LÜDERS'S RULE AS A DESCRIPTION OF INDIVIDUAL STATE TRANSFORMATIONS*

SERGIO MARTINEZ†‡

Instituto de Investigaciones Filosoficas
Universidad Nacional Autonoma de Mexico

Usual derivations of Lüders's projection rule show that Lüders's rule is the rule required by quantum statistics to calculate the final state after an ideal (minimally disturbing) measurement. These derivations are at best inconclusive, however, when it comes to interpreting Lüders's rule as a description of individual state transformations. In this paper, I show a natural way of deriving Lüders's rule from well-motivated and explicit physical assumptions referring to individual systems. This requires, however, the introduction of a concept of individual state which is not standard.

1. Introduction. In 1955, von Neumann added to his formulation of quantum mechanics a postulate that he considered to be indispensable to provide a satisfactory description of the measurement process. Von Neumann's original version of this postulate said that "if a physical quantity R is measured twice in succession on a system then we get the same value each time" (p. 335). Nowadays the postulate is known as "the first kind condition" (Pauli) or "the projection postulate" (Margenau).

It is usually considered noncontroversial to think of the projection postulate as a simplifying assumption describing a class of idealized measurement procedures. For maximal measurements this seems to provide a satisfactory interpretation of von Neumann's postulate, but for nonmaximal measurements, that is, for measurements of magnitudes with degenerate eigenvalues, a further problem arises.

Suppose that we measure a nonmaximal magnitude A. Von Neumann's view was that such a measurement should be understood as a partial description (with loss of information) of a maximal measurement. Accordingly, von Neumann argued that the final state of a system after a nonmaximal measurement should be described by a mixture of all eigenstates of the eigenvalue representing the measurement result. Lüders was the

*Received May 1988; revised June 1989.

†I would like to thank Linda Wessels and Geoffrey Hellman who directed the research on which this paper is based. I also would like to thank an anonymous referee of this Journal for valuable comments.

‡Send reprint requests to Instituto de Investigaciones Filosoficas, Universidad Nacional Autonoma de Mexico, Circ. Mario de la Cueva, 04510 Coyoacan, MEXICO, D.F.

Philosophy of Science, 58 (1991) pp. 359–376.

first to point out that such a rule for the calculation of a final state after a nonmaximal measurement could not be right since at least in some cases it leads to unacceptable results.

Bub (1979) illustrates the main difficulty with von Neumann's proposal by means of the following example. Suppose that a two particle system $S_1 + S_2$ is represented by the Hilbert space $H_1 \otimes H_2$, where H_1 represents S_1 and H_2 represents S_2. Let A be a maximal magnitude in S_1, and suppose we measure nonmaximal magnitude $A \otimes T_2$ in $H_1 \otimes H_2$. According to von Neumann's proposal the state resulting from this measurement depends on the maximal magnitude in $H_1 \otimes H_2$ that we choose to measure. But there is a very natural and compelling sense in which this measurement seems to be independent of the choice of maximal magnitude. After all, measuring $A \otimes I_2$ is from a physical viewpoint nothing more than measuring the maximal magnitude A in S_1. An appropriate rule for the change of state on measurement should be independent of the choice of maximal magnitude.

Lüders (1951) proposed the following alternative rule for the description of the final state after a nonmaximal measurement: the final state after a (first kind) nonmaximal measurement is given by the projection of the original state on the subspace (eigenspace) of the measurement result. In Martinez (1990), I examine in detail the most conspicuous derivations of Lüders's rule and argue that one has to distinguish carefully between justifying an interpretation of Lüders's rule as a purely statistical rule for the calculation of the final (statistical) state and justifying an interpretation of it as a description of individual state transformations. Lüders's arguments, as well as more sophisticated arguments (by Herbut 1969, Bub 1979 and others) show that Lüders's rule is *the* rule required by quantum mechanical statistics. But the usual derivations are at best inconclusive when it comes to justifying Lüders's rule on the second interpretation. In this paper, I show that if the underlying concept of individual state in the interpretation of quantum mechanics is not the usual (standard) concept but a concept of individual state relative to magnitudes, there is a natural way of deriving Lüders's rule from well-motivated and explicit physical assumptions.

2. The Concept of Individual State. An individual state is supposed to describe the values of magnitudes (properties) of a system. In classical statistical mechanics, we can take statistical states to reflect distributions over properties of individual systems in a straightforward way. This procedure is summarized in the concept of phase space. States of individual systems can be identified with pure (dispersion free) statistical states. An analogous reasoning leads us to the usual (standard) concept of individual state as a set of "properties" (closed projections) to which a given "state-

vector" assigns probability one. In the framework of the present investigation, the quantum logical (lattice) formulation of quantum mechanics (see appendix for a brief overview), this is equivalent to say that states are represented by lattice ultrafilters, that is, by lattice filters generated by atoms of the lattice representing the system. A measurement state transformation can be defined accordingly as a function of two variables: $T(a,r) = b$, where a is the (lattice) atom generating the initial state, r is the proposition representing the result of measurement and b is the (lattice) atom generating the final state. The equation $T(a,r) = b$ is supposed to satisfy the condition that $b < r$.

The analysis of the projection rule that I have carried out in Martinez (1988, 1990) strongly suggests that the clarification of the role of the projection rule in quantum mechanics is intimately related with the exploration of a different (nonstandard) concept of individual state. Several alternative representations of the concept of individual state are scattered (most often implicitly) in the literature. Kochen (1979) has proposed a distinction between individual and statistical states. Statistical states for Kochen are probability functions assigning probabilities to state (or property) transitions. Individual states represent the properties a system has *relative to a certain interaction*. Kochen shows how the Hilbert space formalism has a natural way of accommodating such a concept of individual state. A pair of correlated Boolean σ-algebras of projection operators always exists in the two component spaces representing the interacting systems. Individual states for Kochen are represented by ultrafilters in these Boolean σ-algebras. *States which are represented by Boolean ultrafilters* will be here generically referred to as *B-states*. Kochen's individual states are examples of this sort of individual states. I will derive Lüders's rule for *B*-states.

Different interpretations, of course, can be formulated under the assumption that states for individual systems are represented by Boolean filters. Details on the way in which *B*-filters can be linked to particular interpretations are here left aside as this is not relevant to my present purpose. An exposition and discussion of alternative definitions of individual states in the quantum logical framework and their relation with different interpretations of quantum mechanics can be found in Hardegree (1980). In general, Bohrian type interpretations lead to a quantum logical formalization centered on the concept of *B*-state. A relevant discussion of the relation between some of Bohr's semantical insights on the interpretation of quantum mechanics and a quantum logical framework can be found in van Fraassen and Hooker (1976).

The assumption that individual states are represented by Boolean ultrafilters leads to a modification of the concept of state transformation used in the orthodox analysis. A state transformation is now a function

mapping Boolean filters to Boolean filters. It will be shown in the next pages that one can derive Lüders's rule in this framework in such a way that Lüders's rule receives a natural interpretation as a description of "dispersion free" individual state transformations. First, the framework is sketched on the basis of which a semantical reconstruction of the quantum logical structure can be carried out. Details are left aside since they can be filled out following standard reconstructions, (for example, see Piron 1976).

3. The Basic Semantical Framework. The usual motivation for the construction of the quantum logical framework (see, for example, Piron 1976) is to explain the quantum mechanical structure on the basis of purely "logical" assumptions. In order to achieve this goal, however, quantum logicians rely on the usual interpretation of quantum mechanics and, when that is not enough, on purely ad hoc postulates. My reconstruction of the quantum logical structure has an altogether different motivation. I do not think one can get away in quantum mechanics without some basic assumption about the probabilistic structure of the theory which is not reducible to "logical" considerations. Therefore, I explicitly start with axioms of a probabilistic nature. They are very simple and natural axioms and, once they are accepted, little is needed in the way of interpretive commitments to arrive at the quantum logical structure. Furthermore, these interpretive commitments are formulated explicitly via the concept of *B*-state. In this way the logico-algebraic formulation of the foundations of quantum mechanics can serve the purpose of discussing and clarifying interpretive issues on an altogether different basis from what is commonly referred as "Quantum Logic" (like Piron's, but also Bub 1979, Stairs 1982, and Friedman and Putnam 1978).

I assume that a physical system can be represented by a set of discrete magnitudes, each magnitude being represented by a complete atomic Boolean lattice generated by a set of possible exclusive events, the results of the measurements of the magnitude. Each result or outcome corresponds to an atom of the Boolean lattice representing the magnitude. It will be assumed that each magnitude is a submagnitude of (at least one) maximal magnitude and furthermore that all maximal magnitudes have the same (at most countable) dimension. Notice that in each magnitude M_i (via its representation as a Boolean algebra) there is a well-defined orthogonality relation $\perp_i$. I will write $x \perp y$ instead of "$x \perp_i y$ for some $i \in I$", where i indexes the magnitudes of a system S. A basis for a magnitude is a maximal set of (orthogonal) atoms. Since an atomic Boolean lattice is uniquely generated by its atoms, I will usually identify a maximal magnitude with its basis. I postulate that a *measurement situation* can be represented by a pair (a,M), where a is an atom of M and M is a maximal magnitude. This reflects the idea (originating with Bohr)

that values of magnitudes in quantum mechanics must be understood as relative to an "experimental arrangement". Let $[a)_M$ be the Boolean ultrafilter in M generated by a. The expression $[a)_M$ will be taken to represent the *B-state* generated by the situation (a,M). Sometimes I will talk of maximal magnitudes as measurement situations. This ambiguity will allow for a less cumbersome notation, but it should be kept in mind that, when fully described, a measurement situation includes the specification of a state (atom) of the corresponding maximal Boolean algebra.

It might be worth emphasizing that I am not going to argue here for any specific interpretation of quantum mechanics based on the concept of B-state. I believe that a satisfactory interpretation of quantum mechanics (if any can be found) will incorporate a concept of individual state "relative" to experimental situations as this is formalized in the concept of B-state. But precisely what this relativity amounts to is an open question and I do not pretend to answer this question here. Accordingly, I do not address crucial decisions leading to different interpretations. For example, I leave open the question of whether the state of a physical system relative to experimental situations is taken to mean that the system has indeed relational properties with respect to possible physical interactions as Kochen claims, or whether this concept of B-state applies only to the properties elicited by an actual (last) measurement as Bohrian interpretations claim. This is, of course, a crucial interpretive issue, one that must be addressed in order to analyze EPR type situations for example. *My point is that independently of these interpretive commitments, a physical interpretation of Lüders's rule is available for interpretations in which individual states are taken to be B-states.*

Finally, the following fundamental assumption is added: the B-states, that is, the atoms of the maximal Boolean lattices representing a given physical system, are related by a transition probability function. In this way, I formulate the assumption that there is a nonclassical physical propensity for states of quantum systems to "jump" to a different state in a different experimental (measurement) situation. In other words, I postulate the existence of a transition probability function that assigns a number (between zero and one) interpreted as the probability of transition between states for each pair of (Boolean) states. The question of exactly how this transition probability is to be interpreted physically will not be addressed here, (but see section 4 for a relevant discussion). It will be assumed that however such transition is to be interpreted, the corresponding transition probability function $t(.\ ,\ .)$ will satisfy the following axioms: for x,y B-states of a system S,

Axiom 1: $x \perp y$ if and only if $t(x,y) = 0$,

Axiom 2: $\Sigma_{i\geq 1} t(x,y_i) = 1$ where $\{y_i\}$ is any basis (a maximal set of orthogonal atoms).

Let me first motivate these axioms.[1] Notice that axiom 2 is a straightforward requirement on probability assignments. It simply states the assumption that $t(x,y)$ is supposed to describe probabilities. Now consider axiom 1. If $x \perp y$ (i.e., if x and y are B-states in a common Boolean algebra), we certainly want $t(x,y) = 0$. After all, what we mean when we say that two atoms belong to a common magnitude (experimental situation) is that they behave "classically" with respect to each other in the sense that if one of the two states happens to be the case, the other is excluded (in the classical sense of not being the case). The other direction of axiom 1, however, carries with it an important empirical claim. It takes the impossibility of a transition between two Boolean states to imply an orthogonality relation between them.

The semantical framework to be used here relies on an identification of Boolean states based on the following theorem:

THEOREM 3.1. *For x,y B-states, $t(x,y) = 1$ if and only if $x = y$.*

Proof: Assume $t(x,y) = 1$. The element y is a B-state and thus there exists a maximal magnitude with basis $\{y_i\}$ such that $y_1 = y$. Hence, by axiom 2, $t(x,y) = t(x,y) + \Sigma_{i\geq 2} t(x,y_i) = 1$. Since $t(x,y) = 1$, then $t(x,y_i) = 0$ for $i \geq 2$, and thus, by axiom 1, $x \perp y_i$ for $i \geq 2$. Element x is then a B-state orthogonal to all but one of the elements of a basis (of a Boolean lattice) and thus $x = y$. Now suppose $x = y$. Then there is a basis $\{y_i\}$ such that $y_1 = x$ and $\Sigma_{i\geq 1} t(x,y_i) = 1$ by axiom 2. If $t(x,y_1) = t(x,x) \neq 1$, then $t(x,y_i) \neq 0$ for some $i \geq 2$, contradicting axiom 1.

The quantum logical structure will be seen as arising from a process in which the different magnitudes are "glued" together into a physical system represented by an orthomodular structure. *Two B-states (atoms) are "glued" together if the transition probability between them is 1.*

4. The Derivation of Lüders's Rule. In the orthodox view, states get represented by lattice ultrafilters and thus the concept of state transformation can be defined independently of magnitudes (see section 2). But in our present framework, based on the concept of B-states, we have to modify accordingly the underlying concept of state transformation to be used in the description of measurement. Given a system in state $[a)_M$ and assuming that magnitude $_N$ is measured with result (represented by prop-

[1]The function $t(.\ ,\ .)$ is an abstract version of the transition probability function for pure states in quantum mechanics given by the square of the absolute value of the inner product of vectors representing the pure states. See Beltrametti and Cassinelli (1981, chap. 18) for a survey of this topic. For a specific proposal formulating the quantum mechanical structure in terms of transition probability structures, see Mielnik (1968).

osition) r, the final individual state must be $[b)_N$ where $r > b$ and b is an atom of $_N$. We can represent *state transformations* of this sort as follows: $T([a)_M,r) = [b)_N$. However, when the specification of the magnitudes involved is unnecesary (as it will be usually the case), reference to magnitudes will be dropped and we will simply write $T(a,r) = b$. Also, when the specification of r is unnecesary, I will write $T(a) = b$.

State transformations involve an initial and a final magnitude. Suppose that a preparatory measurement of magnitude M has taken place and system is in state $[a)_M$, and that subsequently we measure proposition A. Suppose $\mathcal{A} = \{a_i\}$ is a subset of the atoms of magnitude N generating A (i.e., $\vee \{a_i\} = A$). In general there will be several atoms (states) in $\mathcal{A}$ which are accessible (nonorthogonal) from the initial state a. Each accessible state in $\mathcal{A}$ defines a state transformation which, we will say, is *N-consonant* with the measurement of A. For example, let b_1 and b_2 be two accessible states from a, with $b_1, b_2 \in N$. This means that $t(a,b_1) \neq 0$ and $t(a,b_2) \neq 0$. The equations $T(a) = b_1$ and $T(a) = b_2$ are state transformations *N-consonant* with A. But how can we go from here to establish a connection with Lüders's rule? Let us see first some examples of what can happen.

Consider the lattice generated by the gluing of two magnitudes $M = \{a_1,a_2,a_3\}$ and $N = \{b_1,b_2,b_3\}$. Let us suppose that $t(a_3,b_3) = 1$ but $t(a_1,b_i) \neq 1$ for $i \neq 3$, $j \neq 3$. According to axiom 1 for $t(\ ,\ )$ we identify a_3 with b_3. Clearly $a_3^{\perp}$ holds whenever $b_3^{\perp}$ holds (and vice versa). This leads us to identify also $a_3^{\perp}$ with $b_3^{\perp}$. The Haase diagram (see Appendix, part A) of the gluing of M and N is given in figure 1. Select a state in M and a proposition in N. For example, let the initial state be a_1 and suppose we determine $b_2 \vee b_3$ as a result of measurement. The only state transformation N-consonant with measurement is $T(a_1) = b_2$.

Now consider another example. Let $M = \{a_1,a_2,a_3,a_4\}$ and $N = \{b_1,b_2,b_3,b_4\}$. Suppose that M and N have only one overlapping element $a_3 = b_3$ for example. The Haase diagram of such "gluing" is given in figure 2. With initial state a_1, if the result of measurement is $b_2 \vee b_4$, for example, $T(a_1) = b_2$ and $T(a_1) = b_4$ are N-consonant state transformations. But suppose that there are two overlapping elements, say $a_2 = b_2$ and $a_4 = b_4$. In this case we should also identify $a_2 \vee a_4$ with $b_2 \vee b_4$. The resulting Haase diagram of this gluing is represented in figure 3. Suppose, for example, that initial state is a_1 and we measure $b_3 \vee b_4$. The only measurement transformation consonant with this measurement is $T(a_1) = b_3$. We will say in this case that $T(a_1) = b_3$ is the *pure* or *dispersion free* state transformation representing the measurement. Notice that in this situation, for any initial state (atom) a_i in M there is a unique state transformation $T(a_i) = b$ which is consonant with the measurement of an arbitrary proposition A (generated by N).

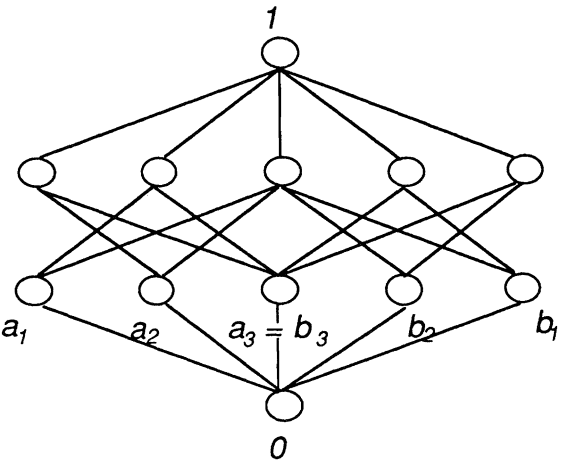

Figure 1. Haase diagram of the lattice generated by the Boolean lattices A (with atoms a_1, a_2, a_3) and B (with atoms b_1, b_2, b_3). Elements are represented by circles, and an element x being "less than" another element y is represented by joining the two elements with a line going up from x to y.

In general, for a system initially in state $[a)_M$ a state transformation $T(a,A) = b$ is called *dispersion free* (or *pure*) with respect to the measurement of a proposition A in N *if* $T(a,A) = b$ *is the only state transformation (into* N*) consonant with this measurement of* A. It seems then natural to say that M and N are *ideal magnitudes relative to each other* if, given that the system is in a state represented by an atom of one of the two magnitudes, after measurement of the other magnitude the system

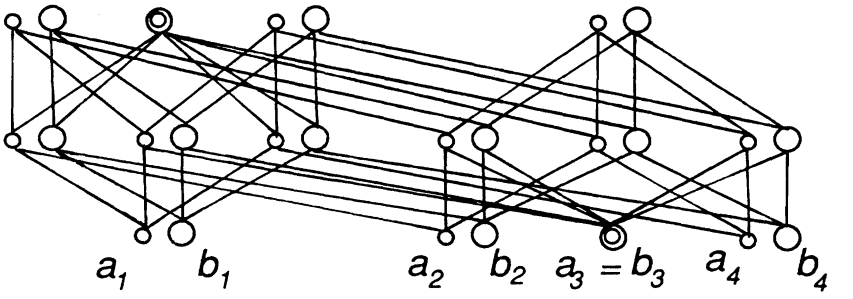

Figure 2. Partial Haase diagram of the pasting of two Boolean lattices with center $C(L) = \{0,1,a_3,a_3\}$. Relations of the form $0 \lessdot x$ and $x \lessdot 1$ are not depicted. For easier reading of the diagram, elements of the different Boolean lattices are represented by circles of different sizes.

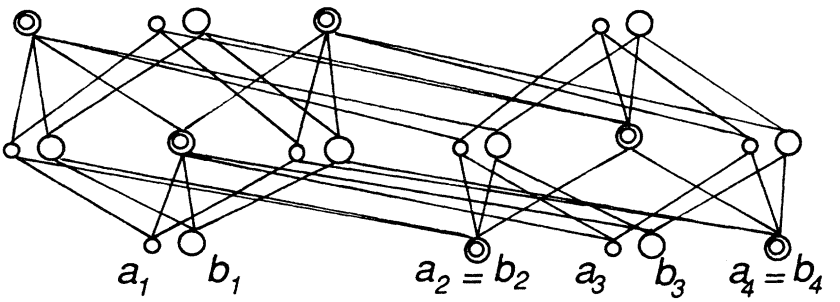

Figure 3. Partial Haase diagram of the pasting of the same Boolean lattices as in figure 2, but now with full center.

either remains in the same state (but maybe in a different measurement situation) or else it changes to a state described by a pure state transformation. The following question arises. Is it possible to find simple necessary and sufficient conditions characterizing (algebraically) magnitudes which are ideal relative to each other? This would lead us to a characterization of ideal measurements via pure B-state transformations. I will answer this question in the affirmative and show how this approach to ideal measurements provides us with an interpretation of Lüders's rule as a description of pure (dispersion free) transformations.

It is clear from the definition of the center of a lattice (see definition [A.5] in Appendix, part A) that the center of the lattice generated by the gluing of magnitudes M and N is constituted by the elements that are "glued" together. The examples above suggest that intuitively two magnitudes are ideal relative to each other when they have a "big enough" (relative) center. This idea is made precise below. The proofs of the following theorems are carried out using terminology and lemmata proved in part B of the appendix.

THEOREM 4.1. *Let $L = Lat(M,N)$ be the sublattice of two blocks generated by maximal magnitudes M and N. $C(L)$ (the center of L) is full if and only if M is ideal relative to N.*

Proof: Suppose M and N are not relatively ideal. That means that there is a proposition A, $A \in N$ but $A \notin C(L)$ such that an N-basis of A includes two atoms x, y accessible from some state a in M. Suppose $x \vee y \in C(L)$, then $x \vee y \in F_c$ (where F_c is the [maximal] central filter generated by f_0, see lemma [B.5] in Appendix, part B and below) and $x \vee y < A$, and thus $A \in C(L)$ contradicting assumption; so, $x \vee y \notin C(L)$. Hence, the "floating atom" f_0 (see lemma [B.6] in Appendix, part B) is generated by at least three atoms and this

implies, by corollary (B.7) in Appendix, part B, that $C(L)$ is not full. This proves (by contraposition) that the fullness of $C(L)$ is a sufficient condition for M to be relatively ideal to N. Now suppose that $C(L)$ is not full. Then $A(F)$ has more than two elements (by corollary [B.7]) and thus there are two atoms $x \in N$, $y \in N$ such that $x \notin C(L)$, $y \notin C(L)$ and $x \vee y \notin C(L)$. Take $b = x \vee y$. There are two state transformations $T(a,x)$ and $T(a,y)$ which are consonant with the measurement of b and thus M and N are not ideal relative to each other.

THEOREM 4.2. *Let $L = Lat(M,N)$ be a two block orthomodular lattice (generated by atomic Boolean lattices M and N). Lüders's rule selects a unique final state in L if and only if $C(L)$ is full.*

Proof: Assume $C(L)$ is not full. Hence, for any x in N, and for any atom a of M, with a not in $C(L)$, $x^{\perp} \vee a \geq e(a) \not> b$, for any atom b in N. From the fact that $e(a)$ does not cover b, it follows that $e(a) \geq b_1 \vee b_m \vee b_n$ for b_1, b_m, b_n atoms of N. Select $x = b_1 \vee b_m$. Clearly $x \wedge (x^{\perp} \vee a) \geq x \wedge e(a) \geq b_1 \vee b_m$. Thus Lüders's rule does not select in this case an atom of N. Assume $C(L)$ is full. We want to show that $x \wedge (x^{\perp} \vee a)$ is an atom (or $x \wedge (x^{\perp} \vee a) = 0$). Now, if $\{x,x^{\perp},a\}$ is a distributive triple, then $x \wedge (x^{\perp} \vee a) = 0$, or $x \wedge (x^{\perp} \vee a) = a$. Suppose that $\{x,x^{\perp},a\}$ is not a distributive triple. In this case, we must have $a \notin I_c$, and $x \in N$. Since $a \notin I_c$, then $e(a) \in F_c$ (by lemma [B.5]) and $e(a) = b_1 \vee b_2$ for b_1, b_2 atoms of N (by lemma [B.2]). The following cases arise: (i) $b_1 \leq x$, $b_2 \leq x$. Since $e(a) \in F_c$, it follows that $x \in C(L)$. This case reduces then to the case in which the triple $\{x,x^{\perp},a\}$ is a distributive triple. (ii) $b_1 \not\leq x$, $b_2 \not\leq x$. In this case, $b_1 \leq x^{\perp}$, $b_2 \leq x^{\perp}$, it follows then that $x^{\perp} \in C(L)$ and thus again it reduces to the case of a distributive triple. (iii) $b_1 \leq x$, $b_2 \not\leq x$. Now, $b_2 \not\leq x$ implies $b_2 \not\leq x \wedge (x^{\perp} \vee a)$; on the other hand, $b_1 \leq x$, and $b_1 \leq e(a)$ by hypothesis for this case, and $e(a) \leq e(a) \vee e(x^{\perp}) = e(a \vee x^{\perp}) = a \vee x^{\perp}$, thus $b_1 \leq x \wedge (x^{\perp} \vee a)$. Suppose $b_1 < x \wedge (x^{\perp} \vee a)$. Then, by atomicity of L, there must exist an atom $b_j \in N$, $b_j \neq b_1$, such that $b_j < x \wedge (x^{\perp} \vee a)$, hence $b_j \leq x$ and $b_j \leq (x^{\perp} \vee a) \leq x^{\perp} \vee e(a)$. But $b_j \not\leq e(a)$ since by hypothesis $e(a) = b_1 \vee b_2$; and $b_j \not\leq x^{\perp}$ since $b_j < x$. This shows that there cannot be such an atom b_j and thus $b_1 = x \wedge (x^{\perp} \vee a)$. (iv) $b_1 \not\leq x$, $b_2 \leq x$. As for (iii), one concludes that $b_2 = x \wedge (x^{\perp} \vee a)$.

COROLLARY 4.3. *If two magnitudes M and N are ideal relative to each other then the only state transformation N-consonant with a measurement of a submagnitude A of N, for arbitrary initial situation (a,M), is given by Lüders's rule.*

This corollary follows immediately from theorems (4.1) and (4.2).

A lattice (with 0) has the covering property if $a \wedge x = 0$ implies $x \prec x \vee a$ for any atom a and $x \in L$ ($x \prec y$ if $x < y$ and for no z, $x < z < y$). If the covering property holds in an orthomodular lattice then Lüders's rule selects an atom in the lattice (see Piron 1976, for example). Using this result and theorem (4.2), the following corollary follows. The corollary can also be proved directly by a slight modification of the proof of theorem (4.2).

COROLLARY 4.4. *Two magnitudes M and N are ideal relative to each other if and only if Lat(M,N) has the covering property.*

The following theorem provides the final step in our analysis.

THEOREM 4.5. *If L is a complete atomic orthomodular lattice the following two conditions are equivalent:*

(i) L has the covering property.

(ii) for any pair (a,x), where a is an atom of L and x is an arbitrary element of L, a lattice Lat(a,x) with the following properties can be constructed (shown below):

(a) if a is compatible with x then Lat(a,x) is a maximal Boolean algebra including a and x.

(b) if a is not compatible to x then Lat(a,x) is an orthomodular lattice of two blocks with the covering property.

Proof: In case (ii,a) the equivalence with (i) is immediate. Assume that a is not compatible with x. Let us prove first that (i) implies (ii). I first show how the construction goes for three-dimensional lattices and then I sketch how this construction can be carried out for the general case. Under the assumption of the covering property, Lüders's rule selects an atom: $L(a,x) = x \wedge (x^{\perp} \vee a) = a_f$. Suppose x is two dimensional (the interesting case). That means that $(x^{\perp} \vee a)$ and x share one and only one atom (a_f). Now, $x^{\perp} \perp L(a,x)$, and since $(x^{\perp} \vee a)^{\perp}$ is also an atom, we can construct

$$B_f = \{x^{\perp}, (x^{\perp} \vee a)^{\perp}, x \wedge (x^{\perp} \vee a)\}$$
$$B_i = \{a, (x^{\perp} \vee a)^{\perp}, a^{\perp} \wedge (x^{\perp} \vee a)\};$$

B_i and B_f generate an orthomodular sublattice of two blocks with full center, and thus with the covering property. This lattice we call Lat(a,x). For more than three dimensions, the construction proceeds as follows: if x is a dual atom (i.e., if $x \prec l$), the construction of B_i and B_f—and of Lat(a,x)—follows the same steps, the only difference is that now the center $(x^{\perp} \vee a)^{\perp}$ is not an atom but a subspace and thus we have to replace $(x^{\perp} \vee a)^{\perp}$ with a basis of it. If x is not

a dual atom, select a dual atom s such that $x < s$ and construct Lat(a,s). Notice that $L(a,x) = L(a,s)$ and $(x^{\perp} \vee a)^{\perp} < (s^{\perp} \vee a)^{\perp}$, whence $L(a,x) \vee (x^{\perp} \vee a)^{\perp} = x \in \text{Lat}(a,s)$. Thus Lat($a$,$s$) has the properties required for Lat(a,x). To show that (i) follows from (ii) we have to show that if $a \wedge x = 0$ then $x \prec a \vee x$, for any atom a and any element x of L. By assumption there is a sublattice Lat(a,x) with the covering property generated by maximal Boolean algebras (in L), for any atom a and element x. Therefore, in the lattice Lat(a,x), if $a \wedge x = 0$ then $x \prec a \vee x$; but clearly there is no z in L such that $x < z < a \vee x$, since Lat(a,x) is generated by maximal Boolean blocks and $x < z < a \vee x$ would imply that z is in Lat(a,x), contradicting the assumption that the covering property holds in the Boolean blocks.

I have shown that a very simple algebraic condition characterizes "dispersion free" B-state transformations. Dispersion free transformations are minimally disturbing in the sense that the center of the orthomodular lattice of two blocks generated by the transformation is full. Alternatively, Corollary (4.3) provides ground at the level of individual states for the usual motivation for Lüders's rule as a rule describing measurements with "no loss of information". Furthermore, *the present derivation, interpreted as a description of a physically distinguishable class of (dispersion free) measurements, allows us to establish a link between individual and statistical state transformations via Lüders's rule*. The notion of "respecting compatibility" for statistical measurement transformations is found to have its counterpart in the notion of "dispersion free" individual state transformations. This connection between individual states and quantum statistics will be explored in more detail elsewhere.

At least something has to be said here, however, about an immediate question elicited by the above derivation. I have derived Lüders's rule for maximal magnitudes and B-states. This derivation then provides an alternative way of understanding von Neumann's original interpretation of nonmaximal measurements, but it confronts us, at least prima facie, with the difficulties that such an interpretation faced and which led to its dismissal. It is important, however, to notice first of all that the derivation here presented refers to B-state transformations and thus the objections raised against von Neumann's original proposal do not apply to this interpretation of nonmaximal measurements as *ideal* maximal measurements of a special sort. To see this, let us consider the objection to von Neumann's rule presented above (see section 1). There it was assumed that a measurement of one of the subsystems of the composite system $S_1 + S_2$ is made with an apparatus that *indeed performs* a maximal measurement in the composite system. This is unacceptable since it seems to imply that

the measurement apparatus has to read our mind to select the total system from which S_1 is a subsystem. The present derivation of Lüders's rule suggests the following different interpretation. The measurement apparatus interacts with the measured system in such a way that it changes the system in a certain way. It is a particular type of change that makes of the measurement an *ideal* measurement.

Returning to the above mentioned difficulty with von Neumann's proposal, as this comes out in Bub's example, we can think that the measurement of $A \otimes I_2$ in one of the subsystems results in a change of the individual state that is the same as the change it would experience if it were measured with a different apparatus, one that would measure the (unique) maximal magnitude associated with Lüders's rule through the present derivation.

5. The Relation Between $t(.\,,.)$ and Quantum Conditional Probabilities. In this section, I want to clarify the relation between the transition probability function $t(.\,,.)$ postulated in section 3 above and the probabilities generated by the fundamental algorithm of quantum mechanics. This should justify the implicit claim in my proposal that the postulated transition probabilities generate (or correspond to) the probabilistic structure of quantum mechanical predictions. Thus, at least in principle, one should be able to show how the usual quantum mechanical probabilities generated by measurement reduce to our framework of transition probabilities.

Several investigations have shown that important conceptual difficulties facing the interpretation of probabilities as distribution over properties can be overcome (or at least seen in a new light) once the probabilities in question are seen as ultimately referring to transition probabilities between states. Here, my aim is not to show how a full-fledged reduction of probabilistic claims in quantum mechanics to transition probabilities can be carried out, even less, how this reduction can be used for overcoming difficulties of interpretation beyond the problem of interpretation of the projection postulate. I assume one necessary component of such a reduction which is independent of any details of a specific reconstruction of the quantum logical (Hilbert space) structure or its philosophical consequences. This is the assumption that on the basis of transition probabilities we must be able to calculate conditional probabilities for elements A in L in a given state using the formula (c) below:

$$P(A/a) = \Sigma t(a_i,a) \qquad \text{(c)}$$

where $\{a_i\}$ is a basis A, and a is an arbitrary B-state.

The conditional probabilities given by (c) are supposed to correspond to the probabilities given by the fundamental algorithm of quantum me-

chanics. That is, $P(A/a)$ is to be interpreted as the probability that, given that the system is in state a, a measurement (of the appropriate kind) will show A to be the case. Formula (c) provides us with conditional probabilities which depend, at least prima facie, on the selection of the basis for A. Mielnik (1968) has shown that a sufficient condition for this formula to be representation invariant is that all bases have the same dimension. Since we are assuming that a system is constituted by maximal magnitudes of the same dimension, Mielnik's proof applies in our framework.

But look at figure 2. In this case, it is clear that the conditional probabilities described by (c) do not fully reduce to transition probabilities unless additional axioms for $t(.\ ,\ .)$ or empirical restrictions are added. For example, formula (c) requires that $P(b_2 \vee b_4/a_1) = t(b_2,a_1) + t(b_4,a_1)$. It seems that the only way in which this reduction of the conditional probability to transition probabilities can take place with full generality is if (at least) in one representation all but one of the terms in the sum—right side of (c)—reduces to zero. Otherwise it could very well happen that $P(b_2 \vee b_4/a_1) > 1$. It has been shown (see theorem 4.5) that the validity of the covering property is the necessary and sufficient condition for this to happen. Furthermore, this theorem shows that whenever the sum (c) reduces to zero for all but one of the terms, the only accessible (nonorthogonal) atom to the initial state is given by Lüders's rule. It is clear then that in quantum lattices the conditional probability formula (c) can be represented as follows:

$$P(A/a) = t(a,L(a,A)) \qquad \text{(cc)}$$

where $L(a,A)$ stands for the atom generated by the Lüders's transformation with initial state a and element A. We see then that the validity of the covering property in quantum lattices insures the reduction of the conditional probability formula to the transition probability framework we have postulated as fundamental. Thus, this result provides an alternative (lattice theoretical) route to the same conclusion reached by Bub's analysis of quantum statistics via generalization of the standard (Kolmogorov) theory of probability (see Bub 1979) and references therein.

To conclude, I have shown that one can derive Lüders's rule from well-motivated explicit physical assumptions within the framework of interpretations of quantum mechanics for which individual states get algebraically represented by B-states. It remains to be seen whether a satisfactory interpretation of quantum mechanics based on the representation of individual states by B-states can be implemented. But the fact that Lüders's rule plays a central role in the conceptual structure of quantum mechanics via B-state interpretations seems to me to point to a promising area of research in the philosophy of quantum mechanics, and in particular to a

way of making sense of the statistical claims of quantum mechanics without hidden variables.

APPENDIX

This appendix has two parts. In part A, I collect well-known terminology and results of lattice theory that are used in the paper. In part B, I prove some lemmata which characterize the structure of orthomodular lattices with two blocks as this is required for the proof of the theorems in sections 3 and 4.

Part A: Lattice Theoretical Background. An algebra $(L,\wedge,\vee)$ is called a lattice if L is a nonempty set and $\wedge$ and $\vee$ are binary idempotent, commutative and associative operations on L (called respectively the "meet" and the "join") satisfying the so-called absorption identities:

$$a \wedge (a \vee b) = a$$
$$a \vee (a \wedge b) = a \tag{A.1}$$

A lattice is *complete* if for any subset S of L the join $\vee$ $(a;\ a \in S)$ and the meet $\wedge$ $(a;\ a \in S)$ exist. Finite lattices can be described using Haase diagrams. *Haase diagrams* rely on the graphic representation (by means of segments of lines) of the relation of *covering*. We say that a *covers* b (notation: $a \succ b$) if $a > b$ and for no x, $a > x > b$. The elements of the lattice are represented by small circles. If x covers y then the circle representing x is higher up than the circle representing y.

An orthocomplemented lattice, or simply, *an ortholattice,* is a lattice with 0 and 1 and with an orthocomplementation (a mapping $a \to a^{\perp}$ of L onto itself satisfying:

i) $a \vee a^{\perp} = 1, \quad a \wedge a^{\perp} = 0;$
ii) $a \leq b \Rightarrow a^{\perp} \geq b^{\perp};$
iii) $a^{\perp\perp} = a.$

An *orthomodular lattice* is an ortholattice that satisfies the orthomodular identity:

$$a \leq b \Rightarrow b = a \vee (b \wedge a^{\perp}). \tag{A.2}$$

A triple (a,b,c) of elements of L is distributive if:

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \vee c). \tag{A.3}$$

A lattice is *distributive* if for any a, b, c in L (A.3) is satisfied. A distributive ortholattice is called a *Boolean lattice*. For distributive lattices orthocomplementation is a dual automorphism, thus the orthocomplement of an element is unique.

DEFINITION A.4. *In an orthomodular lattice two elements b and c are said to be compatible if the sublattice generated by* $\{b,b^{\perp},c,c^{\perp}\}$ *is distributive.*

DEFINITION A.5. *The set of elements in a lattice L which are compatible with all other elements is called the center of the lattice and denoted by C(L).*

An element of $C(L)$ is called a central element. A filter in $C(L)$ is called a central filter. For ortholattices, $C(L)$ is always a Boolean lattice. In a complete orthomodular lattice L there exists a unique least central element $e(a)$ called *the central cover of a* such that $a < e(a)$ for any $a \in L$.

LEMMA A.6. *In a complete orthomodular lattice:*

i) $e(a \vee b) = e(a) \vee e(b);$
ii) *if* $z \in C(L)$, *then* $e(z \wedge a) = z \wedge e(a).$

(See Maeda 1970, 5.11.)

DEFINITION A.7. *A filter F is a set of elements in a lattice L satisfying the following two conditions:*

i) *if $x, y \in F$, then $x \wedge y \in F$;*
ii) *if $x \in F$, then $x \vee y \in F$, for any y in L.*

A filter is *proper* if $0 \notin F$. In this exposition, filters are always proper. A filter F is *maximal* if there is no other filter F' such that F is a proper subset of F'. A maximal filter is also called an ultrafilter.

Since in von Neumann's formulation of quantum mechanics physical systems are represented by Hilbert spaces and subspaces and projection operators play a fundamental role in the formulation of the theory, one is led to characterize the mathematical structure of the theory in purely lattice theoretical terms on the basis of the following theorem.

THEOREM A.8. *Let H be a Hilbert space, the set $L_c(H)$ of all closed subspaces of H forms an irreducible complete atomic orthomodular lattice with the covering property.*

(See, for example, Maeda 1970, theorem 34.8.)

A lattice with 0 is *atomic* when for every nonzero element x there exists an atom a such that $a \leq x$. An atom of L is an element a such that $0 \lessdot a$. A dual atom (or antiatom) is an element x such that $x \lessdot 1$. A lattice (with 0) has the *covering property* if $a \wedge x = 0$ implies $x \lessdot x \vee a$ for any atom a and element x. A complete atomic orthomodular lattice with the covering property is called a *quantum lattice*.

Lattice Formulation of Lüders's Rule. The fundamental algorithm of quantum mechanics states that a measurement of the quantity A in a system in the state Q gives the result a_r with probability $\text{Prob}(a_r) = \text{tr}(QP_r)$. Lüders's rule states that after a measurement with result represented by P, the state of the system is given by $PQP = (\text{tr}QP)P$. Let $E_1(PQP)$ denote the support of the projection PQP, that is, $E_1(PQP)$ is the orthocomplement of the null space of PQP, denoted by $E_0(PQP)$.

THEOREM A.9. *For all projections P, Q in H, $E_0(PQP) = P^{\perp} \vee (P \wedge Q^{\perp})$.*

For the proof see Hardegree (1976).

COROLLARY A.10. $E_0^{\perp}(PQP) = E_1(PQP) = P \wedge (P^{\perp} \vee Q)$.

In our abstract lattice theoretical framework then, the transformation $L(q,p) = p \wedge (p^{\perp} \vee q)$ represents Lüders's transformation, where q is an atom and p represents an arbitrary element in L.

DEFINITION A.11. *A set $B = \{a_1, \ldots, a_k\}$ of compatible atoms in a ortholattice L is a basis for an element $q \in L$, if $q = a_1 \vee \ldots \vee a_k$, that is, B generates q, but no subset of B generates q.*

If every basis of q has the same cardinality then we call k the dimension of q, and write $\dim(q) = k$.

Part B: The Structure of Orthomodular Lattices with Two Blocks. I am interested in a description of the structure of complete atomic orthomodular lattices obtained by pasting together two atomic Boolean lattices M and N. Atomic Boolean lattices M and N are called the components or *blocks* of the resulting orthomodular structure. Pasting involves the identification of a Boolean subalgebra B_0 in M with an isomorphic copy (of B_0) in N. A lattice $L = \text{Lat}(M,N)$ is then constructed where the corresponding elements (under the isomorphism) are identified. The center of the lattice, $C(L)$, is the Boolean subalgebra in L resulting from the identification of the two isomorphic copies of B_0. My aim is to characterize atomic orthomodular lattices of two blocks with the covering property in terms of the structure of the center. The first restriction we have to impose on B_0 is that it is (what I will call) a *FI*-Boolean subalgebra. A Boolean subalgebra of an arbitrary Boolean algebra M is a *FI-subalgebra* (relative to M) if it is a complete Boolean subalgebra that can be represented as the union of a filter and an ideal of M.

An extension of a filter F is a (proper) filter F_e properly containing F. An extension of a Boolean subalgebra B is a Boolean algebra B_e properly containing B. An *extension F_e of a filter F is maximal* in a Boolean algebra B if F_e is a maximal filter in B.

In the proofs of the lemmata below, I use freely two properties of atomic Boolean algebras: (i) if $a < b$, then there exists an atom p such that $p \leq a$ and $p < b$; (ii) $x < y^{\perp}$ if and only if $x \wedge y = 0$. Property (i) says that the algebra is atomistic; property (ii) says that for Boolean algebras, complementation is pseudo-complementation.

DEFINITION B.1. *C(L) is full if there are no (proper) extensions of C(L) (as a Boolean subalgebra) in M and N.*

Let B be a *FI*-Boolean subalgebra (of M) that can be represented as the union of a principal filter and a principal ideal, $B = [x) \cup (x^{\perp}]$, where $[x)$ is a maximal filter in B, that is, x is an atom of B and $x^{\perp}$ is a dual atom of B. The following lemma provides necessary and sufficient conditions for the principal filter $[x)$ (and thus for B) to have only maximal extensions in M.

LEMMA B.2. *For a,b atoms of M, $x = a \vee b$ if and only if $[x)$ has only maximal extensions in M.*

Proof: Suppose $x = a \vee b$. Let F be an extension of $[x)$, then there is $y \in F$ such that $y \notin [x)$. Notice that $y \notin (x^{\perp}]$ since $y \in (x^{\perp}]$ if and only if $x \wedge y = 0$, and this would imply $x \wedge y = 0 \in F$, contradicting the assumption that F is a proper filter. Now, $x \wedge y = x$ iff $x \leq y$. Since $x \nleq y$, then $x \wedge y \neq x$, hence $x \wedge y < x$ and thus $[x \wedge y)$ is an extension of $[x)$ including y. Since $x \wedge y = a$ or $x \wedge y = b$, there are only (two) maximal extensions of $[x)$ (in M). Now, for the converse, if there are only maximal extensions of $[x)$ then x covers a for some atom a of M. This can only be the case if $x = a \vee b$, for a,b atoms of M.

The above lemma shows that if $C(L)$ can be represented as the union of a principal filter and a principal ideal of M, and if $C(L)$ is maximal, then there are no proper extensions of $C(L)$. Using the hypothesis of completeness, one can show that every *FI*-subalgebra of M can be represented in this way. Furthermore, this representation is unique (relative to M). This is the content of the following lemmata.

LEMMA B.4. *Let B be a FI-Boolean subalgebra of M. Let I_B be the ideal generated by the atoms of B which are atoms of M. Element I_B is a maximal ideal in B.*

Proof: Suppose I_B is not maximal, then there is an extension I_B' with an atom a of B which is not an atom of M. Since a is not an atom of M, then there is an atom b in M such that $b < a$. But then b would also belong to I_B' contradicting the assumption that a is an atom of B.

We can then represent B as the filter-ideal pair $B = \langle F_B, I_B \rangle$, where F_B is the filter dual to I_B. We are interested in the case $B = C(L)$ $(\neq M)$. For brevity we will use C instead of $C(L)$ in the lemmata below.

LEMMA B.5. *If a is an atom of M, then e(a) is an atom of C.*

Proof: Let $x \in C$ and $0 < x \leq e(a)$. Then $e(x \wedge a) = x \wedge e(a) = x > 0$ (see lemma [A.6]). Since $e(y) = 0$ iff $y = 0$, then $x \wedge a \neq 0$, and since a is an atom of M (an atom of L) $x \wedge a = a$, hence $a \leq x$, whence $e(a) \leq e(x) = x$. Thus $e(a) = x$.

The result of Lemma (B.4) can be sharpened using completeness and Lemma (B.5). Let *A(F) be the set of atoms of M which are not in I_c*. I show that all elements of $A(F)$ have a common cover. By completeness $\vee \{x; x \in A(F)\}$ exists. Suppose $y, z \in A(F)$ and $e(y) \neq e(z)$. Since y, z are not in I_c then y, z as well as $e(y), e(z)$ are in F_c—since F_c is a filter and $x \leq e(x)$ for all x—and thus $0 \neq e(y) \wedge e(z) \in F_c$ (since F_c is a proper filter). The assumption that $e(y) \neq e(z)$ implies that $e(y) < e(z)$ or $e(y) \wedge e(z) < e(z)$—or alternatively $e(z) < e(y)$ or $e(y) \wedge e(z) < e(y)$—contradicting our earlier result that $e(z)$ and $e(y)$ are atoms of C. Let us denote by *f_0 the common cover for the elements of A(F)*, that is, $f_0 = e(x)$, for all $xA(F)$.

LEMMA B.6. $\vee\{x;\ x \in A(F)\} = f_0$.

Proof: Clearly $\vee\{x;\ x \in A(F)\} \leq f_0$. Suppose $\vee\{x;\ x \in A(F)\} < f_0$, then there is an atom a of M such that $a < f_0$ and $a \notin A(F)$. But $a \notin I_c$ since $a < f_0$ and f_0 is an atom of C, contradicting our assumption.

The element f_0 can be called the "floating atom". The element *f_0 is the only atom in the center which is not an atom of the lattice*. Suppose P is a proper subset of $A(F)$. Then it follows from Lemma (B.6) that $\vee\{x;\ x \in P\} \notin C$. If $A(F)$ has more than two elements then the center C (of L) admits nonmaximal extensions. This proves the following corollary.

COROLLARY B.7. *$A(F)$ has more than two elements if and only if $C(L)$ is not full.*

REFERENCES

Beltrametti, E. G. and Cassinelli, G. (1981), *The Logic of Quantum Mechanics*. Reading, MA: Addison-Wesley.

Bub, J. (1979), "The Measurement Problem of Quantum Mechanics", in T. di Francia (ed.), *Physics*. Reading, MA: Addison-Wesley, pp. 71–124.

Friedman, M. and Putnam, H. (1978), "Quantum Logic, Conditional Probability, and Interference", *Dialectica 32*: 305–315.

Hardegree, G. M. (1980), "Micro-States in the Interpretation of Quantum Theory", in P. D. Asquith and R. N. Giere, (eds.), *PSA 1980*, vol. 1. East Lansing: Philosophy of Science Association, pp. 43–54.

Herbut, F. (1969), "Derivation of the Change of State in Measurement from the Concept of Minimal Measurement", *Annals of Physics 55*: 271–300.

Kochen, S. (1979), "The Interpretation of Quantum Mechanics", Department of Mathematics, Princeton University. Unpublished manuscript.

Lüders, G. (1951), "Über die Zustansänderung durch den Messprozess", *Annalen der Physik 8*: 322–328.

Maeda, F. and Maeda, S. (1970), *Theory of Symmetric Lattices*. Berlin: Springer-Verlag.

Martínez, S. (1988), "Minimal Disturbance in Quantum Logic", in A. Fine and J. Leplin (eds.), *PSA 1988*, vol. 1. East Lansing: Philosophy of Science Association, pp. 83–88.

———. (1990), "A Search for the Physical Content of Lüders's Rule", *Synthese 82*: 97–125.

Mielnik, B. (1968), "Geometry of Quantum States", *Communications in Mathematical Physics 9*: 55–80.

Piron, C. (1976), *Foundations of Quantum Physics*. Reading, MA: Benjamin.

Stairs, A. (1982), "Discussion: Quantum Logic and the Lüders Rule", *Philosophy of Science 49*: 422–436.

van Fraassen, B. C. and Hooker, C. A. (1976), "A Semantic Analysis of Niels Bohr's Philosophy of Quantum Theory", in W. L. Harper and C. A. Hooker (eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*. Vol. 3, *Foundations and Philosophy of Statistical Theories in the Physical Sciences*. Dordrecht: Reidel, pp. 221–241.

von Neumann, J. ([1932] 1955), *Mathematical Foundations of Quantum Mechanics*. Translated by R. T. Beyer. Originally published as *Mathematische Grundlagen der Quantenmechanik*. (Berlin: Springer-Verlag.) Princeton: Princeton University Press.