"Theory of Measurement" in Differential-Space Quantum Theory

ARMAND SIEGEL,* *Department of Physics, Boston University, Boston, Massachusetts*

AND

NORBERT WIENER, *Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts*

(Received August 17, 1955)

The "reduction of the wave packet" is an essential feature of the quantum description of measurement. Since it means an abrupt change in the wave function, it therefore also implies an abrupt change in the statistical description of the system. If the differential-space theory of quantum systems rests on a description of reality more detailed than quantum mechanics, it must contain an account of this change which is consistent with the strictly probabilistic point of view of the theory. It is here shown that the theory does satisfy this necessary condition. The need of a special postulate for the measurement process is eliminated in this formulation, and a new explicitness is introduced into the description.

I.

IN quantum mechanics, the phrase "theory of measurement"[1–3] has become a handy designation for those considerations having to do with the process called "reduction of the wave packet." In general, one speaks of reduction of the wave packet under the following circumstances: Two systems have been placed in interaction, as a result of which their states as superimposed in their wave function are correlated. An observation is made on one of them,[4] and the pure-state wave function assigned to the combined system is thereby changed to one consistent with the result of the observation. When the observations on one system are made for the purpose of yielding information about the state of the other, we have a further consequence, and this is the really interesting thing about this process: The new pure-state wave function of the combined system yields a new (in general, mixed) quantum-mechanical state for the second of the two systems concerned, with respect to observables of the second system, when those of the first are disregarded. And in the special case where each distinct state of one system is coupled to only one state, or comparatively few states, of the other, the new state of the second system is, respectively, almost pure or pure: In this case we call the process a *measurement* on the second system.

In the differential-space theory of quantum systems,[5,6] the statistics ordinarily obtained by the Born postulate from the wave function are embedded in a statistical ensemble of "representative" systems identical in properties with the system of interest but suitably distributed over values of observables. When the system of interest is compounded of two or more subsystems, the representatives are too. We may give the statistical description of the measurement process in three stages, as follows (our discussion will deal only with *exact* measurements). First stage: There are two independent systems and one has a differential-space ensemble for each, obtained by the polychotomic algorithm[5,6] from the two wave functions. Second stage: The two systems are interacting, and are thus to be considered as a single composite system. The appropriate ensemble has as its representative systems duplicates of this composite system, with distributions obtained from the wave function of the combined system. Third stage: An exact measurement has been made of the instrument system; assuming it to be nondegenerate, one may then treat the composite system in terms of a wave function for each part (the composite wave function is then a simple product). Thus we return to a differential-space description in terms of two ensembles, exactly as in the first stage. (If the instrument states are degenerate, one has a mixed state for the other system.)

If the foregoing were the only way to carry the differential-space description through the three stages of the process, one would have to look upon this method as incapable of giving a fundamental account of physical processes. As stated, the replacement of one statistical description by another as one goes from stage to stage requires an appeal to quantum mechanics, and seems to lie outside the domain of the differential-space method. Even on the formal level, one would want the differential-space description to be autonomous in this respect. Physically, one sees the meaning of this requirement even more clearly: If the method is a fundamental description of reality, then the ensembles represent statistically representative populations in a real sense; i.e., one should be able to use in place of any ensemble a set of real systems duplicating the distribution of observables in the ensemble, with the individual systems

* Research of this author supported in part by Office of Naval Research.

[1] F. London and E. Bauer, *La Théorie de l'observation en mécanique quantique* (Hermann et Cie, Paris, 1939).

[2] J. v. Neumann, *Mathematischen Grundlagen der Quantenmechanik* (Verlag Julius Springer, Berlin, 1932; Dover Publications, New York, 1943).

[3] David Bohm, *Quantum Theory* (Prentice-Hall, Inc., New York, 1951), Chap. 22.

[4] This part of the process has an aspect of undefinability to it. In order to avoid an infinite chain of interacting systems, at some point one posits an arbitrary act of cognition, called "l'acte d'objectivation," by London and Bauer (see reference 1). In the case of measurement, this is the "taking of the pointer reading" of the instrument; the required classical character of this part of the process has been emphasized by Bohr.

[5] N. Wiener and A. Siegel, Phys. Rev. **91**, 1551 (1953).

[6] N. Wiener and A. Siegel, Nuovo cimento (to be published).

all obeying a common dynamics [given by Eq. (2′) of reference 5]. This implies, in particular, that in the measurement process described above, it should be possible to trace through the transitions in the ensembles from stage to stage in terms of individual systems that maintain a continuous existence and always obey the appropriate dynamics. The purpose of this paper is to show that this requirement is satisfied: The transition from the first stage to the second corresponds to no change in membership of the ensembles involved, from the differential-space point of view, being merely a change in description—individual systems of the two separate ensembles are combined into pairs of composite systems. The transition from the second to the third corresponds to a selection of a subensemble from the ensemble of composite systems, followed by a change in description exactly the reverse of the one mentioned in the previous sentence; each system of the subensemble selected maintains a continouous existence (in fact, all systems do), subject to the known dynamics, throughout the process of selection.

The remainder of this paper consists of the formal demonstration of these remarks, with detailed interpretations. We give only the transition from the second to the third stage. As stated above, the changes in mode of description in the two transitions are the reverses of one another.

II.

We need first a mathematical description of composite systems suitable for the differential-space description. This is readily obtained in the product Hilbert space of the Hilbert spaces of the individual systems. For simplicity we use discrete representations, as always; the generalization to the continuous case is described in our previous papers.

Let us introduce a verbal distinction between the two systems described: The "first" system of Sec. I will be called the *instrument*; the "second" system, the *particle*. The work "particle" is used only in a manner of speaking; the "second" system might in reality be a gas of 6×10^{23} molecules, and we would still call it a particle; we know of no accurate, brief equivalent for the clumsy phrase "system under observation". The use of the words "instrument" and "particle" does not, moreover, restrict the argument to the case of the measurement process narrowly conceived. It applies equally well to any physical process described by the formal procedure of "reducing the wave packet."

The instrument and particle are both considered as having one degree of freedom. (Any system may be reduced to such a one by the formal procedure given in reference 6, Sec. 3.)

Let $\mathbf{R}$ be an l.h.c. operator[7] for the particle, with eigenvalues R_i, eigenkets $|R_i\rangle$; $\mathbf{S}$ an l.h.c. operator for the instrument, with eigenvalues S_i and eigenkets $|S_i\rangle$. $\mathbf{R}$ and $\mathbf{S}$ commute. They are taken to be nondegenerate and to have discrete spectra only. $\mathbf{RS}=\mathbf{T}$ is an l.h.c. operator for the system, particle+instrument. The eigenvalues

$$T_\mu=R_iS_j, \tag{1}$$

where $\mu=\mu(i,j)$ are either nondegenerate or may be made so by a replacement of either $\mathbf{R}$ or $\mathbf{S}$ by a suitable nondegenerate function of the operator replaced. The numbers $\mu(i,j)$ are to be distinct for distinct pairs (i,j); hence the inverses $i(\mu)$ and $j(\mu)$ are uniquely defined.

The differential space of the composite system has the eigenbasis

$$\{|T_\mu\rangle\},$$

and a point P in it has components

$$\zeta_\mu=\langle P|T_\mu\rangle, \tag{2}$$

whose real and imaginary parts are Gaussianly distributed, with measure

$$\prod_\mu(2\pi)^{-1}\exp(-|\zeta_\mu|^2/2)d\zeta_{\mu R}d\zeta_{\mu I} \tag{3}$$

assigned to the interval

$$\prod_\mu d\zeta_{\mu R}d\zeta_{\mu I} \tag{4}$$

(R and I in this formula stand for real and imaginary, respectively).

Imagine that the instrument and particle have been brought together so that they interact, after which they separate so that we may consider the interaction to have ceased. As a result the ket for the composite system must in general be expanded as a series of eigenkets $|T_\mu\rangle$, with coefficients c_μ:

$$|\psi\rangle=\sum_{i,j}c_\mu|T_\mu\rangle. \tag{5}$$

The observable $[\mathbf{S}]$ (we distinguish between the operator, $\mathbf{S}$, and the observable, $[\mathbf{S}]$)[6] is then measured on the instrument, and found to have the value $S_{j'}$. In ordinary quantum theory one then proceeds as follows: The sum (5) is changed instantaneously to the restricted sum over values of μ compatible with $j=j'$:

$$|\psi\rangle\rightarrow|\psi'\rangle=\sum_i c_{\mu(i,j')}|T_\mu\rangle. \tag{6}$$

This contraction of $|\psi\rangle$ into $|\psi'\rangle$ is just an immediate consequence of the Born statistical postulate applied to the observable $[\mathbf{S}]$. Since

$$|T_{\mu(i,j)}\rangle=|R_i\rangle|S_j\rangle, \tag{7}$$

a state vector $|S_{j'}\rangle$ may be factored out of each term of (6). The state of the particle (in the quantum-mechanical sense) is then independent of that of the instrument and its statistical properties are given by the state vector

$$|\varphi\rangle=\sum_i c_{\mu(i,j')}|R_i\rangle. \tag{8}$$

[7] We use the term "l.h.c. operator" to denote "linear Hermitian operator having a complete set of eigenstates," i.e., an "observable" in Dirac's nomenclature. (See reference 6.)

Now, the detailed description of the differential-space theory would go as follows: Before the measurement of $[\mathbf{S}]$ we must think in terms of an ensemble of *composite systems* working out their destinies in the *product* differential space of the operator **T**. Afterwards, as far as the particle is concerned, it is no longer necessary to be so general; all its statistical properties are given by an ensemble of *particles* in the *factor* differential space of **R**. But the determination of $[\mathbf{S}]$ as equal to $S_{j'}$ is in the differential-space description a selection of a subensemble of composite systems all of which have $[\mathbf{S}]=S_{j'}$ for their instrument components. Each of these is coupled to a particle. Hence we have selected a subensemble of particles, with known statistical distributions over all values of all particle observables. These distributions are the same for the subensemble chosen as they were before the measurement of $[\mathbf{S}]$, since the particle is not disturbed by this measurement (the interaction has ceased). We now come to the point of this paper, as indicated in Sec. I: It must be possible to identify each system of this subensemble in a one-to-one manner with a system of the subensemble obtained, independently, from $|\varphi\rangle$. In more precise language: It must be possible to identify all subsets of the differential-space set associated with the selected subensemble, weight for weight, with subsets of the differential-space ensemble independently constructed from $|\varphi\rangle$.

III.

Our task is in effect to show that the differential space of **R** pre-exists as a factor differential space in that of **T**, with regard to both measure and distribution of values of $[\mathbf{R}]$, for **R** any operator on the particle. In order to do this we must assign eigenvalues of **S** to points in the **T** differential space by applying the polychotomic algorithm not with respect to individual axes in **T** space, but with respect to manifolds each consisting of all eigenaxes of **T** having the same eigenvalue of **S**. [This is necessary for our present proof, but it should also be noted that, as shown in the appendix of reference 6, the former method would give nonunique results for $S(P,\psi)$, since **S** is a degenerate operator in the factor space of **T**; it is nondegenerate only in its own Hilbert space.] The mathematical method for doing this is contained in reference 5, in the paragraph containing Eq. (39); the following is simply an adaptation of this method to the present situation.

Let

$$b_j=(\textstyle\sum_i|c_{\mu(i,j)}|^2)^{\frac{1}{2}}, \tag{9}$$

and

$$|\psi_j\rangle=\frac{1}{b_j}\sum_i c_{\mu(i,j)}|T_\mu\rangle. \tag{10}$$

Then

$$|\psi\rangle=\textstyle\sum_j b_j|\psi_j\rangle. \tag{11}$$

$|\psi_j\rangle$ is an eigenfunction of **S**, with eigenvalue S_j. The $|\psi_j\rangle$ are normalized and orthogonal. Thus the components $\langle P|\psi_j\rangle$ of a differential-space point P have, as a consequence of the distribution of the ζ_μ, Eq. (3), Gaussianly distributed real and imaginary parts, each with mean square 1, and these quantities are all distributed independently of one another. Thus the polychotomic construction of the function $S(P,\psi)$ can be carried out according to Sec. 5 of reference 6, with the $\langle P|\psi_j\rangle$ as the Gaussianly distributed quantities and the b_j as the fixed coefficients.

Consider the subspace defined by a particular value j' of j, $\{P|S(P,\psi)=S_{j'}\}$. The quantities [see Eq. (2)]

$$\zeta_{\mu(i,j)}=\langle P|T_{\mu(i,j)}\rangle \tag{12}$$

will have (for each i) a certain distribution in this subspace. They are, of course, distributed over the entire space according to Eq. (3). But they are, in fact, distributed in exactly the same way over the subspace (apart from normalization, naturally). This may seem strange, since the polychotomic choice of the subscript j' means that we have imposed a bias that makes $\langle P|\psi_{j'}\rangle$ small relative to the other coordinates of P. This might seem to distort the distribution of the $\zeta_{\mu(i,j')}$, in view of the relation

$$\langle P|\psi_j\rangle=\frac{1}{b_j}\sum_i c_{\mu(i,j)}\zeta_\mu. \tag{13}$$

However, the $\langle P|\psi_j\rangle$ with $j\neq j'$ do not contain the $\zeta_{\mu(i,j')}$; any condition of relative smallness of $\langle P|\psi_{j'}\rangle$ is equivalent to a condition of largeness of the $\langle P|\psi_j\rangle$ for $j\neq j'$, relative to $\langle P|\psi_{j'}\rangle$, and in the latter form it is apparent that such a condition may be regarded solely as one on those $\langle P|\psi_j\rangle$ for which $j\neq j'$; these are independent of the $\zeta_{\mu(i,j')}$, hence the latter are unaffected in their distributions.

Now the vectors $|T_{\mu(i,j')}\rangle$ with fixed j' and all values of i are a complete set for the Hilbert space of the particle. The quantities $\zeta_{\mu(i,j')}$ can therefore be put into one-to-one correspondence, according to the subscript i, with the quantities

$$\xi_i=\langle P'|R_i\rangle, \tag{14}$$

which are the Gaussianly-distributed coordinates of a point P' in the differential-space of the particle; moreover we have just seen that they are also distributed in the same way. We now recall that (reference 5, Sec. 5) the assignment of eigenvalues of **R** to points in the subspace $j=j'$ will take place according to the polychotomic theorem with the use of the distributed quantities $\zeta_{\mu(i,j')}$ and the coefficients $c_{\mu(i,j')}$ of Eq. (6).

We now have everything we need to complete the correspondence between the subspace j' of the differential space of the combined systems and its associated distribution of values of $[\mathbf{R}]$, and the whole differential space of the particle and its associated distribution of values of $[\mathbf{R}]$ obtained from the "reduced" state vector $|\varphi\rangle$ [Eq. (8)]. The quantities $\zeta_{\mu(i,j')}$ and ξ_i of the two spaces respectively are identically distributed, and the

coefficients $c_{\mu(i,j')}$ used in the polychotomic algorithm are the same in both cases. [**R**] may moreover be any observable of the particle. We associate points having equal values of the $\zeta_{\mu(i,j')}$ and the ξ_i in the two spaces pairwise with one another. (This association is invariant to the eigenbasis used, hence the two points that are associated are the same for any [**R**].) A pair of associated points will thus have the same value for any observable. A mapping of the spaces on one another according to values of observables preserves measure; the spaces correspond to identical distributions in all respects.

From the last sentence follows the basic conclusion of this article: If a reading of the instrument after the interaction has ceased corresponds, as it should, to nothing more than the selection of the appropriate subensemble of particles from the over-all ensemble, the subensemble obtained is statistically identical with that obtained from the "reduced" wave function.

It should be noted that the measurement process as a whole is not unaccompanied by a disturbance of the particle, but this occurs during the interaction, not at the moment of the reading of the instrument. At this point, without going into details, we make one remark: In our theory, if the form of the interaction is properly chosen for the purposes of measurement, the disturbance affects only variables other than that being measured. Thus the measurement is only a measurement and nothing more so far as this observable is concerned, and does not force a value on the particle that it did not previously possess; it is a forcing process only so far as other variables, particularly that conjugate to the one being measured, are concerned. In this respect our theory differs concretely from that of Bohm,[8] since in the latter the position (or functions thereof) is the only variable that can ever be measured on an individual system without in general simultaneously altering its value.

In a certain sense, the picture given here furnishes a definition of the "reduction of the wave packet" in terms of everyday concepts. The only proviso is that one must admit the "real existence"[9] of the individual systems represented by points in differential space. (If one is not willing to do so, a mental picture still remains, however unverified.) In our first paper we claimed to have a postulate equivalent to the Born postulate. We might now change this assertion as follows: One introduces a certain dynamical description, a mathematical measure (probability) defined in terms of it, and a method of constructing statistical ensembles, such that *no* new postulate is needed to account for the change in statistical description that results from "taking the pointer reading": One has merely selected a subensemble by separating out systems having a common value of some variable, and the postulate invoked in doing so is at least as old as the theory of statistics—it is so familiar, in fact, that so far as we know no one has ever even bothered to give it a name.

[8] D. Bohm, Phys. Rev. **85**, 180 (1952).

[9] This phrase is of course fraught with epistemological implications. We mean to use it in a quite commonplace sense.