CONDITIONAL PROBABILITIES IN NON-BOOLEAN POSSIBILITY STRUCTURES

I

The thesis that the transition from classical to quantum mechanics is to be understood as the transition from a Boolean to a non-Boolean possibility structure of events raises several problems concerning the representation and interpretation of probabilities, since classical probability theory is essentially a Boolean theory. The problem I want to consider here concerns the interpretation of conditional probabilities relating incompatible properties.

To illustrate, consider a system associated with a 2-dimensional Hilbert space. I denote magnitudes of the system by A, B, . . ., possible values by a_1, a_2; b_1, b_2; . . ., and corresponding eigenvectors by α_1, α_2; β_1, β_2; . . . . Suppose the system is represented by the vector α_1 (i.e., statistical operator P_{α_1}), assigning unit probability to a_1 and zero probability to a_2. The statistical operator P_{α_1} assigns probability, $p_{\alpha_1}(b_1) = \mathrm{Tr}(P_{\alpha_1}P_{b_1}) = |(\beta_1, \alpha_1)|^2$ to the property b_1, and probability $p_{\alpha_1}(b_2) = \mathrm{Tr}(P_{\alpha_1}P_{b_2}) = |(\beta_2, \alpha_1)|^2$ to the property b_2. *Informally*, these probabilities are understood as conditional probabilities in *some* sense, i.e., the probability $p_{\alpha_1}(b_i)$ is understood as the probability of the property b_i conditional on the property a_1. But, evidently the probability assigned to b_i when the system is known to have the property a_1 is not a conditional probability in the sense of a probability proportional to the joint probability of a_1 and b_i – such a joint probability does not exist.

For example, α_1 might represent a spin eigenstate associated with 'spin up in the direction **a**'. Then $p_{\alpha_1}(b_1)$ is the probability assigned to the property 'spin up in the direction **b**', given that the spin of the system is up in the direction **a**; and $p_{\alpha_1}(b_2)$ is the probability of the property 'spin down in the direction **b**', given that the spin of the system is up in the direction **a**. I do not think that this problem of interpretation disappears if we simply agree to speak in the conventional idiom of 'the probability of finding spin up in the direction **b** if a measurement is made on a system in the state α_1'. For the property 'spin up in the direction **a**' is surely presumed to obtain if the system is in the quantum state α_1, and the probability in question is now presented as a rather more complicated conditional probability: the probability of finding spin up in the direction **b** conditional on (i) the property 'spin up in the direction **a**' for the system, and (ii) an appropriate spin measurement. However the notion of measurement is construed here, we still have the problem of understanding the sense in which this is a conditional probability, i.e., the sense in which stipulating 'spin up in the direction **a**' for the system, together with an appropriate measurement, conditionalizes an initial probability assignment to yield the probabilities $p_{\alpha_1}(b_1)$, $p_{\alpha_1}(b_2)$ for the alternative spins in the **b** direction.

The comment that the probability $p_{\alpha_1}(b_i)$ cannot be represented as a conditional probability in the usual sense (i.e., proportional to the joint probability of a_1 and b_i) requires some elaboration. Ultimately the non-existence of hidden variables underlying the quantum statistics depends on nothing more than this elementary feature of the theory.

The question at issue is whether the probabilities of the 2-dimensional case are representable as *conditional* probabilities on a classical probability space $(X, \mathscr{F}, \mu)$. Let X_{a_1}, X_{a_2} be two mutually exclusive and collectively exhaustive subsets of X, which partition X into two regions associated respectively with the two possible values a_1 and a_2 of the magnitude A. Similarly, partition X into X_{b_1} and X_{b_2} for the magnitude B, and so on. The sets X_{a_1}, X_{a_2}; X_{b_1}, X_{b_2}; etc. generate the field $\mathscr{F}$. The problem is whether there exists an initial measure μ assigning equal probabilities to any pair of alternative properties a_1, a_2 or b_1, b_2, etc. (i.e., $\mu(X_{a_1}) = \mu(X_{a_2}) = \frac{1}{2}$; $\mu(X_{b_1}) = \mu(X_{b_2}) = \frac{1}{2}$, etc.), with respect to which the family of probabilities $p_{\alpha_i}(b_j)$, $p_{\beta_i}(a_j)$ for all i, j, and all magnitudes A, B, . . ., can be represented as conditional probabilities in the usual sense, i.e.,

$$p_{\alpha_i}(b_j) = |(\beta_j, \alpha_i)|^2 = \frac{\mu(X_{\alpha_i} \cap X_{b_j})}{\mu(X_{\alpha_i})}$$

Now, for *any* pair of magnitudes A, B there exists a measure μ such that

$$\mu(X_{a_i} \cap X_{b_j}) = \frac{|(\alpha_i, \beta_j)|^2}{2}$$

(Notice that $|(\alpha_i, \beta_j)|^2 = |(\beta_j, \alpha_i)|^2$.)

Since

$$X_{a_i} = X_{a_i} \cap (X_{b_1} \cup X_{b_2}) = (X_{a_i} \cap X_{b_1}) \cup (X_{a_i} \cap X_{b_2})$$

it follows that

$$\mu(X_{a_i}) = \frac{|(\alpha_i, \beta_1)|^2}{2} + \frac{|(\alpha_i, \beta_2)|^2}{2} = \frac{1}{2}$$

(assuming α_i is normalized). Similarly, $\mu(X_{b_j}) = \frac{1}{2}$. Thus, it *is* possible to represent the probabilities $p_{\alpha_i}(b_j)$, $p_{\beta_i}(\alpha_j)$ $(i, j = 1, 2)$ as conditional probabilities in the usual sense:

$$p_{\alpha_i}(b_j) = |(\beta_j, \alpha_i)|^2 = \frac{\mu(X_{a_i} \cap X_{b_j})}{\mu(X_{a_i})}$$

$$p_{\beta_i}(a_j) = |(\alpha_j, \beta_i)|^2 = \frac{\mu(X_{b_i} \cap X_{a_j})}{\mu(X_{a_j})}$$

if we consider only a *single pair* of magnitudes A and B.

This measure cannot, however, be extended to three magnitudes A, B, C. It is easy to see that any measure satisfying the required conditions must violate what might well be termed a 'condition of coherence'. For any subsets X_s, X_t, X_u of X any measure μ, we have:

$$\begin{aligned}
\mu(X_s \cap X_t) &= \mu(X_s \cap X_t \cap X_u) + \mu(X_s \cap X_t \cap X'_u)\\
\mu(X_s \cap X_u) &= \mu(X_s \cap X_u \cap X_t) + \mu(X_s \cap X_u \cap X'_t)\\
\mu(X_t \cap X'_u) &= \mu(X_t \cap X'_u \cap X_s) + \mu(X_t \cap X'_u \cap X'_s)
\end{aligned}$$

and so

$$\mu(X_s \cap X_t) \leqslant \mu(X_s \cap X_u) + \mu(X_t \cap X'_u).$$

Hence, taking $s = a_2$, $t = c_1$, $u = b_1$, it follows that

$$\mu(X_{a_2} \cap X_{c_1}) \leqslant \mu(X_{a_2} \cap X_{b_1}) + \mu(X_{c_1} \cap X_{b_1}')$$

i.e.

$$\mu(X_{a_2} \cap X_{c_1}) \leqslant \mu(X_{a_2} \cap X_{b_1}) + \mu(X_{c_1} \cap X_{b_2}).$$

But this condition cannot be satisfied in general for all triples of quantum mechanical magnitudes A, B, C in $\mathscr{H}_2$, i.e., there exist vectors α_2, β_1, β_2, γ_1 such that

$$|(\gamma_1, \alpha_2)|^2 \nleqslant |(\beta_1, \alpha_2)|^2 + |(\gamma_1, \beta_2)|^2$$

or, alternatively, there exist directions **a**, **b**, **c** in real space (corresponding to the directions of the spin magnitudes A, B, C) such that

$$\sin^2 \tfrac{1}{2}\theta_{\mathbf{ac}} \nleqslant \sin^2 \tfrac{1}{2}\theta_{\mathbf{ab}} + \sin^2 \tfrac{1}{2}\theta_{\mathbf{bc}}$$

(where $\theta_{\mathbf{ac}}$ is the angle between **a** and **c**, etc.).

From this standpoint the central interpretative problem of quantum mechanics may be expressed as follows: The probabilities associated with pairs of properties in quantum mechanics (e.g., the probability of spin up in the direction **b**, given spin up in the direction **a**, etc.) cannot be represented as conditional probabilities on a classical probability space. Yet, *in some sense*, a probability like $p_{\alpha_i}(b_j)$ is to be understood as the conditional probability of b_j given a_i. How do we make sense of this probability as a conditional probability?

II

I want to sketch a representation of classical probability theory as an operator calculus, analogous to the operator calculus of quantum mechanics, and show that the classical conditionalization rule in this calculus is just von Neumann's projection postulate (actually, a corrected form of this postulate first proposed by Lüders).[1] This construction will lend support to the claim that von Neumann's projection postulate (more correctly, the Lüders rule) is the appropriate rule for conditionalizing probabilities in the non-Boolean possibility structure of quantum mechanics. The sense in which $p_{\alpha_i}(b_j)$ is the conditional probability of b_j given a_i is just this: $p_{\alpha_i}(b_j)$ is the probability assigned to b_j by the application of the Lüders conditionalization rule to the *a priori* equiprobable measure represented by the normalized unit statistical operator $I/\mathrm{Tr}(I)$.

Consider, for simplicity, a countable classical probability space $(X, \mathscr{F}, \mu)$. I shall label the atomic events or elementary possibilities by $x_1, x_2, \ldots$. These are associated with singleton subsets $X_1, X_2, \ldots$, or indicator functions (characteristic functions) $I_1, I_2, \ldots$. I shall label other, possibly non-atomic, events by $a, b, \ldots$. Thus, the set $a_1, a_2, \ldots$ might denote a set of non-atomic mutually exclusive and collectively exhaustive events ($\sum_i I_{a_i} = I$; $I_{a_i}I_{a_j} = 0$, $i \neq j$).

Now, for any probability measure μ, it is possible to introduce a 'statistical operator' $W = \sum_i w_i I_i$, where $\sum_i w_i = 1$, $w_i \geqslant 0$, for all i, in terms of which the probability of an event a may be represented as:

$$p_\mu(a) = \mu(X_a) = \sum_j \Big(\sum_i w_i I_i(x_j) \Big) I_a(x_j)$$

I shall write $p_W(a)$ for $p_\mu(a)$, where W corresponds to μ, i.e.

$$p_W(a) = \sum_j W(x_j) I_a(x_j).$$

To simplify notation, I shall abbreviate this expression as

$$p_W(a) = \sum W I_a,$$

where a summation sign without an index is understood as summing over all the atomic events x_j. This convention will be used below.

In the terms of the statistical operator, the conditional probability (relative to an initial measure μ associated with the statistical operator W) of an event b given an event a_i, may be represented as:

$$p_W(b \mid a_i) = \frac{\sum W I_{a_i} I_b}{\sum I_{a_i}}$$

$$\left(\text{i.e.,} \quad p_W(b \mid a_i) = \frac{\sum_i W(x_j) I_i(x_j) I_b(x_j)}{\sum_i W(x_j) I_i(x_j)} \right).$$

To see this, simply notice that

$$p_\mu(b \mid a_i) = \frac{\mu(X_{a_i} \cap X_b)}{\mu(X_{a_i})}$$

$$= \frac{\sum W I_{a_i} I_b}{\sum W I_{a_i}}$$

Thus, the transition

$$\mu \to \mu'$$

on conditionalization with respect to a_i (where μ' is defined by

$$\mu'(X_e) = \frac{\mu(X_{a_i} \cap X_e)}{\mu(X_{a_i})} = p_\mu(e \mid a_i) \text{ for any event } e)$$

may be represented as the transition

$$W \to W' = \frac{WI_{a_i}}{\sum WI_{a_i}}$$

so that

$$p_W(b \mid a_i) = \sum W'I_b.$$

The statistical operator construction allows the replacement of the measure function μ, which is a set function whose domain is the field of measurable subsets of X, by a corresponding random variable W, a point function whose domain is X. If we regard the probability space as associated with a physical system with magnitudes A, B, etc., whose possible values $a_1, a_2, \ldots; b_1, b_2, \ldots$, etc. correspond to the possible events represented by the field $\mathscr{F}$, then the statistics of this system is now represented by a physical magnitude W belonging to the algebra of magnitudes of the system. The advantage of this construction is that it provides a purely algebraic way of representing the statistics of a system, which is appropriate whether or not a representation of the algebra of magnitudes as real-valued functions on a space is possible. I want to suggest that we take W as representative of the statistics in a primary sense – the measure function μ exists only if the algebra of magnitudes is commutative. In this special case, the subalgebra of indempotent magnitudes forms a Boolean algebra, which has a representation as a field of subsets of a set, by Stone's theorem. The measure function defined as a set function on this field is essentially the 'Stone representative' of the statistical operator W, which is the element in the algebra of magnitudes incorporating the statistics. Bearing in mind the possibility of non-commutative algebras of magnitudes (i.e., non-Boolean possibility structures), it seems appropriate to represent the transition corresponding to conditionalization with respect to an event a_i by the symmetrical expression:

$$W \to W' = \frac{\sum I_{a_i} W I_{a_i}}{I_{a_i} W I_{a_i}}.$$

Now, *this is just von Neumann's projection postulate* (in the corrected Lüders version). In quantum mechanics, the statistics of a system is represented by a statistical operator W which may be represented as:

$$W = \sum_i w_i P_{\alpha_i}$$

where P_{α_i} are projection operators onto atomic events (i.e., projection operators onto 1-dimensional subspaces spanned by the vectors α_i). In terms of this operator, the probability of an event b may be represented as:

$$p_W(b) = \mathrm{Tr}(WP_b).$$

Notice that the trace of an operator O is just the sum of the eigenvalues of O, i.e., the sum of the possible values of O at each atom in the maximal Boolean subalgebra defined by O. Thus, the operation Tr in the algebra of operators of a quantum mechanical system is completely analogous to the operation Σ in the commutative algebra of magnitudes considered above.

The conditional probability relative to an initial measure associated with the statistical operator (W) of an event b given an event c_i is:

$$p_W(b \mid c_i) = \mathrm{Tr}(W'P_b)$$

where

$$W' = \frac{P_{c_i}WP_{c_i}}{\mathrm{Tr}(P_{c_i}WP_{c_i})}$$

This expression is due to Lüders.[2] In the following section, I shall deal in detail with the relation between the Lüders rule and von Neumann's rule. Here I wish to point out the following: If we assume an *a priori* probability assignment given by the statistical operator $W = I/\mathrm{Tr}(I)$, representing an equiprobable initial distribution over every complete set of orthogonal atomic properties (associated with the possible values of a maximal magnitude), then the conditionalization with respect to an atomic property c_i yields the transition

$$W \to W' = P_{c_i}$$

where P_{c_i} is the projection operator onto the 1-dimensional subspace spanned by the eigenvector γ_i, say, corresponding to c_i. This means that the probability of a property b conditional on c_i (where b may be incompatible with c_i) is to be computed according to the rule:

$$p_W(b \mid c_i) = \mathrm{Tr}(P_{c_i}P_b).$$

If b is atomic, corresponding to the vector β, we have

$$p_W(b \mid c_j) = |(\beta, \gamma_j)|^2.$$

Thus, the probability assigned by the 'state vector' γ_j (representing the association of the property c_j with the system) to an incompatible property b, according to the quantum mechanical rule, may be interpreted as the conditional probability of the property b given the property c_j relative to an initial probability distribution which is equiprobable with respect to every complete set of atomic properties of the system.

III

Von Neumann introduces the rule which has become known as the 'projection postulate' in Section 3 of Chapter III of his book *Mathematical Foundations of Quantum Mechanics*. In its simplest form, the postulate states that if a measurement of a maximal magnitude A with eigenvalues $a_1, a_2, \ldots$ and corresponding eigenvectors $\alpha_1, \alpha_2, \ldots$ yields the result a_i, then the initial quantum state of the system is transformed to the state α_i. Von Neumann goes on to consider the case of a non-maximal measurement. If the eigenvalue a_i has multiplicity k_i, then the corresponding eigenvectors span a k_i-dimensional subspace $\mathscr{K}_{a_i}$, the range of a projection operator P_{a_i}. Von Neumann argues that after a measurement yielding the result a_i, the system is represented by the statistical operator

$$\frac{P_{a_i}}{\mathrm{Tr}(P_{a_i})} = \frac{P_{a_i}}{k_i}.$$

Note that this represents a mixture, not a pure state. In the general case of a magnitude A represented by an operator with a continuous spectrum, he concludes that after a measurement yielding the result $a \in S$, relative probabilities are generated by the unnormalizable statistical operator $P_A(S)$, where $P_A(S)$ is the projection operator in the spectral measure of A corresponding to the range S $(A = \int r \, \mathrm{d}P_A(r))$.

Now, quite apart from any objections to a measurement postulate of this sort, it is generally agreed that von Neumann's rule can only be correct for *maximal* measurements. The accepted rule was first proposed by Lüders.[3] The Lüders rule states that a (possibly non-maximal) measurement of a magnitude A yielding the result a_i leads to the following transition in the statistical operator W of the system:

$$W \to W' = \frac{P_{a_i} W P_{a_i}}{\mathrm{Tr}(P_{a_i} W P_{a_i})} \qquad \text{(Lüders)}$$

and *not*

$$W \to W' = \frac{P_{a_i}}{\mathrm{Tr}(P_{a_i})} \qquad \text{(von Neumann)}$$

What is the difference between these two rules? The two rules agree in only two cases: (i) if $W = I/\mathrm{Tr}(I)$, where I is the unit operator, and (ii) for maximal measurements, i.e., when each P_{a_i} is the projection operator onto a (different) 1-dimensional subspace spanned by the vector α_i.

Case (i) is immediately obvious: $W \to W' = P_{a_i}/\mathrm{Tr}(P_{a_i})$. For case (ii), the von Neumann rule yields:

$$W \to W' = P_{\alpha_i},$$

for the transition corresponding to the result a_i. The Lüders rule yields this transition too, by Lemma 2 (see Appendix):

$$W \to W' = \frac{P_{\alpha_i} W P_{\alpha_i}}{\mathrm{Tr}(P_{\alpha_i} W P_{\alpha_i})} = P_{\alpha_i}.$$

To bring out the difference between the Lüders rule and the von Neumann rule, consider an initial pure statistical operator $W = P_\psi$ and a non-maximal measurement, i.e., where the projection operators P_{a_i} are not in general 1-dimensional.

The Lüders rule yields

$$W = P_\psi \to W' = \frac{P_{a_i} P_\psi P_{a_i}}{\mathrm{Tr}(P_{a_i} P_\psi P_{a_i})}.$$

Let $P_{a_i}\psi = \theta_i$, then for any vector ϕ:

$$\begin{aligned} P_{a_i} P_\psi P_{a_i}\phi &= P_{a_i} P_\psi (P_{a_i}\phi) \\ &= P_{a_i}(\psi, P_{a_i}\phi)\psi \\ &= (P_{a_i}\psi, \phi)\theta_i \\ &= (\theta_i \phi)\theta_i \\ &= P_{\theta_i}\phi \end{aligned}$$

i.e.

$$P_{a_i} P_\psi P_{a_i} = P_{\theta_i}.$$

Now,

$$\mathrm{Tr}(P_{a_i} P_\psi P_{a_i}) = \|P_{a_i}\psi\|^2 = \|\theta_i\|^2,$$

and so

$$W' = \frac{P_{\theta_i}}{\|\theta_i\|^2} = P_{\theta'_i}$$

where

$$\theta'_i = \frac{\theta_i}{\|\theta_i\|} \qquad \text{(by Lemma 5).}$$

The von Neumann rule yields:

$$W = P_\psi \to W' = \frac{P_{a_i}}{\mathrm{Tr}(P_{a_i})}$$

According to the von Neumann rule, W' is a *mixture which does not depend on the initial quantum state* ψ. According to the Lüders rule, W' is a *pure state* θ'_i, *which does depend on the initial state* ψ. In fact, θ'_i is the normalized projection of ψ onto the subspace which is the range of P_{a_i}, i.e., $\theta'_i = P_{a_i}\psi/\|P_{a_i}\psi\|$.

IV

In the previous section, I contrasted two proposed rules for a transition in the state of a quantum mechanical system following a measurement on the system. On the basis of the analysis in Section II, I want to suggest that the Lüders rule is to be understood as the quantum mechanical rule for conditionalizing an initial probability assignment (specified by a statistical operator) with respect to an element in the non-Boolean possibility structure of the theory. It is crucial to this conception that we understand the quantum mechanical specification of a system by its state vector as a *statistical* specification, i.e., the state vector ψ determines a statistical operator P_ψ, which is to be understood in the sense of Section II as the algebraic counterpart of the classical measure function μ (which, of course, does not exist in this case, since the possibility structure is non-Boolean).

I shall illustrate this conception of the 'quantum state' (pure or mixed) of a system, and the associated interpretation of the Lüders projection postulate as the conditionalization rule, by analyzing the 2-slit experiment as a problem in conditional probabilities. I shall show that the von Neumann rule actually gives the wrong result – no interference – while the Lüders rule interpreted as a conditionalization rule gives the correct result.

We have a screen with two slits, A and B, and a second detecting screen or photographic plate. A photon in a pure quantum state represented by a plane wave moves towards the slits. Each slit can be regarded as localizing the photon to a region, Δ_A or Δ_B, in the plane of the slit screen. In other words, there is a magnitude M, representing position in the slit screen plane, and the passage of a particle through a slit is a measurement of the magnitude M, in the sense that a range, Δ_A or Δ_B, is assigned to M for the photon at the time of passage. We are interested in the probability that the photon will arrive at a certain region on the detecting screen, conditional on localization to a certain range of values of M. Localization to a region Δ on the detecting screen is a measurement of a magnitude N, representing position in the detecting screen plane. N may be taken as the same magnitude M, if the regions Δ are the same size as the slits, or at least as compatible with M otherwise.

Suppose slit A is open and slit B closed. What is the probability that the photon will be found in the region Δ on the detecting screen? According to the von Neumann conditionalization rule, the conditionalized statistical operator for the photon, yielding the photon statistics immediately after the photon has passed through slit A, is

$$W_A = \frac{P_M(\Delta_A)}{\mathrm{Tr}(P_M(\Delta_A))},$$

where $P_M(\Delta_A)$ is the projection operator in the spectral measure of M corresponding to the range Δ_A. The probability that the photon will arrive at region Δ on the detecting screen after a travel time t is:

$$p_{W_A}(n \in \Delta) = \mathrm{Tr}(U_t^{-1} W_A U_t P_N(\Delta)),$$

where U_t is the unitary time transformation associated with the photon's motion.

With slit B open and slit A closed, we have

$$p_{W_B}(n \in \Delta) = \mathrm{Tr}(U_t^{-1} W_B U_t P_W(\Delta)).$$

With both slits open, the conditionalized statistical operator for the photon immediately after the photon has passed through the slits is

$$W_{AB} = \frac{P_M(\Delta_A) + P_M(\Delta_B)}{\mathrm{Tr}(P_M(\Delta_A) + P_M(\Delta_B))},$$

since $P_M(\Delta_A) + P_M(\Delta_B)$ is the projection operator in the spectral measure of M corresponding to the range $\Delta_A \cup \Delta_B$. (Note that this depends on the disjointness of Δ_A and Δ_B, i.e., on the orthogonality of $P_M(\Delta_A)$ and $P_M(\Delta_B)$.) Assuming $\Delta_A = \Delta_B$, i.e., that intervals are equal in length, and hence

$$\mathrm{Tr}(P_M(\Delta_A)) = \mathrm{Tr}(P_M(\Delta_B)) = k$$

we have:

$$W_{AB} = \frac{P_M(\Delta_A) + P_M(\Delta_B)}{2k} = \tfrac{1}{2}W_A + \tfrac{1}{2}W_B.$$

It follows that:

$$\begin{aligned} p_{W_{AB}}(n \in \Delta) &= \mathrm{Tr}(U_t^{-1} W_{AB} U_t P_N(\Delta)) \\ &= \mathrm{Tr}(U_t^{-1}(\tfrac{1}{2}W_A + \tfrac{1}{2}W_B)U_t P_N(\Delta)) \\ &= \tfrac{1}{2}p_{W_A}(n \in \Delta) + \tfrac{1}{2}p_{W_B}(n \in \Delta), \end{aligned}$$

i.e., the probability of the photon arriving at region Δ on the detecting screen with both slits open is simply one half the sum of the probabilities with either slit A or slit B open. This is what one would expect on a classical analysis, and is contradicted by the interference pattern.

Notice that W_{AB} represents a mixture which does not in any way depend on the initial quantum state of the photon. But the initial state is *required* to be a plane wave (and not, say, a mixture of plane waves, as might be obtained by placing a candle to the left of the slit screen) in order to obtain the interference pattern.

Now, using the Lüders conditionalization rule, the conditionalized statistical operator yielding the photon statistics immediately after the photon has passed through slit A (with slit B closed) is the *pure* statistical operator $W_A = P_{\psi_A}$, where ψ_A is the normalized projection of ψ onto the subspace which is the range of the projection operator $P_M(\Delta_A)$. Immediately after the photon has passed through slit B (with slit A closed), the conditionalized statistical operator is $W_B = P_{\psi_B}$, where ψ_B is the normalized projection of ψ onto the subspace defined by $P_M(\Delta_B)$. Since ψ represents a plane wave, which assigns equal probabilities to equal ranges of M, it follows that the projection of ψ onto the subspace which is the range of the projection operator $P_M(\Delta_A) + P_M(\Delta_B)$ bisects the angle between ψ_A and ψ_B. Thus, immediately after the photon has passed through the slit system with both slits open, the conditionalized statistical operator is the *pure* statistical operator $W_{AB} = P_\theta$, where

$$\theta = \frac{\psi_A + \psi_B}{\|\psi_A + \psi_B\|}.$$

Since ψ_A and ψ_B are orthogonal unit vectors, $\|\psi_A + \psi_B\|^2 = \|\psi_A\|^2 + \|\psi_B\|^2 = 2$, and so

$$\theta = \frac{\psi_A + \psi_B}{\sqrt{2}}.$$

With slit A open and B closed, the probability of the photon arriving at a region Δ on the detecting screen after a time t is given by

$$p_{W_A}(n \in \Delta) = \mathrm{Tr}(U_t^{-1} P_{\psi_A} U_t P_N(\Delta)).$$

With slit B open and A closed, the probability is:

$$p_{W_B}(n \in \Delta) = \mathrm{Tr}(U_t^{-1} P_{\psi_B} U_t P_N(\Delta)).$$

With both slits open, the probability is:

$$p_{W_{AB}}(n \in \Delta) = \mathrm{Tr}(U_t^{-1} P_\theta U_t P_N(\Delta)).$$

To see the difference between this result and the calculation based on von Neumann's rule, let

$$\begin{aligned} U_t^{-1} P_{\psi_A} U_t &= P'_{\psi_A} \\ U_t^{-1} P_{\psi_B} U_t &= P'_{\psi_B} \\ U_t^{-1} P_\theta U_t &= P'_\theta \end{aligned}$$

and to simplify notation I shall write:

$$\begin{aligned} P_M(\Delta_A) &= P_A \\ P_M(\Delta_B) &= P_B \\ P_N(\Delta) &= Q \\ \|P_A \psi\| = \|P_B \psi\| &= l. \end{aligned}$$

Then:

$$\begin{aligned} p_{W_A}(n \in \Delta) &= \|Q\psi'_A\|^2 \\ p_{W_B}(n \in \Delta) &= \|Q\psi'_B\|^2 \\ p_{W_{AB}}(n \in \Delta) &= \|Q\theta'\|^2 \\ &= \|Q\left(\frac{\psi'_A + \psi'_B}{\sqrt{2}}\right)\|^2 \\ &= \tfrac{1}{2}\|Q\psi'_A + Q\psi'_B\|^2. \end{aligned}$$

It is important to notice that

$$\|Q\psi_A + Q\psi_B\|^2 = \|Q\psi_A\|^2 + \|Q\psi_B\|^2$$

but

$$\|Q\psi'_A + Q\psi'_B\|^2 = \|Q\psi'_A\|^2 + \|Q\psi'_B\|^2 + (Q\psi'_A, Q\psi'_B) + (Q\psi'_B, Q\psi'_A),$$

i.e., the existence of non-zero 'interference terms' depends on there being a non-zero distance between the slit screen and the detecting screen, and hence on a non-zero travel time t between the slit screen and the detecting screen.

Evidently, since M and N are compatible magnitudes, P_A commutes with Q and P_B commutes with Q. So:

$$\|Q\psi_A + Q\psi_B\|^2 = \|Q\psi_A\|^2 + \|Q\psi_B\|^2 + (Q\psi_A, Q\psi_B) + (Q\psi_B, Q\psi_A).$$

But

$$\begin{aligned}(Q\psi_A, Q\psi_B) &= (\psi_A, Q\psi_B)\\ &= \left(\frac{P_A\psi}{l}, Q\frac{P_B\psi}{l}\right)\\ &= \frac{1}{l^2}(\psi, P_AQP_B\psi)\\ &= \frac{1}{l^2}(\psi, QP_AP_B\psi)\\ &= 0\end{aligned}$$

since $P_AP_B = 0$.

Similarly: $(Q\psi_B, Q\psi_A) = 0$.

If $t \neq 0$, we have

$$\begin{aligned}(Q\psi'_A, Q\psi'_B) &= (\psi'_A, Q\psi'_B)\\ &= (U_t\psi_A, QU_t\psi_B)\\ &= \left(U_t\frac{P_A\psi}{l}, QU_t\frac{P_B\psi}{l}\right)\\ &= \frac{1}{l^2}(\psi, P_AU_t^{-1}QU_tP_B\psi).\end{aligned}$$

Although $(\psi, P_AQP_B\psi) = 0$, it is not the case that $(\psi, P_A(U_t^{-1}QU_t)P_B\psi)=0$, since P_A does not commute with $Q' = U_t^{-1}QU_t$, and P_B does not commute with Q'.

This is perhaps physically clear if we assume that Δ is a very small region, so that $P_N(\Delta) = Q$ is effectively a δ-function centred at a point x. Then $\|Q\psi_A'\|^2$ (which is the square of the projection of the Hilbert space vector ψ_A' onto the subspace defined by Q) may be expressed as $|\phi_A'(x)|^2$, where ϕ_A' represents the wave function emanating from slit A at time t (when the wave front reaches the detecting screen). Similarly, $\|Q\psi_B'\|^2$ may be expressed as $|\phi_B'(x)|^2$, and $\|Q\theta'\|^2$ as $|((\phi(_A'x) + \phi_B'(x))/\sqrt{2}|^2$. Thus, with both slits open, the probability of the photon hitting the point x on the detecting screen is:

$$\tfrac{1}{2}|\phi_A'(x) + \phi_B'(x)|^2 = \tfrac{1}{2}|\phi_A'(x)|^2 + \tfrac{1}{2}|\phi_B'(x)|^2 + \tfrac{1}{2}(\phi_A'(x), \phi_B'(x)) + \tfrac{1}{2}(\phi_B'(x), \phi_A'(x),$$

and although $(\phi_A(x), \phi_B(x)) = 0$, it is clearly not the case that $(\phi_A'(x), \phi_B'(x)) = 0$, since the transition $\phi_A \rightarrow \phi_A'$ and $\phi_B \rightarrow \phi_B'$ represents a unitary transformation generated by a Hamiltonian involving a momentum operator which is incompatible with the position magnitude represented by M or N.

More generally, if $P_A\psi \perp P_B\psi$, i.e., $(P_A\psi, P_B\psi) = 0$, then $QP_A\psi \perp QP_B\psi$, i.e., $(QP_A\psi, QP_B\psi) = 0$, if and only if the projection operator Q is compatible with P_A and P_B. This is geometrically obvious if we think of Q as defining a plane in a 3-dimensional Hilbert space. If the plane $\mathscr{K}_{AB}$ defined by the orthogonal pair $P_A\psi$ and $P_B\psi$ is orthogonal to the plane $\mathscr{K}$ defined by Q, then the angle between the projections of $P_A\psi$ and $P_B\psi$ into $\mathscr{K}$ is π. If we rotate the plane $\mathscr{K}_{AB}$ about the line defined by the intersection between $\mathscr{K}_{AB}$ and $\mathscr{K}$, then the angle between the projections of $P_A\psi$ and $P_B\psi$ onto $\mathscr{K}$ decreases continuously to $\pi/2$, when the planes $\mathscr{K}$ and $\mathscr{K}_{AB}$ coincide (and Q is compatible with P_A and P_B).

This analysis of the 2-slit experiment makes clear the role played by (i) the initial quantum state, and (ii) the non-zero distance between the slit screen and the detecting screen.[4]

The explanation of the interference effect depends on the difference between the Lüders conditionalization rule and the von Neumann rule. Von Neumann's rule

$$W \rightarrow W' = \frac{P_{a_i}}{\mathrm{Tr}(P_{a_i})}$$

is the analogue of the classical rule

$$W \rightarrow W' = \frac{I_{a_i}}{\sum (I_{a_i})}$$

representing a conditionalization and *randomization* of the initial measure within the subset X_{a_i}. The natural generalization of the classical conditionalization rule appropriate to non-Boolean possibility structures is the Lüders rule. Thus, the 'paradox' involved in the 2-slit experiment is resolved by showing precisely how the assumption of a non-Boolean possibility structure explains the existence of the 'anomalous' interference effects.

Appendix

The following Lemmas develop some properties of statistical operators and projection operators which are essentially trivial, but may not be familiar to some readers. They are collected here for convenience in following the argument of Section III.

Lemma 1

$$\mathrm{Tr}(WP_\psi) = (\psi, W\psi).$$

Let $\{\phi_j\}$ be a complete orthonormal set of vectors spanning the Hilbert space such that $\phi_1 = \psi$. Then

$$\begin{aligned}\mathrm{Tr}(WP_\psi) &= \sum_i (\phi_j, WP_\psi \phi_j) \\ &= (\psi, W\psi).\end{aligned}$$

Lemma 2

$$P_{a_i} WP_{\alpha_i} = (\alpha_i, W\alpha_i)P_{\alpha_i} = \mathrm{Tr}(WP_{\alpha_i})P_{\alpha_i}.$$

For any vector θ,

$$P_\psi \theta = (\psi, \theta)\psi$$

and so

$$\begin{aligned}P_{\alpha_i} WP_{\alpha_i}\theta &= P\alpha_i W(\alpha_i, \theta)\alpha_i \\ &= (\alpha_i, \theta)P_{\alpha_i} W\alpha_i \\ &= (\alpha_i, \theta)(\alpha_i, W\alpha_i)\alpha_i \\ &= (\alpha_i, W\alpha_i)P\alpha_i\theta\end{aligned}$$

i.e.,

$$P_{\alpha_i} WP_{\alpha_i} = (\alpha_i, W\alpha_i)P_{\alpha_i}.$$

By Lemma 1

$$\mathrm{Tr}(WP_{\alpha_i}) = (\alpha_i, W\alpha_i).$$

LEMMA 3

$$\mathrm{Tr}(P_\psi P_{a_i}) = (\psi, P_{a_i}\psi) = \|P_{a_i}\psi\|^2.$$

By Lemma 1

$$\begin{aligned}\mathrm{Tr}(P_\psi P_{a_i}) = \mathrm{Tr}(P_{a_i}P_\psi) &= (\psi, P_{a_i}\psi)\\ &= (\psi, P_{ai}^2\psi)\\ &= (P_{a_i}\psi, P_{a_i}\psi)\\ &= \|P_{a_i}\psi\|^2.\end{aligned}$$

If $P_{a_i} = P_{\alpha_i}$, i.e., if P_{α_i} is a projection operator onto a 1-dimensional subspace spanned by the vector α_i, then

$$\begin{aligned}\mathrm{Tr}(P_\psi P_{a_i}) &= (\psi, P_{\alpha_i}\psi)\\ &= (\alpha_i, \psi)(\psi, \alpha_i)\\ &= |(\alpha_i, \psi)|^2\end{aligned}$$

LEMMA 4

$$P_\phi P_\psi P_\phi = |(\phi, \psi)|^2 P_\phi$$

For any vector θ

$$\begin{aligned}P_\phi P_\psi P_\phi \theta &= P_\phi P_\psi(\phi, \theta)\phi\\ &= P_\phi(\phi, \theta)(\psi, \phi)\psi\\ &= (\phi, \theta)(\psi, \phi)(\phi, \psi)\phi\\ &= |(\phi, \psi)|^2 P_\phi \theta\end{aligned}$$

i.e., $$P_\phi P_\psi P_\phi = |(\phi, \psi)|^2 P_\phi.$$

(This result also follows from Lemmas 2 and 3.)

LEMMA 5

If $\theta' = \theta/\|\theta\|$, then $P_\theta/\|\theta\|^2 = P'_\theta$.
For any vector θ

$$P_\theta/\|\theta\|^2\theta = (\theta, \phi)/\|\theta\|^2\theta = (\theta', \phi)\theta' = P'_\theta\phi$$

NOTES

[1] A non-commutative extension of the classical notion of conditional probability (more generally, conditional expectation) has been extensively investigated by Umegaki (H. Umegaki, 'Conditional Expectation in an Operator Algebra', I. *Tôhoku Math. J.* **6**, 177–181 (1954); II. **8**, 86–100 (1956); III. *Kodai Math. Semi. Rep.* **11**, 51–64 (1959); IV. *Kodai Math. Semi. Rep.* **14**, 59–85 (1962)). Umegaki's theory has recently been extended to magnitudes with continuous spectra by Davies and Lewis (E. B. Davies and J. T. Lewis, *Commun. Math. Phys.* **17**, 239–260 (1970)). Nakamura and Umegaki have shown that von Neumann's projection postulate is just the conditionalization of the statistical operator relative to an event in the non-Boolean possibility structure. (M. Nakamura and H. Umegaki, 'On von Neumann's Theory of Measurement in Quantum Statistics', *Math. Jap.* **7**, 151–157 (1961–62)). Their demonstration considers only maximal (i.e., non-degenerate) magnitudes with discrete spectra, in which case the Lüders rule coincides with von Neumann's rule. For a discussion of the Lüders rule *vis à vis* von Neumann's rule, see Section III.

[2] By the properties of the trace operation,

$$\mathrm{Tr}(P_{c_i} W P_{c_i}) = \mathrm{Tr}(W P_{c_i}) = \mathrm{Tr}(P_{c_i} W).$$

I shall continue to write such expressions in symmetrical form below.

[3] G. Lüders, *Ann. d. Physik* **8**, 322 (1951)). The Lüders rule is discussed at some length by Furry in W. H. Furry, 'Some Aspects of the Quantum Theory of Measurement', *Lectures in Theoretical Physics Volume VIIIA, Statistical Physics and Solid State Physics*, University of Colorado Press, Boulder, 1966.

[4] Compare this analysis with Putnam's discussion in H. Putnam, 'Is Logic Empirical?', *Boston Studies in the Philosophy of Science V*, R. Cohen and M. Wartofsky (eds.), Reidel, 1969. Putnam's solution to the problem posed by the phenomenon of interference is to block the application of the distributive law in transforming the conditional probability on passage of the photon through both slits, to a sum of conditional probabilities for each of the slits separately. This solution is spurious, however, because the usual classical notion of conditional probability is inapplicable if the possibility structure of events is non-Boolean. Notice that the initial quantum state plays no role in Putnam's analysis, and there is no explicit recognition of the significance of the distance between the slit screen and the detecting screen (although a non-zero distance is implicitly required for the non-distributivity of the events considered).