A New Form for the Statistical Postulate of Quantum Mechanics

NORBERT WIENER AND ARMAND SIEGEL
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received May 15, 1953)

A new type of representation of the wave function of quantum mechanics, and a new form of the statistical interpretation of the wave function, are presented. The variable in the new representation is a mapping of points of differential space, closely related to Hilbert space, but having an associated measure for suitably defined sets of points in it. In the new statistical interpretation, an eigenvalue of every observable defined in the system is associated with each point in differential space. This can be done in such a way that, for each observable, the measure (=probability) of the set of all differential-space points belonging to a given set of eigenvalues is equal to the quantum-mechanical probability, calculated in the usual way (Born statistical postulate), that an experiment will yield an eigenvalue in this set. Thus we obtain an interpretation of quantum mechanics in terms of probability densities or probabilities instead of probability amplitudes.

I. INTRODUCTION

THE statistical interpretation of the wave function, introduced by Born,[1] consists of the following postulate: If $\varphi(q)$ is the wave function in a representation in which any operator $\mathbf{q}$ (having eigenvalues q) is diagonal, the probability that the corresponding observable q will be found on measurement to lie in the range q_1 to q_2 is

$$\int_{q_1}^{q_2} |\varphi(q)|^2 dq, \tag{1}$$

if $\mathbf{q}$ has a continuous eigenvalue spectrum with an analogous expression for discrete spectra. Since the entity in terms of which the dynamics of the system is expressed, through the time-dependent Schrödinger equation, is the wave function $\varphi(q)$ and not the probability density $|\varphi(q)|^2$, actual physical probabilities—i.e., frequencies of occurrence of given events—play a noncentral role in the conventional quantum theory.[2]

In this paper we propose an explicit statistical postulate distinct from that of Born, which is equivalent to the Born postulate so far as final results are concerned, but which has the effect of altering the structure of quantum mechanics in such a way that the dynamics is capable of being formulated as a time-rate of change of probabilities.

II. MATHEMATICAL APPARATUS

We first put the wave function into a representation such that its change with time, under any time-independent Hamiltonian,[3] is a measure-preserving point transformation of its argument. The mathematical techniques involved have been fully described in the literature.[4] However, in order to make the basic ideas readily accessible to those unfamiliar with them we shall give here a full, but nonrigorous, account of all the underlying mathematics.

We have first to consider Wiener's mathematical generalization of the Brownian motion. We are actually not concerned with the literal, physical Brownian motion, but with a set of functions having statistical properties very similar to those of the set of functions (describing the displacement of a Brownian particle in its dependence on the time) that would be obtained if we selected a large number of paths of such particles at random from a given solution.

Imagine a fluid containing a large number of particles undergoing Brownian motion. Describe the path of each particle as a function of time. The ensemble of all these functions has rather easily describable statistical properties when considered on a comparatively gross scale; this is due to the *randomness* (again on a gross scale) of the impacts to which the Brownian particles are continually subjected, this randomness being reflected in a corresponding randomness of the paths. With a finer scale of description, however, a nonrandom element begins to appear. This is, of course, due primarily to the nonvanishing mass of the Brownian particle, which gives it the ability at least partially to maintain its direction and velocity through a series of not too large a number of molecular impacts.

Wiener postulates a set of functions which is an abstraction from the ensemble of physical Brownian motion paths. Roughly speaking, the abstraction consists mainly of retaining the randomness of the "displacements" under *indefinite* subdivision of the paths. Wiener's set also contains discontinuous functions, which of course would not exist at all in the physical Brownian motions; but all such functions, relative to the whole set, constitute a set of zero probability—in the

[1] M. Born, Z. Physik **37**, 863 (1926).

[2] We speak here in terms of the Schrödinger picture. In the Heisenberg picture, where the dynamics is applied to the operators representing physical observables rather than to the wave function, probabilities are (at least formally) still further removed from any role in the dynamic part of the theory. The operators are here entirely independent of the probabilities, which vary with the initial conditions of the problem.

[3] This is not an essential restriction, provided we may regard any system with a time-dependent Hamiltonian as part of a larger, isolated system having a time-independent Hamiltonian. The former case then involves an inessential alteration in the proposed statistical interpretation.

[4] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain* (American Mathematical Society, New York, 1934), Chap. 9; N. Wiener, Acta Math. **55**, 117–258 (1930), Sec. 13; N. Wiener, J. Math. and Phys. **2**, 131 (1923).

terminology of measure theory, "almost all" the Brownian motions are continuous. The statistical properties possessed as an approximation by the physical Brownian motions become *exact* for Wiener's set, i.e., they subsist even under indefinitely minute scrutiny of the "paths."

From now on, we shall refer to the generalized set of functions referred to above simply as "Brownian motion functions." Whenever we refer to the *physical* Brownian motion, we shall explicitly include the additional adjective.

We now proceed to give a partly heuristic, quantitative account of the statistical properties of the set of Brownian motion functions. Consider a one-dimensional Brownian motion taking place in a discrete "time" described by the variable x, which can take on the equally spaced values 0, $1/N$, $2/N$, $\cdots 1 (N = \text{integer})$. (This variable is not destined to play the role of time in the theory we shall ultimately describe, hence we avoid using the symbol t for it. It is, however, the analog of the time of the physical Brownian motion.) The "displacement" is a real function $X(\alpha, x)$. Here α is a parameter whose value singles out a particular function from the whole set. Wiener has shown that a set of real numbers α from zero to one is sufficient for this purpose, for Brownian motions not only in one dimension, but in any finite number of dimensions. We shall not prove this here, although we shall describe and utilize certain important properties of this mapping of the functions on the unit interval.

Let

$$\begin{aligned} X_j(\alpha) &= X(\alpha, j/N), \qquad (j=0, 1, \cdots N), \\ \Delta_j(\alpha) &= X_{j+1}(\alpha) - X_j(\alpha), \quad (j=0, 1, \cdots N-1). \end{aligned} \tag{2}$$

We now postulate that the "number"[5] of Brownian motions, in the whole set, having its jth increment Δ_j (for any given value of j) in the range $a<\Delta_j<b$, is proportional to

$$\frac{1}{(2\pi/N)^{\frac{1}{2}}} \int_a^b \exp\left(-\frac{\Delta^2}{2/N}\right) d\Delta,$$

independently of what ranges may be fixed simultaneously for the Δ's having other values of j. In other words, each Δ_j has a Gaussian distribution of values from $-\infty$ to $+\infty$, centered at zero, with mean square $1/N$, and statistically independent of all other increments; we shall say that the increments have *independent Gaussian distributions*. This will be recognized as the same as the gross statistical behavior of the increments of displacement in the physical Brownian motion: In the physical Brownian motion the independence of increments results from the large-scale lack of influence of all prior states of motion on the motion of the particle at any given moment, the increment in a given time interval thus being given entirely by the integrated molecular impulses received during that interval; while the Gaussian distribution can be shown from general probability laws to result from the fact that the individual molecular impulses are random in nature.

It can readily be shown that, with the above distribution for the elementary increments of the process, the distribution of the increment for an arbitrary multiple of the elementary "time interval" $1/N$ will obey the same Gaussian law, and will always have its mean square equal to the reciprocal of the length of the interval. We could multiply $1/N$ in the elementary distribution law by any constant, and have an analogous proportionality of the mean-square increment for arbitrary intervals; but for our present purposes such a coefficient need not be included explicitly.

Since the distribution law given above is normalized to unity, it may be used in the sense of a probability, i.e., the probability that $a \leqq \Delta_j(\alpha) \leqq b$ for α chosen at random is

$$\Pr\{a \leqq \Delta_j(\alpha) \leqq b\} = \frac{1}{(2\pi/N)^{\frac{1}{2}}} \int_a^b \exp\left(-\frac{\Delta^2}{2/N}\right) d\Delta. \tag{3}$$

In the above assertion, the basic property of the mapping of Brownian motions on the parameter α, which has not been given up to now, is implicitly assumed: The set of Brownian motions is *uniformly* distributed over the interval $0 \leqq \alpha \leqq 1$, i.e., the probability of a given subset is equal to the Lebesgue measure of the corresponding set of α's. If, moreover, the distribution law for the increments within each of the elementary "time" intervals $1/N$ is normalized to unity, as above, it is easy to show that the distribution law for any interval equal to an integer multiple of the basic interval is also normalized to unity.

Since the distributions of the Δ_j are mutually independent, it is a simple matter to obtain the probability that a randomly chosen Brownian motion will have all of its increments within a given set of ranges,

$$\begin{aligned} &\Pr\{a_1 \leqq \Delta_1(\alpha) \leqq b_1, \quad a_2 \leqq \Delta_2(\alpha) \leqq b_2, \\ &\qquad \cdots, \quad a_{N-1} \leqq \Delta_{N-1}(\alpha) \leqq b_{N-1}\} \\ &= \left(\frac{2\pi}{N}\right)^{-\frac{1}{2}(N-1)} \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_{N-1}}^{b_{N-1}} \exp\left(-\frac{\Delta_1^2}{2/N}\right. \\ &\qquad \left. -\frac{\Delta_2^2}{2/N} - \cdots - \frac{\Delta_{N-1}^2}{2/N}\right) d\Delta_1 d\Delta_2 \cdots d\Delta_{N-1}. \end{aligned} \tag{4}$$

This is equal to the measure of the set of α's for which $X(\alpha, x)$ satisfies the given inequalities.

The set of Brownian motions in a continuous "time" variable x varying from 0 to 1 may now be characterized in its statistical properties by going to the limit $N \to \infty$ in

[5] Actually, measure rather than number. The "number" here involved is infinite. It must thus be dealt with quantitatively in terms of measure, just as the "number" of points in a line segment is dealt with in terms of the measure of the segment—a segment being, of course, an example of a simple type of set for which the measure is given merely by its length.

the above equations. This gives the measure of any set of Brownian motions all of whose X *vs* x graphs lie between those of any two given Brownian motions, or the measure of any sum of (a denumerable number of) such sets.

The general reasoning so far given can be readily extended to any range of the x variable, including an infinite range, without any essential changes.

The multidimensional Brownian motion is a simple generalization of the one-dimensional case. X will in this case have a number of components, all functions of one time variable x. The increments of any component are independent of one another as well as of those of the other components; the measure of a set is given by an obvious generalization of the method already given, based on the independence of all increments. As already mentioned, a measure-preserving mapping onto the interval (0, 1) is still possible.

An important special case of multidimensional Brownian motion, which is particularly important to us, is the *complex* Brownian motion. Here X is a complex number, a function of the real variable x and the parameter α. The increments of its real and imaginary parts are all independent of one another.

The usefulness for us of the set of Brownian motions lies in the fact that it is capable of giving a measure for a set of functions. If a function is thought of as a vector in function space in the usual way, namely, if the variable x in the function $f(x)$ is considered as a component index, i.e., $f(x_0)$ is the x_0th component, etc., of the representative vector, then one might expect to have a measure of certain simple sets of functions, simply by using a volume element in the function space. This expectation is, of course, not borne out, because a volume element in a space of an infinite number of dimensions can only be equal to zero or infinity. The statistical weighting technique used with the Brownian motions, on the other hand, gives a manageable definition of measure for a set of functions.

This measure is made visualizable in another kind of space, the "differential space." Imagine all the increments Δ_j divided by the increment $\Delta x=1/N$. Then the set

$$\frac{\Delta_1(\alpha)}{\Delta x}, \quad \frac{\Delta_2(\alpha)}{\Delta x}, \quad \cdots, \quad \frac{\Delta_{N-1}(\alpha)}{\Delta x}$$

is a set of difference quotients which approaches, in the limit $N\to\infty$, the values of the derivative of $X(\alpha, x)$. (Since almost all of the $X(\alpha, x)$ are nondifferentiable, most of the difference quotients will approach infinity; but one may imagine the derivative as obtained by termwise differentiation of an expansion of $X(\alpha, x)$ in an orthogonal series of continuous functions.) The ordinary function-space representation of the derivative of $X(\alpha, x)$ has components equal to the above difference quotients, in the limit $N\to\infty$, $\Delta x\to 0$. Again, a set of finite components may be obtained, in a nonrigorous sense of course, by a Fourier transformation of $X(\alpha, x)$ followed by termwise differentiation: Let the range of x be from 0 to 1, and expand $X(\alpha, x)$. In terms of the orthogonal set $e^{2\pi inx}$,

$$X(\alpha, x)=\sum_{n=-\infty}^{\infty} a_n e^{2\pi inx},$$

$$dX(\alpha, x)/dx=\sum 2\pi ina_n e^{2\pi inx}.$$

Thus $dX(\alpha, x)/dx$ is represented in Hilbert space by the components $2\pi ina_n$ with respect to the orthogonal axes given by the functions $e^{2\pi inx}$. The norm of dX/dx, $\sum|2\pi ina_n|^2$, will not converge, of course, except for a set of α of measure zero, hence dX/dx is undefined, strictly speaking.

Differential space, referred to above, is not quite the same as the function space of the derivative; it is defined as the space in which a function is represented by its *differentials* ("limits" of the increments Δ_j) as coordinates, rather than by the values of its derivative. This makes the norm of the vector almost always finite, since

$$\lim_{N\to\infty}\sum_{j=0}^{N-1}|\Delta_j|^2$$

is the quadratic variation of the function, which is finite for almost all of the Brownian functions as a result of their definition.

If now we select a large number of Brownian motions at random, and imagine the fine dust formed by their representative points in differential space, we would find a Gaussian distribution of the density of this dust, centered at the origin, as we moved out along any axis. The number of dust particles in a given region of the differential space is proportional to the measure (in physical terms, probability) assigned to this region by going to the limit $N\to\infty$ in Eq. (4), or the counterpart of Eq. (4) for a Brownian motion of whatever number of components may be involved. This in turn is the measure assigned to the functions satisfying the inequalities given in the curly brackets on the left-hand side of Eq. (4).

The mean squares of all increments are kept equal to one another in the limiting process $N\to\infty$, and therefore the density of the "dust" in differential space is hyperspherically symmetrical. Considering for a moment a real Brownian motion, this corresponds to the fact that in Eq. (4) the exponent of the Gaussian function is proportional to the squared magnitude of the differential-space vector. Hence any rotation of the axes in differential space, or any rigid rotation of all vectors in the space, would leave this exponent invariant, and we would have the result that the independent Gaussian distribution of differential-space components of the set of real Brownian motions is invariant to such rotations, i.e., to any real orthogonal transformation of these

functions. It is, in fact, readily apparent that such transformations transform the set into itself.

In the case of complex Brownian motions, the Gaussian exponent is equal to the sum of the squares of the real and imaginary parts of the complex increments in the limit $N\to\infty$, or the sum of the quadratic variations of the real and imaginary parts of the Brownian function. In this case, the analog of the rigid rotation in the real case is a unitary transformation; hence a unitary transformation of the differential-space vector conserves the independent Gaussian distributions of the differentials and transforms the set of complex Brownian motions into itself.

Consider now a subdivision of the interval (0, 1) of definition of the x variable of the Brownian motion labeled as follows:

$$x_0=0,\quad x_1=1/N,\quad x_2=2/N,\quad \cdots,\quad x_N=1,$$

and a matrix depending on the two indices x_i, y_m, having matrix elements $\varphi(x_i, y_m)$. These satisfy the unitarity conditions,

$$\begin{aligned}\sum_{y_m}\bar{\varphi}(x_i, y_m)\varphi(x_j, y_m)&=\delta_{ij},\\ \sum_{x_i}\bar{\varphi}(x_i, y_m)\varphi(x_i, y_n)&=\delta_{mn}.\end{aligned}\tag{5}$$

The quantity

$$\psi(\alpha, y_m)=\sum_{i=0}^{N}\bar{\Delta}_i(\alpha)\varphi(x_i, y_m)\tag{6}$$

then has the meaning of a component, namely, the mth of the vector representing the αth Brownian motion in the differential space (of Brownian functions of the discrete time variable x) with respect to a new set of axes labeled by the indices y_m; or alternatively, as a component of the vector obtained by rotating the original Brownian motion vector. In the case of a unitary transformation represented by a function $\varphi(x, y_m)$ which is continuous in x, the counterpart of Eq. (6) is the Stieltjes integral

$$\psi(\alpha, y_m)=\int d\bar{X}(\alpha, x)\varphi(x, y_m).\tag{7}$$

Because the Stieltjes integral is not rigorously defined for all Brownian motions $X(\alpha, x)$, since not all are of bounded variation, it is necessary to state that (7) is to be interpreted only in a formal sense. The right-hand side of Eq. (7) obtains its rigorous definition through integration by parts

$$\psi(\alpha, y_m)=-\int\bar{X}(\alpha, x)d\varphi(x, y_m).\tag{8}$$

As a result of the interpretation given $\psi(\alpha, y_m)$ in the case of discrete x, immediately following Eq. (6), we would expect that the distribution of the values of the real and imaginary parts of $\psi(\alpha, y_m)$, obtained by letting α range through all its values, is Gaussian and independent of the values of the similarly distributed real and imaginary parts of all $\psi(\alpha, y_n)$ for which $n\neq m$. This is indeed the case; the proof is given in the references cited, and we shall not repeat it here.

The functions $\psi(\alpha, y_m)$ may be interpreted in another way: They are the coefficients in the formal expansion of $d\bar{X}(\alpha, x)/dx$ in terms of the orthogonal set of functions $\varphi(x, y_m)$. If for $\varphi(x, y_m)$ we take the exponential function $e^{2\pi imx}$ (the interval of x being from 0 to 1), the term-by-term integration of the series gives the expansion of $\bar{X}(\alpha, x)$ in terms of these functions. Going back to Eq. (7), with $\varphi(x, y_m)$ an arbitrary orthogonal set other than the complex exponentials, the expansion of $X(\alpha, x)$ in complex exponentials may be used in this equation to yield a formal proof of a second basic property of the functions $\psi(\alpha, y_m)$: They are a *normal orthogonal set in* α, if the $\varphi(x, y_m)$ are normal and orthogonal in x,

$$\begin{aligned}\int_0^1\bar{\psi}(\alpha, y_m)\psi(\alpha, y_n)d\alpha&=\int\bar{\varphi}(x, y_m)\varphi(x, y_n)dx\\ &=\delta_{mn}.\end{aligned}\tag{9}$$

Hence the quantities $\Delta_i(\alpha)$ are in a generalized sense also representatives of a unitary transformation between the discrete "axes" represented by the letters i and the continuous variable α; and the improper functions $d\bar{X}(\alpha, x)/dx$ represent a generalized unitary transformation between the two continuous variables x and α. Hence $\psi(\alpha, y_m)$ has a double meaning: It may be regarded as (*a*) the representative of $\varphi(x, y_m)$ transformed from the variable x to α, or (*b*) the representative of the formal derivative $d\bar{X}(\alpha, x)/dx$ transformed from the variable x to y_m. The independence of the $\psi(\alpha, y_m)$ follows from (*b*), and their orthogonality follows from (*a*); but it should be noted that, even if the $\varphi(x, y_m)$ are a complete set in the variable x, the $\psi(\alpha, y_m)$ are not a complete set in α.

III. THE WAVE FUNCTION IN THE "α REPRESENTATION"

We now consider x as the spatial variable of a normalizable wave function $\varphi(x, t)$, where t represents the time.

Let

$$\psi(\alpha, t)=\int d\bar{X}(\alpha, x)\varphi(x, t).\tag{10}$$

We shall call this the "wave function in the α representation." The use of the term "representation" is justified by the remarks above to the effect that $d\bar{X}(\alpha, x)/dx$ behaves like a unitary transformation representative. This is an extension of the usual meaning of a representation, since α is equivalent to an infinity of ordinary variables of the type of the spatial variable x and,

although $d\bar{X}(\alpha, x)/dx$ behaves, as we saw above, like the representative of a unitary transformation, it completely lacks the symmetry that would be expected between α and x in an ordinary unitary transformation.

Suppose now that $\varphi(x, t)$ transforms with time according to the equation

$$\varphi(x, t)=\int u(x, x'; t)\varphi(x', 0)dx', \tag{11}$$

where $u(x, x'; t)$ is a unitary kernel. If (11) is substituted into (10), the result can be formally altered according to the following steps:

$$\begin{aligned}\psi(\alpha, t)&=\int d\bar{X}(\alpha, x)\int u(x, x'; t)\varphi(x', 0)dx'\\&=\int dx'\left[\int d\bar{X}(\alpha, x)u(x, x'; t)\right]\varphi(x', 0)\\&=\int d\bar{X}(T^t\alpha, x')\varphi(x', 0).\end{aligned} \tag{12}$$

The first line shows the result of direct substitution. The second is obtained by interchanging the order of the integrations. The quantity

$$dx'\int d\bar{X}(\alpha, x)u(x, x'; t) \tag{13}$$

can be seen, by a reasoning analogous to that given in connection with Eq. (7), to be the formal expression for the x' component in differential space of a vector obtained from that having components $d\bar{X}(\alpha, x)$ by a "rotation"—i.e., a unitary transformation in differential space.[6] Hence $dx'\int d\bar{X}(\alpha, x)u(x, x'; t)$ is the x' component of a new Brownian motion, and the transformation merely rotates the entire differential space. From this it follows that the set of Brownian motions is transformed into itself. This transformation is denoted in the third line of Eq. (12) by the operator T^t; $T^t\alpha$ is the new value of α obtained from the rotation of the Brownian function having parameter α by the unitary kernel $u(x, x'; t)$. As suggested by the hyperspherical symmetry of the distribution in differential space, this rotation sends a given region of differential space into another region with preservation of measure. Crudely speaking, the "number" of Brownian motions in the original region and the transformed region is unaltered by the transformation; the rigorous statement of this fact is that for all α belonging to a given measurable set on the interval $0\leqq\alpha\leqq 1$ the measure of the set of the corresponding $T^t\alpha$ is equal to that of the original set.

[6] The formal parallelism between Eqs. (7) and (13) is established by associating ξ_m with x' and $\varphi(x, \xi_m)$ with $u(x, x'; t)dx'$; the differential in the latter quantity is necessary in doing this because x' is a continuous variable while ξ_m is discrete.

Comparing the third line of Eq. (12) with Eq. (10), we see that

$$\psi(\alpha, t)=\psi(T^t\alpha, 0), \tag{14}$$

i.e., *the change with time in $\psi(\alpha, t)$ is in general a measure-preserving point transformation on the argument α.* Such a result is impossible for the wave function in any of the usual representations of quantum mechanics. It would contradict the characteristic spreading-out in time of wave functions in representations such that operators diagonal in them do not commute with the Hamiltonian, and would be similarly inconsistent with the mode of change with time [through exponentials $\exp(iEt/\hbar)$] of wave functions in representations such that the diagonal operators do commute with the Hamiltonian.

The transformation $\psi(\alpha, 0)\rightarrow\psi(\alpha, t)=\psi(T^t\alpha, 0)$ may be "visualized" in differential space as follows: $\psi(\alpha, 0)$ associates with each point α in this space a complex number $\psi(\alpha, 0)$. The time-transformation brought about by the dynamics in time t corresponds to a "rigid rotation" of the generalized contours in differential space that characterize this function, i.e., to a rotation of the structure of complex numbers that are values of $\psi(\alpha, 0)$ for the different points α.

IV. IDENTIFICATION OF DIFFERENTIAL-SPACE POINTS WITH INDIVIDUAL SYSTEMS

With quantum dynamics in the form of a transformation of points in differential space into one another, the final step that we want to make is the identification of these points with individual systems having well-defined values of all observable quantities. This section is devoted to establishing an algorithm for this purpose, which takes the form of a new statistical interpretation for the wave function, as promised in the introductory section.

Consider an observable $\mathfrak{R}$ to which corresponds an operator $\mathbf{R}$ acting on the variable x of the wave function $\varphi(x)$. Suppose at first that $\mathbf{R}$ has only discrete eigenvalues R_1, R_2, $\cdots$ corresponding to eigenfunctions $\chi_1(x)$, $\chi_2(x)$, $\cdots$, i.e.,

$$\mathbf{R}\chi_1(x)=R_1\chi_1(x), \quad \mathbf{R}\chi_2(x)=R_2\chi_2(x), \quad \cdots. \tag{15}$$

We suppose the $\chi_i(x)$ normalized and orthogonal. Expand the wave function in terms of the $\chi_i(x)$,

$$\varphi(x)=\sum_{i=1}^{\infty} a_i\chi_i(x). \tag{16}$$

Now let

$$\xi_i(\alpha)=\int d\bar{X}(\alpha, x)\chi_i(x). \tag{17}$$

Then, integrating term by term, we get

$$\psi(\alpha)=\int d\bar{X}(\alpha, x)\varphi(x)=\sum_{i=1}^{\infty} a_i\xi_i(\alpha). \tag{18}$$

We now consider each $\xi_i(\alpha)$ as a random variable, a function in turn of the random variable α. As we have seen (end of Sec. II), the distributions of the real and imaginary parts of $\xi_i(\alpha)$ are mutually independent of one another and of those of all other $\xi_j(\alpha)$ and are Gaussian in form, i.e., the relative probability that any $\mathrm{Re}\xi_i(\alpha)$ or $\mathrm{Im}\xi_i(\alpha)$ lies between a and b is

$$\frac{1}{(2\pi)^{\frac{1}{2}}}\int_a^b \exp(-y^2/2)dy, \tag{19}$$

independently of what ranges may be fixed simultaneously for the other part of the same ξ_i and the real and imaginary parts of the others.

We now assign an eigenvalue of $\mathbf{R}$ to any given value of α in the following way: Divide $\psi(\alpha)$ into two parts, say

$$\psi(\alpha)=\psi_{11}(\alpha)+\psi_{12}(\alpha), \tag{20}$$

$$\psi_{11}(\alpha)=\sum_{i=1}^{n_1} a_i\xi_i(\alpha), \quad \psi_{12}(\alpha)=\sum_{i=n_1+1}^{\infty} a_i\xi_i(\alpha). \tag{21}$$

Since ψ_{11} and ψ_{12} are orthogonal, the distributions of their real and imaginary parts are independent and Gaussian. The mean square of $\mathrm{Re}\psi_{11}$, $\langle(\mathrm{Re}\psi_{11})^2\rangle$, equals that of $\mathrm{Im}\psi_{11}$, and similarly for $\mathrm{Re}\psi_{12}$ and $\mathrm{Im}\psi_{12}$. Hence the mean square is

$$\sigma_{11}=\langle(\mathrm{Re}\psi_{11})^2\rangle=\langle(\mathrm{Im}\psi_{11})^2\rangle=\tfrac{1}{2}\langle|\psi_{11}|^2\rangle$$
$$=\tfrac{1}{2}\int_0^1 |\psi_{11}(\alpha)|^2d\alpha=\tfrac{1}{2}\int|\chi_{11}(x)|^2dx=\tfrac{1}{2}\sum_{i=1}^{n_1}|a_i|^2 \tag{22}$$

for $\mathrm{Re}\psi_{11}$ and $\mathrm{Im}\psi_{11}$, and

$$\sigma_{12}=\langle(\mathrm{Re}\psi_{12})^2\rangle=\langle(\mathrm{Im}\psi_{12})^2\rangle=\tfrac{1}{2}\sum_{i=n_1+1}^{\infty}|a_i|^2 \tag{23}$$

for $\mathrm{Re}\psi_{12}$ and $\mathrm{Im}\psi_{12}$. We shall take $\varphi(x)$ normalized, hence

$$\int_0^1|\psi(\alpha)|^2d\alpha=\int[|\psi_{11}(\alpha)|^2+|\psi_{12}(\alpha)|^2]d\alpha$$
$$=2(\sigma_{11}+\sigma_{12})=1.$$

We now determine which is bigger, $|\psi_{11}(\alpha)|$ or $|\psi_{12}(\alpha)|$. If the first, we take R to lie in the set $R_1,\cdots R_n$, which we shall call S_{11}; otherwise, in the remaining set S_{12}. This completes the first step in an infinite dichotomic process, consisting of steps similar to the above, whereby the choice of an R belonging to α is successively narrowed down, leading in the limit to a unique value (except for a set of α of measure zero) of R, which we shall call $R(\alpha)$. In the next step, we take ψ_{11} or ψ_{12}—depending on which was chosen in the first step—and proceed with it exactly as we did in the first step, dividing it into two parts, finding which is of bigger modulus, and assigning $R(\alpha)$ to the corresponding range; and so on.

We now wish to know, first, the *probability* of $R(\alpha)$ lying in S_{11} or S_{12}. Since the real and imaginary parts of $\psi_{11}(\alpha)$ and $\psi_{12}(\alpha)$ are all mutually independent, the probability that $|\psi_{11}(\alpha)|\geqq|\psi_{12}(\alpha)|$, i.e. $|\psi_{11}(\alpha)|^2\geqq|\psi_{12}(\alpha)|^2$, is

$$\iiiint_{(u^2+v^2\geqq x^2+y^2)}\exp\left(-\frac{u^2+v^2}{2\sigma_{11}}\right)\exp\left(-\frac{x^2+y^2}{2\sigma_{12}}\right)dudvdxdy$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left(-\frac{u^2+v^2}{2\sigma_{11}}\right)\exp\left(-\frac{x^2+y^2}{2\sigma_{12}}\right)dudvdxdy$$

$$=\frac{4}{\sigma_{11}\sigma_{12}}\int_0^{\infty}\exp\left(-\frac{t}{2\sigma_{12}}\right)dt\int_t^{\infty}\exp\left(-\frac{s}{2\sigma_{11}}\right)ds=\frac{\sigma_{11}}{\sigma_{11}+\sigma_{12}}=2\sigma_{11}=\sum_{i=1}^{n_1}|a_i|^2, \tag{24}$$

which is equal to the relative probability according to the Born postulate that the measured value of $\mathfrak{R}$ lie in the set S_{11}. The probability that $R(\alpha)$ is in S_{12} is obviously $2\sigma_{12}$, and the sum of the two probabilities is 1.

At any later step, moreover, we shall obtain a set of probabilities for $R(\alpha)$ lying in the ranges defined at this step (there will be at this step 2^n ranges in the subdivision of the total infinite range of R) having the following property: The relative probabilities of the two ranges obtained by subdivision of any single one of the ranges of the previous step are proportional to the Born probabilities. Since the absolute probabilities of the two ranges obtained in the first step agree with the Born probabilities, it follows that those of the second step also agree, and so on, as may be shown by mathematical induction.

We are evidently at liberty to choose the indices defining the successive subdivisions in any way we wish. In particular, with normalizable wave functions the eigenfunctions can be numbered and n_1 chosen so that the probability of S_{12} is as small as we wish, and the set S_{11} is simultaneously finite. This choice insures that, after some finite number of dichotomies for each α, the set of all α that are still not assigned an $\mathbf{R}$ eigenvalue is of measure as small as we wish, i.e., that the still unassigned $\mathbf{R}$ eigenvalues are of correspondingly small probability.

The case where $\mathbf{R}$ has continuous eigenvalues, i.e.,

where

$$\varphi(x)=\int_{-\infty}^{\infty}\chi(x,R)a(R)dR, \tag{25}$$

$\chi(x,R)$ satisfying the orthogonality and normalization conditions

$$\int\bar{\chi}(x,R)\chi(x,R')dx=\delta(R-R'),$$

$$\int\bar{\chi}(x,R)\chi(x',R)dR=\delta(x-x'),$$

is treated along the same lines as the discrete case, by exploiting the formal parallel between $a(R)$ and a_i, and between $\chi(x,R)$ and $\chi_i(x)$. Specifically, we proceed as follows:

Let

$$\psi(\alpha)=\int d\bar{X}(\alpha,x)\varphi(x), \tag{26}$$

as before. Then

$$\psi(\alpha)=\int_{-\infty}^{\infty}a(R)dR\int d\bar{X}(\alpha,x)\chi(x,R) \tag{27a}$$

$$=\int_{-\infty}^{\infty}d\bar{\xi}(\alpha,R)a(R), \tag{27b}$$

where

$$d\bar{\xi}(\alpha,R)=dR\int d\bar{X}(\alpha,x)\chi(x,R). \tag{28}$$

Now put

$$\psi_{11}(\alpha)=\int_{-\infty}^{A_1}d\bar{\xi}(\alpha,R)a(R), \tag{29a}$$

$$\psi_{12}(\alpha)=\int_{A_1}^{\infty}d\bar{\xi}(\alpha,R)a(R). \tag{29b}$$

These two quantities have, like the like-named quantities in the discrete case, independent Gaussian distributions for their real and imaginary parts, and are orthogonal to one another. The mean squares of the real and imaginary parts of, respectively, ψ_{11} and ψ_{12}, are

$$\sigma_{11}=\tfrac{1}{2}\int_{-\infty}^{A_1}|a(R)|^2dR,\quad \sigma_{12}=\tfrac{1}{2}\int_{A_1}^{\infty}|a(R)|^2dR, \tag{30}$$

i.e., again the usual quantum-mechanical probabilities for finding a value of R in the ranges indicated by the limits of integration. From here on, the calculation is the same as in the discrete case: One assigns to a given α the range $(-\infty, A_1)$ or $(A_1, +\infty)$, depending on whether $|\psi_{11}(\alpha)|$ or $|\psi_{12}(\alpha)|$ is the greater. The measures of the two sets of values of α corresponding to these ranges will be $2\sigma_{11}$ and $2\sigma_{12}$, respectively (if $\varphi(x)$ is normalized), agreeing with the quantum-mechanical probabilities for these two ranges. Successive steps in the sequence of dichotomies are modeled on the first one, exactly as in the case of discrete eigenvalues.

Although we have in the above development given a privileged status, so to speak, to the variable x, formally speaking x and R are really on the same footing. This is readily apparent from the completely parallel form of Eqs. (26) and (27b): $\xi(\alpha,R)$ is a complex Brownian motion in R, and $X(\alpha,x)$ and $\xi(\alpha,R)$ may be thought of as representatives, in terms of different variables, of a single element of an ensemble of abstract elements represented by points in differential space. Using Dirac's bra and ket notation, it is possible to demonstrate this equivalence in a formal way. We take advantage of the fact already shown that $d\bar{X}(\alpha,x)/dx$ and $d\bar{\xi}(\alpha,R)/dR$ behave to some extent like unitary transformation coefficients; let us then denote them by $\langle\alpha|x\rangle$ and $\langle\alpha|R\rangle$, respectively, although it must be borne in mind that this notation is not completely justified in view of the lack of symmetry between the α variable and the ordinary variables of quantum theory. Let us put

$$\chi(x,R)=\langle x|R\rangle,$$
$$\varphi(x)=\langle x|\varphi\rangle,$$
$$a(R)=\langle R|\varphi\rangle,$$

and

$$\psi(\alpha)=\langle\alpha|\varphi\rangle.$$

The first equation is of standard type. The second and third are the ordinary consequences, expressed in the Dirac notation, of the assumption that the state vector of the system is $|\varphi\rangle$. The fourth equation has the same interpretation as the second and third, once one accepts the treatment of α as on the same footing as x or R as is implied by the very use of the symbols $\langle\alpha|x\rangle$ and $\langle\alpha|R\rangle$. Equations (25), (26), (27a), and (27b) may now be written

$$\langle x|\varphi\rangle=\int\langle x|R\rangle dR\langle R|\varphi\rangle, \tag{25'}$$

$$\langle\alpha|\varphi\rangle=\int\langle\alpha|x\rangle dx\langle x|\varphi\rangle \tag{26'}$$

$$=\int\int\langle\alpha|x\rangle dx\langle x|R\rangle dR\langle R|\varphi\rangle \tag{27'a}$$

$$=\int\langle\alpha|R\rangle dR\langle R|\varphi\rangle. \tag{27'b}$$

In order to go from (27'a) to (27'b) directly, one has to assume

$$\langle\alpha|R\rangle=\int\langle\alpha|x\rangle dx\langle x|R\rangle, \tag{28'}$$

which is entirely consistent with the previous identifications and corresponds to (28).

When **R** has a discrete spectrum, the situation is formally not changed, and we shall not trouble to write

down all the equations corresponding to (25′) to (27′b). The transformation equations

$$\langle\alpha|R_i\rangle=\xi_i(\alpha)=\int\langle\alpha|x\rangle dx\langle x|R_i\rangle,$$

$$\langle\alpha|x\rangle=d\bar{X}(\alpha,x)/dx=\sum_i\langle\alpha|R_i\rangle\langle R_i|x\rangle,$$

where

$$\langle x|R_i\rangle=\chi_i(x),$$

reveal that $\xi_i(\alpha)$ is the ith increment of a Brownian motion in the *discrete* "time" variable R which takes on the values $R_1, R_2, \cdots$.

We can now give a concise general statement of the new statistical postulate for quantum mechanics: Given a division of the (continuous or discrete) set of eigenvalues of a quantum-mechanical operator into two non-overlapping subsets,[7] and a normalizable state vector $|\varphi\rangle$. The state vector is projected onto the two subspaces of Hilbert space corresponding to the two subsets of eigenvalues, and the "wave functions" in the α representation for these two projections are evaluated. The "wave function" having the larger modulus is then selected and its subspace divided in turn into two non-overlapping sub-subspaces. The "wave functions" in α representing the projections on these smaller subspaces are again compared in modulus, the larger selected, and so on. At each stage the eigenvalue of R that we seek is assigned to the subset corresponding to the subspace chosen, and its range thus narrowed down, in the limit, to a single number $R(\alpha)$. At each stage, moreover, it can be shown that the relative probability of the two subsets of eigenvalues involved is equal to that given by the Born postulate; the total probability of all subsets is one, so that the absolute probability of any subset, with this normalization, also agrees with that of the Born postulate.

In this way we may obtain a function $R(\alpha)$ for each observable $\mathfrak{R}$, such that

$$\int_0^1 R(\alpha)d\alpha=\langle\mathfrak{R}\rangle, \tag{34}$$

where $\langle\mathfrak{R}\rangle$ is the quantum-mechanical expectation value. Since the set of such functions for all observables also gives the probabilities for any defined ranges of values of any observables, we have a postulate entirely equivalent to the Born postulate. Moreover, we now have quantum dynamics expressed as a transformation of the same random variable α as that which furnishes, via the measure of appropriate subsets, the probability basis of the theory. This conclusion, stated only briefly here, will be discussed more fully in a subsequent paper.

V. NONUNIQUENESS OF $R(\alpha)$. THE POLYCHOTOMIC METHOD

The particular way in which the Hilbert space is divided at each stage of the succession of dichotomies is evidently entirely arbitrary. This means that the R value assigned to a given α is correspondingly arbitrary. For example, if at any stage R is found to be in a given subset of eigenvalues, the extension at this stage of the subspace corresponding to this subset at the expense of the other subspace might well throw R into the other subset, even though it has become smaller; and by the very nature of the dichotomous process, R cannot at any later stage re-enter the region from which it was thus ejected.

A brief calculation will show that the division of the subspace at each stage of the sequence into more than two parts will not work; for example, if one divides the subspace into *three* parts and assigns R to the subset of eigenvalues for which the projection of the wave function is the largest in modulus of the three, the relative probabilities of the three subsets do not satisfy Born's postulate.

It is interesting to note that this limitation can be removed by an alteration of the method in a different respect, as follows (for the sake of illustration, we treat first the case of discrete eigenvalues): We choose as $R(\alpha)$ the eigenvalue R_k for which $|\xi_k(\alpha)/a_k|$ (see Eq. (18)) is the *smallest of the entire set* of these quantities; the probability that $|\xi_k(\alpha)/a_k|$ is the smallest of all the $|\xi_i(\alpha)/a_i|$ is just $|a_k|^2$, as given by Born's postulate. To prove this, consider first a wave function for which only n coefficients a_i fail to vanish, and let these be named $a_1, a_2, \cdots a_n$. We first find the probability that

$$|\xi_n/a_n|<|\xi_{n-1}/a_{n-1}|<\cdots<|\xi_1/a_1|,$$

which we shall call for brevity

$$\Pr\{n<n-1<\cdots<1\}.$$

The quantities $|\xi_i(\alpha)/a_i|$ have mutually independent Gaussian distributions of their real and imaginary parts, with mean squares $\frac{1}{2}|a_i|^2$. Let $A_i=|a_i|^2$. Then

$$\Pr\{n<n-1<\cdots<1\}=\frac{\int_0^\infty e^{-A_nu_n}du_n\int_{u_n}^\infty e^{-A_{n-1}u_{n-1}}du_{n-1}\cdots\int_{u_2}^\infty e^{-A_1u_1}du_1}{\int_0^\infty e^{-A_nu_n}du_n\int_0^\infty e^{-A_{n-1}u_{n-1}}du_{n-1}\cdots\int_0^\infty e^{-A_1u_1}du_1}$$

$$=\frac{A_n}{A_1+\cdots+A_n}\cdot\frac{A_{n-1}}{A_1+\cdots+A_{n-1}}\cdots\frac{A_2}{A_1+A_2}. \tag{35}$$

[7] If the eigenvalues are discrete, no further characterization is necessary. If they are continuous, the subsets must be measurable.

The probability that $|\xi_n/a_n|$ is the smallest of all, $\Pr\{n<n-1, n-2, \cdots, 1\}$, is the sum of the above over all permutations of $A_1, A_2, \cdots A_{n-1}$. To carry out this summation, we note that since (35) must hold for any n,

$$\Pr\{n-1<n-2<\cdots<1\} = \frac{A_{n-1}}{A_1+\cdots+A_{n-1}} \frac{A_{n-2}}{A_1+\cdots+A_{n-2}} \cdots \frac{A_2}{A_1+A_2}; \quad (36)$$

hence

$$\Pr\{n<n-1<\cdots<1\} = \frac{A_n}{A_1+\cdots+A_n} \Pr\{n-1<n-2<\cdots<1\}, \quad (37)$$

and for the sum of the permutations we obtain

$$\begin{aligned}\Pr\{n<n-1, n-2, \cdots, 1\} &= \sum_{\substack{\text{perm. of}\\ A_1, \cdots, A_{n-1}}} \Pr\{n<n-1<\cdots<1\} \\ &= \frac{A_n}{A_1+\cdots+A_n} \sum_{\substack{\text{perm. of}\\ A_1, \cdots, A_{n-1}}} \Pr\{n-1<n-2<\cdots<1\} \\ &= \frac{A_n}{A_1+\cdots+A_n}, \quad (38)\end{aligned}$$

which is the Born probability of the eigenvalue R_n. Since the numbering of the eigenvalues is arbitrary, the theorem holds for any of the R_i in the set $R_1, R_2, \cdots R_n$. It holds, moreover, for any value of n, and therefore holds in the limit $n\to\infty$.

It is not necessary to carry the breakdown all the way to individual eigenvalues; we can assign $R(\alpha)$ merely to one or another subset of the set of n eigenvalues, with the correct probability for each subset as follows: If the n eigenvalues are to be divided into n' subsets, they may be renumbered in such a way as to make the indices consecutive in each group; then one has a set of numbers $1\leqq k_1<k_2\cdots<k_{n'-1}<n$ such that the particular inequality satisfied by i from the set

$$\begin{aligned}1&\leqq i\leqq k_1,\\ k_1&<i\leqq k_2,\\ &\cdots\\ k_{n'-1}&<i\leqq n,\end{aligned}$$

determines the subset to which R_i belongs. Let

$$\begin{aligned}b_1&=(\sum_1^{k_1}|a_i|^2)^{\frac{1}{2}},\\ b_2&=(\sum_{k_1+1}^{k_2}|a_i|^2)^{\frac{1}{2}},\\ &\cdots\\ b_{n'}&=(\sum_{k_{n'-1}+1}^{n}|a_i|^2)^{\frac{1}{2}};\end{aligned}$$

then

$$\begin{aligned}\xi_1'&=\frac{1}{b_1}\sum_1^{k_1}a_i\xi_i,\\ \xi_2'&=\frac{1}{b_2}\sum_{k_1+1}^{k_2}a_i\xi_i,\\ &\cdots\\ \xi_{n'}'&=\frac{1}{b_{n'}}\sum_{k_{n'-1}+1}^{n}a_i\xi_i,\end{aligned}$$

is a set of orthonormal functions such that

$$\psi(\alpha)=\sum_{i=1}^{n'}b_i\xi_i'(\alpha). \quad (39)$$

Then the preceding method may be used, with b_i replacing a_i, ξ_i' replacing ξ_i, and n' replacing n. This determines $R(\alpha)$ only as belonging to one or another of the n' sets, but gives the correct probability $|b_i|^2$ for the ith set, for any i.

The method for continuous eigenvalues is modeled on that for groups of discrete eigenvalues. If

$$\psi(\alpha)=\int_{-\infty}^{\infty}d\bar{\xi}(\alpha, R)a(R), \quad (40)$$

we break up the infinite range of integration into n' subsets; for simplicity, let these be intervals, although they need only be measurable, nonoverlapping subsets. Then we choose a set of numbers $r_1<r_2<\cdots r_{n'-1}$, corresponding to $k_1, k_2, \cdots k_{n'}$. If

$$\begin{aligned}b_1&=\left(\int_{-\infty}^{r_1}|a(R)|^2dR\right)^{\frac{1}{2}},\\ b_2&=\left(\int_{r_1}^{r_2}|a(R)|^2dR\right)^{\frac{1}{2}},\\ &\cdots\\ b_{n'}&=\left(\int_{r_{n'-1}}^{\infty}|a(R)|^2dR\right)^{\frac{1}{2}},\end{aligned}$$

then

$$\begin{aligned}\xi_1&=\frac{1}{b_1}\int_{-\infty}^{r_1}d\bar{X}(\alpha, R)a(R),\\ \xi_2&=\frac{1}{b_2}\int_{r_1}^{r_2}d\bar{X}(\alpha, R)a(R),\\ &\cdots\\ \xi_{n'}&=\frac{1}{b_{n'}}\int_{r_{n'-1}}^{\infty}d\bar{X}(\alpha, R)a(R),\end{aligned}$$

is a set of orthonormal functions such that

$$\psi(\alpha)=\sum_{i=1}^{n}b_i\xi_i(\alpha), \quad (41)$$

and the real and imaginary parts of the ξ_i's have independent Gaussian distributions with mean squares equal to $\frac{1}{2}$. Then the method of discrete eigenfunctions can be applied, determining $R(\alpha)$ to within one of the n subsets, with the correct probability $|b_i|^2$ for the ith subset.

The general method presented in this section, which we will call the polychotomic method, has certain advantages over the dichotomic method. First, it has

the advantage, so far as manipulation of its results is concerned, of yielding a result at one step. Second, for a given subdivision of the total set of eigenvalues, it gives a unique result, unlike the dichotomic method. In particular, when a discrete set of eigenvalues is divided up into individual eigenvalues, it gives a unique assignment of eigenvalues to values of α, except for those values of α for which two or more $|\xi_i(\alpha)/a_i|$ are smaller than all others but equal to one another, but these values of α form a set of measure zero and can be neglected. On the other hand, it does not yield a unique $R(\alpha)$ in the continuous case: A sequence of finer and finer subdivisions, each yielding a range for $R(\alpha)$ by the polychotomic method, will not give convergence to a unique single value of $R(\alpha)$ in the limit. The dichotomic method, on the other hand, does in the limit converge to a single value of $R(\alpha)$ for almost all values of α, for any given sequence of dichotomies.

The polychotomic method, on the other hand, can be used in a way analogous to the dichotomic method; this is because, using the polychotomic method with subsets of eigenvalues, one can take a set of α for which all R values belong to one of the subsets of R and make more precise assignments of R values within the subset, by a further polychotomy or sequence of polychotomies of arbitrary type. In particular, the mode of fixing upon a subset of R eigenvalues (to be associated with an α) characteristic of the polychotomic method could if desired be used with a sequence of dichotomies (to avoid confusion, such a procedure might be dubbed a dicho-polychotomy). Thus the polychotomic method is a generalization of the dichotomic method. A fixed succession of finer and finer dicho-polychotomies will converge for almost all α values. Here every finer division is a subdivision of the coarser divisions which precede it.

VI. CONCLUDING REMARKS

At this point, we wish to call attention to the fact that it was only for the sake of simplicity that we presented the preceding material as if x or R were single variables. No essential change is required if one has a system of more than one degree of freedom, in which case x and R stand for sets of more than one variable, some of which may even be discrete while others are continuous. The Brownian motion functions $X(\alpha, x)$ or $\eta(\alpha, R)$ are then still single complex numbers, but functions of several variables—complex Brownian motions of a multidimensional "time." Integrals like $\int d\bar{X}(\alpha, x)\varphi(x)$ are multidimensional. The set of values of α is still the unit interval. Where in the case of a single variable the division of $\psi(\alpha)$ into two parts is a division of the line into two subsets, a division of the multidimensional space of eigenvalues into subsets takes its place when x or R stands for more than one variable

It should be emphasized that the purpose of our theory is simply to achieve an interpretation of quantum mechanics in terms of probability densities, and not to reconcile quantum concepts with classical ones. The method rests on the identification of observables with operators in Hilbert space, and achieves quantization purely as a result of this assumption, which has basic, not derived, status.

The following simple example will suffice to show the nonclassical nature of the observables in our theory: By the method we have given, simultaneous values of position q and momentum p can be assigned to almost all values of α. Consider, on the other hand, the operator for p^2+q^2; its eigenvalues are quantized, and the value obtained for this observable, for a given α, will not in general coincide with the sum of the squares of the position and momentum variables belonging to this value of α. Thus our method does not restore the classical meaning of observables. In fact, it would seem reasonable to picture the process of determining an eigenvalue to go with an α value as the mathematical counterpart of "forcing" the system into an eigenstate by the measuring process. This "forcing" is closely associated with the uncontrollable disturbance of the system, emphasized by Bohr, which necessarily accompanies measurement of an observable when the system is not in an eigenstate of that observable.

Closely associated with the preceding considerations is the lack of invariance of the individual eigenvalues tied to values of α under transformations of the associated variables. If, for example, the zero point of the scale of $\mathbf{R}$ eigenvalues is translated an amount R_0, i.e., the wave function $a(R)$ is transformed into $a(R+R_0)$, the value of R belonging to an individual α, say $R(\alpha)$, does not go into $R(\alpha)-R_0$ (unless, of course, the state is an eigenstate of $\mathbf{R}$). It is, however, trivially evident—from the fact that our postulate is equivalent to the Born postulate—that the *distribution function* of $R(\alpha)$ is correctly displaced by an amount R_0, and this is all we can reasonably expect, given the essential impossibility of a classical type of assignment of R values to individual systems.

In this connection, we wish further to point out that, since the operator formalism of relativistic (Dirac) quantum mechanics is subsumed under that of nonrelativistic quantum mechanics, the application of our method to Dirac particles is straightforward. On the other hand, the preceding remarks show that, although distributions of relativistic dynamic variables in differential space transform covariantly under Lorentz transformations, given the correct transformations of the wave functions, the individual values attached to points in differential space do not do so (except, again, for eigenstates of the variable concerned). Here again it is impossible to attribute any meaning to these individual values independent of the act of measurement. No contradiction can come from this, since the verification of covariance of momentum and energy of a single particle would imply at least two measurements of the components in two different Lorentz frames, and since the disturbance due to a quantum-mechanical measurement prevents the second quantity measured from ever being "the same" as the first.