspondence between those observables in classical mechanics which depend upon the coordinates, but not the velocities, and certain quantum-mechanical observables. It turns out *not* to be possible to extend this correspondence to one between all classical-mechanical observables and all quantum-mechanical observables. Indeed since classical mechanics is only a limiting case of quantum mechanics there is no reason to expect such a correspondence. Two different quantum-mechanical observables can well coincide in the classical limit. On the other hand there are, as it turns out, well defined quantum analogues for certain of the more important velocity dependent classical-mechanical observables. For example, the formal differential operator $(\hbar/im_j)\partial/\partial x_j$ defines† an observable whose mean value coincides with the time derivative of the mean value of A_{x_j} in every state for which these things have a meaning. In this sense it is the analogue of the classical-mechanical observable known as the x component of the velocity of the j-th particle. To remind ourselves of the analogy we give it the same name in quantum mechanics and do likewise for $(\hbar/im_j)\partial/\partial y_j$ and $(\hbar/im_j)\partial/\partial z_j$. The analogy may be pursued further. The observable corresponding to

$$\sum_{j=1}^{n} m_j\left(\frac{1}{i}\,\frac{\hbar}{m_j}\,\frac{\partial}{\partial x_j}\right) = \frac{\hbar}{i}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}$$

is an integral just as is the corresponding sum in classical mechanics. For obvious reasons it is called the x component of the total momentum and the observable corresponding to the term $(\hbar/i)\partial/\partial x_j$ is called the x component of the momentum of the j-th particle. The y and z components of momentum are defined analogously. In a similar manner one is led to define angular momentum observables and to the discovery that the total angular momentum about any axis is an integral. The angular momentum observables are interesting in that they all have pure point spectra; the spectrum in each case being the set of all integral multiples of h. Finally, if one takes the classical energy integral $\sum_{j=1}^{n} \frac{1}{2}m_j(\dot{x}_j^2 + \dot{y}_j^2 + \dot{z}_j^2) + V$ and tries to construct a corresponding quantum-mechanical observable by substituting the analogues already discovered for the $\dot{x}_j$, $\dot{y}_j$, $\dot{z}_j$ and V, one is led to the formal differential operator

$$f \to -\frac{\hbar^2}{2}\sum_{j=1}^{n}\frac{1}{m_j}\left(\frac{\partial^2 f}{\partial x_j^2} + \frac{\partial^2 f}{\partial y_j^2} + \frac{\partial^2 f}{\partial z_j^2}\right) + Vf$$

which is just $\hbar$ times the differential operator D with which we started. For this reason the observable $\hbar H_D$ is called the energy observable. As we have already remarked it is obviously an integral. The fact that one has natural analogues in quantum mechanics for the energy and momentum observables

† The correspondence between formal differential operators and self adjoint operators on Hilbert space is rather complicated in general. However, for the members of a wide class of first-order operators, including those with which we shall deal, there is a canonical way of passing to a corresponding self-adjoint operator.

lies deeper than the above formal considerations indicate. It can be traced back to the relationship of the observables in question to certain one parameter groups of automorphisms of the system.

12. The quantum mechanics of the atom. The quantum-mechanical model of an atom is not precisely what one would expect from the considerations of the preceding paragraph and the classical picture of an atom as a nucleus surrounded by n electrons. In order to produce results in reasonable agreement with experiment the model resulting from this picture must be modified in three respects. (We simplify the discussion by replacing the (relatively heavy) nucleus by a fixed force field so that an n electron atom leads to an n-body problem.) First $\mathfrak{L}^2(R^{3n})$ must be replaced by the Hilbert space of all square summable functions from R^{3n} to a 2^n-dimensional Hilbert space. Second this new Hilbert space must be replaced by the subspace consisting of those functions f such that $f(T_{ij}(p)) = -f(p)$ for all p in R^{3n} and all i and j with $i \neq j$. Here T_{ij} is the map of R^{3n} onto itself which takes each point into the point obtained by interchanging the coordinates of i-th and j-th particles. Thirdly the restriction of H_D to this subspace must be changed by adding a "small perturbation" which we shall not describe explicitly. The first change has the effect of adding certain new observables—the so-called spin angular moments. The second change eliminates a good many observables—all of those in which the n electrons are not treated symmetrically. The fact that there are observables corresponding to $x_1+x_2+\cdots+x_n$, $x_1^2+x_2^2+\cdots+x_n^2$ *etc.*, but none corresponding to x_1, $x_2, \cdots, x_n$ individually, is interpreted as indicating a new sense in which an electron is not a classical particle. An electron is supposed to have no individuality. It is supposed to be meaningless to make a distinction between a pair of electrons and the same pair with the electrons interchanged. The situation is analogous to that of a pair of identical kinks traveling along in a stretched string. It is possible to have motion in which the kinks seem to have changed places although in fact the state of the string is just as before.

With the indicated modifications, the quantum-mechanical refinement of the classical picture of the atom leads to a model which explains atomic phenomena, including chemistry, in a truly remarkable manner.†

† *Added in proof.* Since this article was sent to the printer, Professor Gleason has improved the result referred to in the footnote on page 51. A proof that there are no measures on the questions other than those associated with states will appear in the Journal of Mathematics and Mechanics.