PROBABILITY RELATIONS BETWEEN SEPARATED SYSTEMS

By E. SCHRÖDINGER

[Communicated by Prof. P. A. M. Dirac]

[*Received* 21 April, *read* 26 October 1936]

1. An earlier paper† dealt with the following fact. If for a system which consists of two entirely separated systems the representative (or wave function) is known, then current interpretation of quantum mechanics obliges us to admit *not only* that by suitable measurements, taken on *one* of the two parts only, the state (or representative or wave function) of the *other* part can be determined without interfering with it, *but also* that, in spite of this non-interference, the state arrived at *depends* quite decidedly on *what* measurements one chooses to take—not only on the results they yield. So the experimenter, even with this indirect method, which avoids touching the system itself, controls its future state in very much the same way as is well known in the case of a direct measurement. In this paper it will be shown that the control, with the indirect measurement, is in general not only *as* complete but even more complete. For it will be shown that *in general* a sophisticated experimenter can, by a suitable device which does *not* involve measuring non-commuting variables, produce a non-vanishing probability of driving the system into any state he chooses; whereas with the ordinary direct method at least the states orthogonal to the original one are excluded.

The statement is hardly more than a corollary to a theorem about "mixtures"‡ for which I claim no priority but the permission of deducing it in the following section, for it is certainly not *well* known.

2. Supposing we knew that a system at a given moment were in either one or other of the sequence of states corresponding to the following sequence of wave functions ϕ_i, finite or infinite in number, normalized but in general not orthogonal; supposing further we knew the probabilities of the system being in any one of these states—they must, of course, sum up to unity and shall be the real positive numbers w_i, written in the second line below the symbol of the function or state:

$$\left.\begin{array}{ccccc} \phi_1 & \phi_2 & \phi_3 & \phi_4 & \cdots \\ w_1 & w_2 & w_3 & w_4 & \cdots \\ (c'_{1k}) & (c'_{2k}) & (c'_{3k}) & (c'_{4k}) & \cdots \end{array}\right\}. \tag{1}$$

† *Proc. Camb. Phil. Soc.* 31 (1935), 555–63.

‡ The valuable conception of a mixture and the appropriate way of handling a mixture by the Statistical Operator is due to Johann von Neumann; see his *Mathematische Grundlagen der Quantenmechanik*, Berlin, Springer, 1932; especially pp. 225ff.

In the third line I have written the kth development coefficient of the function above, with respect to an arbitrary complete orthogonal system, chosen as a frame of reference to start with. The brackets are to indicate that the kth is meant as a representative of all of them; (c'_{3k}), for instance, means all the development coefficients of ϕ_3.

The mean value (or expectation value) of a physical variable A, represented in our frame of reference by the matrix A_{lk}, is

$$\bar{A} = w_1 \sum_{l,k} c'^{*}_{1l} A_{lk} c'_{1k} + w_2 \sum_{l,k} c'^{*}_{2l} A_{lk} c'_{2k} + \dots$$

$$= \sum_{l,k} A_{lk} \sum_{i} w_i c'_{ik} c'^{*}_{il}.$$

Let us call U the object, represented by the matrix whose (k, l)th element is

$$U_{kl} = \sum_{i} w_i c'_{ik} c'^{*}_{il}, \tag{2}$$

then

$$\bar{A} = \text{Trace}\,(AU). \tag{3}$$

This trace (i.e. sum of diagonal terms) is obviously independent of the frame of reference, since the w_i are, by their meaning, invariants and U therefore transforms like a matrix representing a physical variable, e.g. like A.

Since mean values are all that quantum mechanics predicts at all, the knowledge of U in a definite frame of reference *exhausts* our real knowledge of the situation, just as in the case of a "state" the wave function exhausts it. The detailed times (1) from which U_{kl} is composed may refer to the origin of our knowledge. But if another set of similar times leads to the same U, then it would be entirely meaningless to distinguish between the two physical situations. U is von Neumann's Statistical Operator. Its matrix is hermitian. It has the formal character of a real physical variable, but the physical meaning of a wave function, that is to say it describes the instantaneous physical situation of the system.

We propose to find *all* the different ways (or detailed data like (1)) which lead to the *same* U. Mark first, that the hermitian U_{kl} is composed linearly, with positive coefficients w_i, whose sum total is unity, from matrices each of which obviously has the eigenvalues 1, 0, 0, 0, ..., and therefore the trace 1. It follows that U_{kl} has itself the trace 1 and non-negative eigenvalues. Now let us change the frame of reference by a canonical transformation, which makes U diagonal. Let

$$p_1, p_2, p_3, p_4, \dots, \tag{4}$$

be this diagonal, the eigenvalues of U; according to what has just been said, they are non-negative and their sum is 1. This shows that the *same* mixture is obtained by mixing *the orthogonal functions* (or states) which form the basis of the new frame of reference, in the proportions (or with the probabilities) p_k.

Now let c_{ik} (*without* a dash) be the kth development coefficient of ϕ_i in the

https://doi.org/10.1017/S0305004100019137 Published online by Cambridge University Press

new frame. That the statistical matrix in the new frame actually *is* the diagonal (4) is expressed by

$$\sum_i w_i c_{ik} c_{il}^* = \delta_{kl} \sqrt{(p_k p_l)}. \tag{5}$$

Taking first $k=l$, then since the w_i are positive, we see that $c_{ik}=0$ whenever $p_k=0$. So the basic states corresponding to zero eigenvalues of U are not needed for the expansion of any function which can appear as a constituent of our mixture. It is convenient to drop them altogether, *confining oneself to that subspace* of Hilbert space which is spanned by the basic functions with non-vanishing eigenvalue with respect to U. It then follows that the quantities

$$c_{ik} \sqrt{\left(\frac{w_i}{p_k}\right)} = \alpha_{ik} \tag{6}$$

form *columns* (labelled by k) of a unitary matrix or, let us say, "unitary columns", meaning just only that each of them is normalized in itself and orthogonal to all the others. (More cannot be asserted, because we did not oblige the ϕ_i to be linearly independent.) Since the c_{ik} are normalized we must have

$$1 = \sum_k |c_{ik}|^2 = \frac{1}{w_i} \sum_k p_k |\alpha_{ik}|^2$$

or

$$w_i = \sum_k p_k |\alpha_{ik}|^2 \tag{7}$$

and

$$c_{ik} = \frac{\alpha_{ik} \sqrt{p_k}}{\sqrt{(\sum_l p_l |\alpha_{il}|^2)}}. \tag{8}$$

Obviously *any* sequence of functions which produce the same mixture must be given by (8) (and their mixing fractions by (7)), with α_{ik} certain unitary columns. On the other hand (8) and (7) with any unitary columns inserted will produce as their statistical matrix the diagonal (4) (expurgated from zeros). Therefore the answer to our question is:

The most general way of obtaining the same mixture as indicated by the diagonal (4) *is given by* (8) *and* (7) *with* α_{ik} *any unitary columns.* Note: There have to be as many columns (k) as there are basic states in the subspace retained, thus one for every non-vanishing p_k. The range of the line subscript (i) is arbitrary except that it must suffice to form that many normalized and mutually orthogonal columns; but the range of i is allowed to be larger than that.

Can (with a given U) an arbitrary state be made to enter the circle of composing states?

Well, as we have seen, certainly only a state which belongs to our subspace. In order to show that there is no further restriction, let us *now* take for α_{ik} a true *unitary matrix*. This, by the way, has the effect of leading to *linearly independent* constituents. Indeed if there were a relation

$$\sum_i \lambda_i c_{ik} = 0 \quad (k = 1, 2, 3, \ldots)$$

https://doi.org/10.1017/S0305004100019137 Published online by Cambridge University Press

that would, from (6), imply a similar relation

$$\sum_i \mu_i \alpha_{ik} = 0 \quad (k = 1, 2, 3, \ldots)$$

which cannot hold for a unitary matrix.

Now from (6)
$$1 = \sum_k |\alpha_{ik}|^2 = w_i \sum_k \frac{|c_{ik}|^2}{p_k},$$

and therefore
$$w_i = \left(\sum_l \frac{|c_{il}|^2}{p_l}\right)^{-1}, \tag{9}$$

$$\alpha_{ik} = \frac{c_{ik}}{\sqrt{(p_k)}} \left(\sum_l \frac{|c_{il}|^2}{p_l}\right)^{-\frac{1}{2}}. \tag{10}$$

If you take here *one* function, say the $i=j$th, entirely arbitrary, you get a *line* α_{jk} $(k = 1, 2, 3, 4, \ldots)$, which is evidently fit to be taken as a line of a unitary matrix, leaving the other lines still widely indeterminate. It is interesting to observe that nevertheless the probability fraction with which a function contributes to a given mixture, when regarded as one of a possible set of linearly independent constituents, depends only on the mixture (i.e. on U) and on the function itself. One is allowed to speak in a very general way of the fraction by which an arbitrary state contributes to a given mixture—but of course the sum total of *all* these fractions is by no means unity.

It will be remembered that in the formulae (10), vanishing p_l have been excluded. But the formula for w_j allows us to follow the vanishing of w_j, if we let one of the p_l *tend* to zero and the corresponding basic state does not happen to be absent in that particular function ($c_{jl} \neq 0$). Moreover, the expression of w_j must have an invariant form, independent of the frame of reference. And so it has. In a familiar symbolism
$$w = (U^{-1}\phi, \phi)^{-1}. \tag{11}$$
(U must refer to the *sub*space, in order that U^{-1} shall exist.) Formally the reciprocal probability for an arbitrary function is calculated as if it were the mean value of a physical variable under that function, the physical variable being represented by the reciprocal of the Statistical Operator.

We sum up the main result: any state which belongs to the subspace spanned by the really contributing orthogonal constituents can be regarded as a constituent, with a probability (or mixing fraction) that only depends on that state itself provided the constituents are chosen linearly independent.

3. It is just the multitude of non-orthogonal compositions of one and the same mixture, which gives rise to the strange situation described in the beginning of section 1. As in eq. 1, p. 556, *loc. cit.*, let x stand for all the coordinates of the first, y for those of the second system. It was shown that the wave function of the combined system can be given the form
$$\Psi(x, y) = \sum_k a_k g_k(x) f_k(y) \tag{12}$$
with *both* sets, the $g_k(x)$ and the $f_k(y)$ normalized *and* orthogonal. They need not

form *complete* sets, unless we admit dummy terms with $a_k = 0$. This we will not. We prefer to confine all the following considerations to the two subspaces of Hilbert space which are spanned by the $g_k(x)$ for the first system, by the $f_k(y)$ for the second system.

Consider the *possibility* of measuring in the second system a physical variable of which the $f_k(y)$ are eigenfunctions belonging to *different* eigenvalues. The probability of finding the one that belongs to $f_k(y)$ is

$$p_k = |a_k|^2, \tag{13}$$

and *if* we find it, we have to assign to the *first* system the wave function $g_k(x)$. In view of the *possibility* of following that particular measuring device, with all its possible outcomes, the first system is in the situation of a mixture characterized by the orthogonal functions $g_k(x)$ and the probabilities p_k. From the preceding section we know that such a mixture can be regarded as the result of extremely various compositions. If we restrict this variety to those compositions which consist of linearly independent constituents, then it corresponds exactly to the variety of different measuring programmes (of commuting variables) which can be carried out on the *second* system. Envisage the unitary transformation

$$f_k(y) = \sum_i \alpha'_{ik} h_i(y), \tag{14}$$

and a physical variable of which the $h_i(y)$ are eigenstates belonging to different eigenvalues. From (12) and (14)

$$\begin{aligned}\Psi(x,y) &= \sum_k \sum_i a_k g_k(x)\, \alpha'_{ik} h_i(y) \\ &= \sum_i \Big(\sum_k a_k \alpha'_{ik} g_k(x)\Big) h_i(y).\end{aligned}$$

The functions of x described by the $\sum_k$ lack normalization. Besides from (13)

$$a_k = \sqrt{p_k}\, e^{i\phi k} \quad (\phi_k \text{ real})$$

and we can put

$$\alpha'_{ik} e^{i\phi k} = \alpha_{ik}$$

with α_{ik} another unitary matrix, which is arbitrary because α'_{ik} was arbitrary. So we can write

$$\Psi(x,y) = \sum_i \sqrt{(w_i)} \left(\sum_k \frac{\alpha_{ik} \sqrt{(p_k)}}{\sqrt{w_i}} g_k(x) \right) h_i(y) \tag{15}$$

with

$$w_i = \sum_l p_l |\alpha_{il}|^2.$$

Then the function in the bracket is normalized. It is precisely the wave function which we have to assign to the first system, if we find in the second one the eigenvalue belonging to $h_i(y)$, and the probability of finding it is w_i.

Comparing with (7) and (8) of the preceding section we find complete coincidence; from which we infer that an arbitrary state within the subspace in question can be given a non-vanishing probability of its turning up in the *first*

https://doi.org/10.1017/S0305004100019137 Published online by Cambridge University Press

system by a suitable treatment of the second one. Since it has a finite chance of turning up, it will certainly turn up, if precisely the same experiments are repeated sufficiently often. Moreover, quite apart from special applications, the case that in the expansion (12) *no* coefficients vanish deserves to be called the general one. Then there is no reduction to a subspace and what has just been said holds for an altogether arbitrary state. That was the meaning of the announcement made in section 1.

4. Indubitably the situation described here is, in present quantum mechanics, a necessary and indispensable feature. The question arises, whether it is so in Nature too. I am not satisfied about there being sufficient experimental evidence for that. Years ago I pointed out† that when two systems separate far enough to make it possible to experiment on one of them without interfering with the other, they are bound to pass, during the process of separation, through stages which were beyond the range of quantum mechanics as it stood then. For it seems hard to imagine a complete separation, whilst the systems are still so close to each other, that, from the classical point of view, their interaction could still be described as an unretarded *actio in distans*. And ordinary quantum mechanics, on account of its thoroughly unrelativistic character, really only deals with the *actio in distans* case. The whole system (comprising in our case *both* systems) has to be small enough to be able to neglect the time that light takes to travel across the system, compared with such periods of the system as are essentially involved in the changes that take place.

Though in the mean time some progress seemed to have been made in the way of coping with this condition (quantum electrodynamics), there now appears to be a strong probability (as P. A. M. Dirac‡ has recently pointed out on a special occasion) that this progress is futile.

It seems worth noticing that the paradox could be avoided by a very simple assumption, namely if the situation after separating were described by the expansion (12), but with the additional statement that the knowledge of the *phase relations* between the complex constants a_k has been entirely lost in consequence of the process of separation. This would mean that not only the parts, but the whole system, would be in the situation of a mixture, not of a pure state. It would not preclude the possibility of determining the state of the first system by *suitable* measurements in the second one or *vice versa*. But it would utterly eliminate the experimenters influence on the state of that system which he does not touch.

This is a very incomplete description and I would not stand for its adequateness. But I would call it a possible one, until I am told, either why it is devoid of

† E. Schrödinger, *Annalen der Physik* (4), 83 (1927), 961. *Collected Papers* (Blackie and Son, 1928), p. 141.

‡ P. A. M. Dirac, *Nature*. 137 (1936), 298.

meaning or with which experiments it disagrees. My point is, that in a domain which the present theory does not cover, there is room for new assumptions without necessarily contradicting the theory in that region where it is backed by experiment.

SUMMARY

The paper first scrutinizes thoroughly the variety of compositions which lead to the same quantum-mechanical *mixture* (as opposed to *state* or *pure state*). With respect to a given mixture *every* state has a definite probability (or mixing fraction) between 0 and 1 (including the limits), which is calculated from the mixtures Statistical Operator and the wave function of the state in question.

A well-known example of mixtures occurs when a system consists of two separated parts. If the wave function of the whole system is known, either part is in the situation of a mixture, which is decomposed into *definite* constituents by a *definite* measuring programme to be carried out on the *other* part. All the conceivable decompositions (into linearly independent constituents) of the first system are just realized by all the possible measuring programmes that can be carried out on the second one. *In general every* state of the first system can be given a finite chance by a suitable choice of the programme.

It is suggested that these conclusions, unavoidable within the present theory but repugnant to some physicists including the author, are caused by applying non-relativistic quantum mechanics beyond its legitimate range. An alternative possibility is indicated.

I wish to acknowledge with gratitude the support of my studies by Imperial Chemical Industries, Limited.