An Operational Approach to Quantum Probability

E. B. Davies★

Massachusetts Institute of Technology, Cambridge, U.S.A.

J. T. Lewis★★

Institute for Advanced Study, Princeton, U.S.A.

Abstract. In order to provide a mathmatical framework for the process of making repeated measurements on continuous observables in a statistical system we make a mathematical definition of an instrument, a concept which generalises that of an observable and that of an operation. It is then possible to develop such notions as joint and conditional probabilities without any of the commutation conditions needed in the approach via observables. One of the crucial notions is that of repeatability which we show is implicitly assumed in most of the axiomatic treatments of quantum mechanics, but whose abandonment leads to a much more flexible approach to measurement theory.

§ 1. Introduction

It is well known that Kolmogorov's measure-theoretic formulation of classical probability theory [1] can be expressed in a way which emphasizes its similarities with von Neumann's Hilbert space formulation of quantum theory (see for example [2]); the 'observables' of probability theory are the random variables and the 'states' are the probability measures. Many attempts have been made to extend the statistical interpretation of quantum theory by trying to construct analogues of more complicated objects of probability theory, such as joint probability distributions and conditional expectations. The difficulty with these approaches is that it invariably turns out that the relevant objects exist only in very special circumstances; joint distributions in the sense of Urbanik [3] and Varadarajan [4] exist if and only if the observables commute; a conditioning in the sense of Umegaki [5] and Nakamura and Turumaru [6] exists if and only if the observable has discrete spectrum.

An alternative approach to statistical concepts in quantum theory is provided by von Neumann's ideas on repeated measurements [7]. Accepting that a measurement of an observable causes a transformation of the state and that it is possible to perform a sequence of measurements on a state, a major problem is the difficulty of determining the possible transformations of states associated with the measurement of an observable. We show that von Neumann's repeatability hypothesis is too restrictive and that this is connected with his difficulties over the continuous spectrum. At the same time we generalise to continuous observables the notion of operation, described for discrete observables by Schwinger [8], for example, and introduced in the context of algebraic quantum field theory by Haag and Kastler [9]. The theory is developed at a level of generality which allows both the Kolmogorov theory and the von Neumann theory to be considered as special cases.

We would like to thank M. Kac, E. Nelson, M. Gerzon, F. Holroyd, R. Haag and the referee for helpful discussions, encouragement and advice concerning the paper.

§ 2. Probabilistic Concepts in Quantum Theory

In the standard formulation of quantum theory a state is a positive operator such that trace $[\varrho]$ is finite, on a complex Hilbert space $\mathscr{H}$. An observable is a self-adjoint operator A on $\mathscr{H}$ or, equivalently by the spectral theorem, a projection-valued measure $a(\cdot)$ on the σ-field of Borel subsets of the real line $\mathbb{R}$, related to A by the formula $A = \int_{\mathbb{R}} \lambda a(d\lambda)$.
The probability $P(\varrho, A, E)$ that in the state ϱ the observable A should take values in E is given by

$$P(\varrho, A, E) = \text{trace} [\varrho a(E)]/\text{trace} [\varrho] .$$

According to von Neumann [7, p. 220] an observable A can be measured with absolute precision if and only if its spectrum is purely discrete; for such observables it is straightforward to give an operationally motivated account of conditioning and joint distributions within the von Neumann theory.

Let A be a self-adjoint operator with purely discrete spectrum $\{x_1, x_2, ...\}$ and associated projections $P_1, P_2, ...$ so that $a(E) = \sum_{x_i \in E} P_i$
for all $E \subseteq \mathbb{R}$. For any eigenvalue x_i it is conventionally supposed that a measurement which gives the value x_i transforms the state ϱ into the state $P_i \varrho P_i$. If x_i is non-degenerate this follows from the repeatability hypothesis [7, p. 214] while if x_i is degenerate some extra principle of least interference [10–12] is also needed. From the elementary transfor-

mations $\varrho \to P_i \varrho P_i$ we can build up the transformation $\varrho \to \varrho^{\sharp} = \sum\limits_{i=1}^{\infty} P_i \varrho P_i$ which describes what happens when A is measured but the fractions in which the various values $x_1, x_2, \ldots$ occur are not separated [7, p. 347]. Nakamura and Umegaki [13] point out that the mapping $\varrho \to \varrho^{\sharp}$ is a *conditional expectation* in the sense of Umegaki [5] and hence is a generalisation of the classical conditional expectation. In the same spirit we can define the conditioning of the observable $B = \int\limits_{\mathbb{R}} \lambda b(d\lambda)$. Defining

$$b^{\sharp}(F) = \sum_{i=1}^{\infty} P_i b(F) P_i$$

we have $\text{trace}\,[\varrho b^{\sharp}(F)] = \text{trace}\,[\varrho^{\sharp} b(F)]$ so $F \to \text{trace}\,[\varrho b^{\sharp}(F)]/\text{trace}\,[\varrho]$ is a probability measure on $\mathbb{R}$. We call $b^{\sharp}(\cdot)$ the observable *conditioned* by the measurement of A. Notice that $b^{\sharp}(\cdot)$ is a positive operator-valued measure which is not projection-valued unless A and B commute, in which case $b^{\sharp}(\cdot) = b(\cdot)$. Thus the introduction of the concept of conditioned observables requires an extension of the class of positive operator-valued measures considered as observables.

An operationally motivated definition of the joint distribution of A and B can be given when A has purely discrete spectrum as above. The map

$$F \times E \to c(F \times E) = \sum_{x_i \in E} P_i b(F) P_i$$

can be shown to extend to a unique positive operator-valued measure on $\mathbb{R}^2$, which we call the *joint distribution* of $b(\cdot)$ following $a(\cdot)$. This measure has marginal distributions

$$c(\mathbb{R} \times E) = a(E), \qquad c(F \times \mathbb{R}) = b^{\sharp}(F).$$

In this connection we note that a joint distribution of $a(\cdot)$ and $b(\cdot)$ in the sense of Varadarajan [4] exists if and only if A and B commute; in this case $b = b^{\sharp}$ and c is a projection-valued measure on $\mathbb{R}^2$ which coincides with the joint distribution in Varadarajan's sense. It is convenient to introduce the notation

$$\mathscr{E}(E)\varrho = \sum_{x_i \in E} P_i \varrho P_i.$$

For each $E \subseteq \mathbb{R}$ the mapping $\varrho \to \mathscr{E}(E)\varrho$ is a linear mapping sending states into states and satisfying

(i) trace $[\mathscr{E}(E)\varrho] = \text{trace}\,[\varrho a(E)]$;

(ii) for every countable disjoint family of Borel sets

$$E_1, E_2, \ldots \mathscr{E}\left(\bigcup_{i=1}^{\infty} E_i\right)\varrho = \sum_{i=1}^{\infty} \mathscr{E}(E_i)\varrho.$$

The adjoint maps on the observables are given by

$$\mathscr{E}(E)^* B = \sum_{x_i \in E} P_i B P_i$$

for all bounded operators B. In terms of $\mathscr{E}$ we have the formulae

$$\varrho^\sharp = \mathscr{E}(\mathbb{R})\,\varrho, \qquad b^\sharp(F) = \mathscr{E}(\mathbb{R})^* b(F),$$

$$c(F \times E) = \mathscr{E}(E)^* b(F).$$

The constructions we have used above depend crucially on the assumption that the spectrum of A is purely discrete. If we were not interested in observables with continuous spectrum or were content to adopt von Neumann's device [7, p. 220] of approximating continuous observables with step-function operators having purely discrete spectrum we need go no further. However, there is evidence for believing that notions of conditioning are particularly interesting when the observable has continuous spectrum [14; 15]; moreover von Neumann himself was not entirely satisfied with the step-function device and regarded it as a temporary expedient [7, p. 223].

The basis of the measurement theory we have used up to this point, and the source of the difficulties concerning the continuous spectrum, is the repeatability hypothesis [7, p. 214].

(M.) If the physical quantity is measured twice in succession in a system S then we get the same value each time.

To begin with we abandon this hypothesis completely. Given an observable $a(\cdot)$ an arbitrary family $\mathscr{E}(\cdot)$ of linear mappings which send the states into themselves and which satisfy (i) and (ii) above is called an instrument which determines the observable $a(\cdot)$ and it is used to define a conditioning and a joint distribution.

We make essential use of the following facts. The space of self-adjoint trace class operators on $\mathscr{H}$ forms a real Banach space V under the trace norm [16] and the states form a positive norm closed cone K in V. The Banach space V^* dual to V can be identified with the space of bounded self-adjoint operators on $\mathscr{H}$ and the dual cone K^* is the set of positive operators. The identity operator 1 defines a functional on V such that for all states $\varrho \in K$, $\langle 1, \varrho \rangle = \|\varrho\|$. With this example in mind we start the general theory.

§ 3. General Theory

A *state space* (or complete base norm space [17]) is a triple (V, K, τ) consisting of a real Banach space V, a closed cone K which generates V and a linear functional τ on V such that $\langle \tau, v \rangle = \|v\|$ for all $v \in V$; we

also suppose that for all $v \in V$

$$\|v\| = \inf\{\|v_1\| + \|v_2\| : v = v_1 - v_2, v_1 \in K, v_2 \in K\}.$$

V is endowed with a partial ordering by putting $v \leqq w$ if and only if $(w - v) \in K$. A linear functional φ on V is said to be *positive* if $\langle \varphi, v \rangle \geqq 0$ for all $v \in K$. Every positive linear functional on a state space is necessarily continuous and the set of all positive linear functionals forms a closed cone K^* in the Banach dual space V^* endowed with the weak* topology. A linear mapping $T: V \to V$ is said to be positive if $T(K) \subseteq K$. A positive linear mapping on a state space is necessarily bounded [18] and the set of all positive linear mappings forms a closed cone $\mathscr{L}^+(V, V)$ in the topological linear space $\mathscr{L}(V, V)$ of all bounded linear mappings of V into itself, endowed with the strong operator topology.

A *state* v is a non-zero element of K. An *observable* is a triple $(X, \mathscr{A}, a)$ consisting of a set X with a σ-field $\mathscr{A}$ of subsets of X, and a mapping $a: \mathscr{A} \to V^*$ satisfying

(i) $0 \leqq a(E) \leqq a(X)$ for all $E \in \mathscr{A}$;

(ii) $a(X) = \tau$;

(iii) for each countable family $\{E_i\}$ of pairwise disjoint sets in $\mathscr{A}$,

$$a\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} a(E_i),$$ where the right-hand side converges in the weak* topology of V^*. The Borel space $(X, \mathscr{A})$ is called the *value space* of the observable. An *instrument* is a triple $(X, \mathscr{A}, \mathscr{E})$ consisting of a Borel space $(X, \mathscr{A})$ and a mapping $\mathscr{E}: \mathscr{A} \to \mathscr{L}^+(V, V)$ satisfying

(i) $\langle \tau, \mathscr{E}(X) v \rangle = \langle \tau, v \rangle$ for all $v \in V$;

(ii) for each countable family $\{E_i\}$ of pairwise disjoint sets in $\mathscr{A}$,

$$\mathscr{E}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathscr{E}(E_i),$$ where the right-hand side converges in the strong operator topology. The instrument (observable) is said to be *discrete* if its value space is discrete.

Theorem 1. *To every instrument $\mathscr{E}$ on the value space $(X, \mathscr{A})$ there is a unique observable $a(\cdot)$ such that*

$$\langle \tau, \mathscr{E}(E) \varrho \rangle = \langle a(E), \varrho \rangle$$

for all $\varrho \in V$ and $E \in \mathscr{A}$. Every observable $a(\cdot)$ on $(X, \mathscr{A})$ is determined in such a way by at least one instrument.

Proof. For each $E \in \mathscr{A}$ the above formula defines $a(E)$ as a positive linear functional on V, that is as an element of K^*. It is easy to verify that $E \to a(E)$ is a measure. On the other hand given any observable $a(\cdot)$ on $(X, \mathscr{A})$ choose any countable partition $\{E_i\}$ of X into pairwise disjoint Borel sets and any sequence $\{v_i\}$ of states such that $\langle \tau, v_i \rangle = 1$

for $i = 1, 2, \ldots$ It is easy to check that the formula

$$\mathscr{E}(E)\, v = \sum_{i=1}^{\infty} \langle a(E \cap E_i), v \rangle\, v_i$$

defines an instrument $\mathscr{E}$ whose associated observable is $a(\cdot)$.

We recall that a *standard Borel space* is a Borel space which is Borel isomorphic to a Borel subset of some complete separable metric space. Most of the value spaces which arise in physical applications are standard Borel spaces and their properties have been studied extensively [19].

Theorem 2. *Let $\mathscr{E}$ and $\mathscr{F}$ be instruments on the standard Borel spaces X and Y. Then there exists one and only one instrument $\mathscr{G}$ on $X \times Y$ such that $\mathscr{G}(E \times F) = \mathscr{E}(E)\, \mathscr{F}(F)$ for all Borel subsets $E \subseteq X$ and $F \subseteq Y$.*

Note. The instrument $\mathscr{G}$ is called the *composition* of $\mathscr{E}$ following $\mathscr{F}$ and is denoted by $\mathscr{E} \circ \mathscr{F}$.

Proof. By the classification theorems for standard Borel spaces [19] every standard Borel space is Borel isomorphic to some zero-dimensional separable compact Hausdorff space, and so there is no loss of generality in supposing that both X and Y are both zero-dimensional separable compact Hausdorff spaces.

We denote by $\mathscr{B}(Z)$ the Banach space (under the supremum norm) of all bounded real-valued Borel functions on Z, by $C(Z)$ the space of continuous real-valued functions on Z, and by $\mathscr{A}$ the dense sub-algebra of $C(Z)$ consisting of functions taking only a finite number of distinct values.

As Z is compact and zero-dimensional, every f in can be written in at least one way in the form

$$f(\cdot) = \sum_{r=1}^{n} \alpha_r \chi_{A_r}(\cdot)\, \chi_{B_r}(\cdot)$$

where the Borel sets $A_r \subseteq X$ and $B_r \subseteq Y$ are all open and closed sets; if $f \geq 0$ we can choose $\alpha_r \geq 0$ for all r. It is easy to show that there is a well-defined positive linear map T from $\mathscr{A}$ to $\mathscr{L}(V, V)$ given by the formula

$$Tf = \sum_{r=1}^{n} \alpha_r \mathscr{E}(A_r)\, \mathscr{F}(B_r).$$

From the definitions of the norms on the two spaces it follows that T is continuous with

$$\| T \| = \| T(1) \| \leq \| \mathscr{E}(X) \|\, \| \mathscr{F}(Y) \| \leq 1,$$

and so T has a unique extension to a positive continuous linear map (which we also denote by T) from $C(Z)$ to $\mathscr{L}(V, V)$. Now for each $v \in V$, $\varphi \in V^*$, the formula

$$\langle \varphi, (Tf)\, v \rangle = \int\limits_X f(x)\, \mu_{\varphi, v}(dx)$$

defines a bounded measure $\mu_{\varphi, v}$ on Z with

$$\| \mu_{\varphi, v} \| \leqq \| T \| \, \| \varphi \| \, \| v \| \leqq \| \varphi \| \, \| v \|\,,$$

and so the same formula defines a positive linear map

$$T' : \mathscr{B}(Z) \to \mathscr{L}(V, V^{**}) \quad \text{with} \quad \| T' \| \leqq 1$$

such that T' extends T. We have to show that the image of $\mathscr{B}(Z)$ under T' lies in $\mathscr{L}(V, V)$.

Let $\{f_n\}$ be an increasing sequence of functions in $\mathscr{B}(Z)$ having a pointwise limit $f \in \mathscr{B}(Z)$, and let $v \in V$ be such that $(T'f_n)\, v \in V \subsetneqq V^{**}$ for all n; then as $\{(T'f_n)\, v\}$ is increasing and bounded in norm it converges in norm to some $v' \in V$. Then for all $\varphi \in V^*$

$$\langle \varphi, (T'f)\, v \rangle = \int\limits_X f(x)\, \mu_{\varphi, v}(dx)$$

$$= \lim_{n \to \infty} \left\{ \int\limits_X f_n(x)\, \mu_{\varphi, v}(dx) \right\}$$

$$= \lim_{n \to \infty} \langle \varphi, (T'f_n)\, v \rangle = \langle \varphi, v' \rangle\,,$$

so that $(T'f)\, v = v' \in V$. Therefore for each $v \in V$ the family of f in $\mathscr{B}(Z)$ such that $(T'f)\, v \in V$ is a monotone family containing $C(Z)$, and so is equal to $\mathscr{B}(Z)$. Thus T maps $\mathscr{B}(Z)$ into $\mathscr{L}(V, V)$. We now define $\mathscr{G}$ on Z by

$$\mathscr{G}(G) = T'(\chi_G) \quad \text{for each Borel set} \quad G \subseteq Z\,;$$

we see that $\mathscr{G}$ is an $\mathscr{L}^+(V, V)$-valued measure on Z such that for all open and closed sets $E \subseteq X$ and $F \subseteq Y$

$$\mathscr{G}(E \times F) = T(\chi_E \chi_F) = \mathscr{E}(E)\, \mathscr{F}(F)\,.$$

The validity of this formula for all Borel sets $E \subseteq X$ and $F \subseteq Y$ now follows from the monotone properties of $\mathscr{E}, \mathscr{F}$ and $\mathscr{G}$.

The uniqueness of the map $\mathscr{G}$ having these properities follows from the fact that the sets in Z of the form $E \times F$ generate the σ-field of all Borel sets in Z.

The structure we have described has a statistical interpretation in terms of ensembles. It is also natural to interpret it in terms of filtering operations performed on a beam consisting of *non-interacting* copies of a system. A state v corresponds to a beam and $\langle \tau, v \rangle$ is a measure of the strength of the beam. An instrument $\mathscr{E}$ measuring a physical quantity taking values in X corresponds to a family of filters indexed by the subsets in a family $\mathscr{A}$; the filter corresponding to $\mathscr{E}(E)$ allows a copy of the system to pass if its observed value lies in E and in doing so transforms the input state v into the output state $\mathscr{E}(E) v$. The various filters in the family are related by the additivity condition: if E_1 and E_2 are disjoint the output state $\mathscr{E}(E_1 \cup E_2) v$ is the same as the mixture of $\mathscr{E}(E_1) v$ and $\mathscr{E}(E_2) v$. For each $\mathscr{E}$ and v the mapping $E \to \langle \tau, \mathscr{E}(E) v \rangle / \langle \tau, v \rangle$ is a probability measure on X; the right hand side is the ratio of the strength of the output beam to the strength of the input, and hence can be interpreted as the probability that in the state ϱ the physical quantity observed takes values in E. Successive application of two families of filters corresponds to the composition of the instruments.

Let $(X, \mathscr{E})$ be an instrument which determines the observable (X, a) and let $(Y, \mathscr{F})$ be an instrument which determines the observable (Y, b), where X and Y are standard Borel spaces, and let $(Y \times X, \mathscr{F} \circ \mathscr{E})$ be the composition of $\mathscr{F}$ following $\mathscr{E}$. The V^*-valued measure $F \to \mathscr{E}(X)^* b(F)$ is an observable, which we call the observable $b(\cdot)$ *conditioned* by the measurement of $a(\cdot)$ with the instrument $\mathscr{E}$. The probability measure $M \to \langle \tau, (\mathscr{F} \circ \mathscr{E})(M) v \rangle / \langle \tau, v \rangle$ on $Y \times X$ is that of an observable $c(M) = (\mathscr{F} \circ \mathscr{E})(M)^* \tau$ which we call the *joint distribution* of $\mathscr{F}$ following $\mathscr{E}$. We have the following generalisation of the results of the previous section.

Theorem 3. *Let $\mathscr{E}$ and $\mathscr{F}$ be instruments on the standard Borel spaces X and Y with associated observables $a(\cdot)$ and $b(\cdot)$, respectively. Then the joint distribution of $\mathscr{F}$ following $\mathscr{E}$ is an observable $c(\cdot)$ on $Y \times X$ whose marginal distributions satisfy*

$$c(Y \times E) = a(E), \quad c(F \times X) = \mathscr{E}(X)^* b(F)$$

Proof. Existence of $c(\cdot)$ follows from Theorems 1 and 2. Since $\langle c(Y \times E), \rangle = \langle \tau, (\mathscr{F} \circ \mathscr{E})(Y \times E) v \rangle$

$$= \langle \tau, \mathscr{F}(Y) \mathscr{E}(E) v \rangle = \langle \tau, \mathscr{E}(E) v \rangle = \langle a(E), v \rangle$$

and $\langle c(E \times X), v \rangle = \langle \tau, \mathscr{F}(F) \mathscr{E}(X) v \rangle$

$$= \langle b(F), \mathscr{E}(X) v \rangle = \langle \mathscr{E}(X)^* b(F), v \rangle$$

the marginal distributions have the stated properties.

It is clear that in general the compositions $\mathscr{E} \cdot \mathscr{F}$ and $\mathscr{F} \circ \mathscr{E}$ will have quite different properties; when they happen to be equal we say that $\mathscr{E}$ and $\mathscr{F}$ are *compatible*. Generalising the repeatability hypothesis $(M.)$ to arbitrary value spaces we now say that any instrument $(X, \mathscr{E})$ is *weakly repeatable* if

(R) $\langle \tau, \mathscr{E}(E)\,\mathscr{E}(F)\,v \rangle = \langle \tau, \mathscr{E}(E \cap F)\,v \rangle$

for all Borel subsets $E, F \subseteq X$ and all $v \in V$. We say $\mathscr{E}$ is *repeatable if*

(R') $\mathscr{E}(E)\,\mathscr{E}(F)\,v = \mathscr{E}(E \cap F)\,v$ for all Borel subsets $E, F \subseteq X$ and all $v \in V$.

The existence of repeatable instruments is doubtful even in the case of standard quantum theory, but we can show quite generally that given a non-discrete observable there exists an instrument which determines it and which satisfies an approximate repeatability condition. Further results of this type are contained in [14].

Let $\mathscr{E}$ be an instrument on a separable metric space X with metric d. Let E be a Borel subset of X and for each $\varepsilon > 0$ let E_ε denote the Borel set $\{x \in X : d(x, E) < \varepsilon\}$. We say that $\mathscr{E}$ is *ε-repeatable* if for all $v \in K$ and for all Borel subsets $E \subseteq X$ we have

(ER) $\langle \tau, \mathscr{E}(E_\varepsilon)\,\mathscr{E}(E)\,v \rangle = \langle \tau, \mathscr{E}(E)\,v \rangle$.

In order to formulate the existence theorem we make some definitions concerning observables.

Let A be an observable on a Borel space X. We say that a Borel subset E of X is of A-measure zero if $\langle A(E), v \rangle = 0$ for all $v \in K$. A state v is said to be concentrated on $E \subseteq X$ if $\langle A(E), v \rangle = \langle \tau, v \rangle$. if v_i are concentrated on E_i respectively $(i = 1, 2, \ldots)$, where $\{E_i\}_{i=1}^{\infty}$ is a sequence of pairwise disjoint sets covering X; then

$$\langle A(E_i), v_i \rangle = \langle \tau, v_i \rangle ,$$

and

$$\langle A(X), v_i \rangle = \langle \tau, v_i \rangle ,$$

so that

$$\langle A(E_j), v_i \rangle = 0 \quad \text{for all} \quad i \neq j .$$

In the Hilbert space case a projection-valued measure A on a separable metric space X is such that for each Borel set E which is not of A-measure zero there is at least one state which is concentrated on E, namely $v = \xi \otimes \bar{\xi}$ where ξ is a unit vector in the range of $A(E)$. The example $A(E) = \mu(E)\,1$ where 1 is the identity operator and μ is a probability measure on X shows that this is not true in general.

Theorem 4. *Let A be an observable on a separable metric space X which is such that for each Borel set E which is not of A-measure zero there is at least one state v concentrated on E. Then for each $\varepsilon > 0$ there exists an ε-repeatable instrument which determines A.*

Proof. Given $\varepsilon > 0$ put $X = \bigcup_{n=1}^{\infty} E_n$ where $\{E_n\}$ is a sequence of pairwise disjoint Borel subsets of X having diameters less than ε. Let $\{v_n\}$ be a sequence of normalised states such that v_n is concentrated on E_n for all n for which E_n is not of A-measure zero. Put

$$\mathscr{E}(E)\, v = \sum_{n=1}^{\infty} \langle A(E \cap E_n), v \rangle\, v_n$$

for all $v \in K$ and every Borel subset E of X. Then

$$\langle \tau, \mathscr{E}(E)\, v \rangle = \sum_{n=1}^{\infty} \langle A(E \cap E_n), v \rangle \langle \tau, v_n \rangle = \left\langle \sum_{n=1}^{\infty} A(E \cap E_n), v \right\rangle$$

$$= \langle A(E \cap X), v \rangle = \langle A(E), v \rangle$$

so that $\mathscr{E}$ determines A. Furthermore

$$\langle \tau, \mathscr{E}(E_\varepsilon)\, \mathscr{E}(E)\, v \rangle = \sum_m \sum_n \langle A(E_\varepsilon \cap E_m), v_n \rangle \langle A(E \cap E_n), v \rangle$$

$$= \sum_m \sum_n \delta_{mn} \langle A(E \cap E_n), v \rangle = \langle \tau, \mathscr{E}(E)\, v \rangle,$$

since the E_n have diameter less than ε, and so

$$\langle \tau, \mathscr{E}(E_\varepsilon)\, \mathscr{E}(E)\, v \rangle = \langle \tau, \mathscr{E}(E)\, v \rangle.$$

§ 4. Classical Probability Theory

It is interesting to look in some detail at the particular case of the Kolmogorov model of probability theory. Suppose the sample space $(\Omega, \mathscr{M})$ is a standard Borel space and let V denote the space of real bounded Borel measures on Ω. If we define K as the cone of positive measures in V and τ as the functional $\langle \tau, \mu \rangle = \mu(\Omega)$ then (V, K, τ) is a state space. There is a natural one-one embedding of the space $\mathscr{B}(\Omega)$ of bounded Borel functions on Ω into V^* defined by

$$\langle f, \mu \rangle = \int_{\Omega} f(w)\, \mu(dw)$$

for all $\mu \in V$. This embedding takes positive functions f to positive elements of V^* and takes the function which is constantly one to the element

τ of V^*. Since V^* is a very abstract space we restrict our interest to the *measurable* observables, which we define as those taking their values in $\mathscr{B}(\Omega)$.

A *random variable* is defined as a Borel map $\alpha : \Omega \to X$ where $(X, \mathscr{A})$ is a Borel space, often the real line. We define the associated measurable observable $(X, \mathscr{A}, a)$ by putting $a(E) = \chi_{\alpha^{-1}E}$ for all $E \in \mathscr{A}$. This observable is multiplicative in the sense that $a(E)\, a(F) = a(E \cap F)$ for all $E, F \in \mathscr{A}$, and indeed this property characterises the observables associated with random variables, provided $(X, \mathscr{A})$ is a standard Borel space. For any normalised state (probability law) μ on Ω the *image law* is defined in the Kolmogorov model as the probability measure v on X given by $v(E) = \mu(\alpha^{-1}E)$ for all $E \in \mathscr{A}$; in our notation this is just $\langle a(E), \mu \rangle$ which is precisely the distribution of the observable $a(\cdot)$ in the state μ.

We next examine the instruments. If $(X, \mathscr{A})$ is a value space we define a *kernel* P as a positive real valued function such that $P(M, E, w)$ is defined for all $M \in \mathscr{M}$, $E \in \mathscr{A}$ and $w \in \Omega$, such that

  (i) $P(\Omega, X, w) = 1$ for all $w \in \Omega$;
  (ii) $E \to P(M, E, w)$ is σ-additive for all $M \in \mathscr{M}$ and $w \in \Omega$;
  (iii) $M \to P(M, E, w)$ is σ-additive for all $E \in \mathscr{A}$ and $w \in \Omega$;
  (iv) $w \to P(M, E, w)$ is Borel for all $M \in \mathscr{M}$ and $E \in \mathscr{A}$.

Theorem 5. *If P is a kernel the formula*

$$\{\mathscr{E}(E)\, \mu\}\,(M) = \int_\Omega P(M, E, w)\, \mu(dw)$$

defines an instrument $\mathscr{E}$ on the value space X. We call such instruments measurable; their observables are also always measurable.

Proof. By (iv) the integral is well defined and by (iii) the map $M \to \{\mathscr{E}(E)\, \mu\}\,(M)$ is a measure, so $\mathscr{E}(E)\, \mu \in V$ for all $\mu \in V$ and $E \in \mathscr{A}$. By (ii) $E \to \mathscr{E}(E)$ is σ-additive and by (i) $\langle \tau, \mathscr{E}(X)\, \mu \rangle = \langle \tau, \mu \rangle$ for all $\mu \in V$. If for each $E \in \mathscr{A}$ we define $a(E) \in \mathscr{B}(\Omega)$ by $a(E) = P(\Omega, E, w)$ then $a(\cdot)$ is a measurable observable and for all $v \in V$

$$\langle \tau, \mathscr{E}(E)\, \mu \rangle = \int_\Omega P(\Omega, E, w)\, \mu(dw) = \langle a(E), \mu \rangle$$

so $a(\cdot)$ is the observable determined by $\mathscr{E}$.

We say that $\mathscr{E}$ is *factorisable* if it is measurable for a kernel P can be written in the form $P(M, E, w) = A(E, w) \cdot B(M, w)$ where A and B are functions satisfying

  (i) $E \to A(E, w)$ is σ-additive for all $w \in \Omega$;
  (ii) $M \to B(M, w)$ is σ-additive for all $w \in \Omega$;

(iii) $w \to A(E, w)$ is Borel for all $E \in \mathscr{A}$;

(iv) $w \to B(M, w)$ is Borel for all $M \in \mathscr{M}$;

(v) $A(X, w) = B(\Omega, w) = 1$ for all $w \in \Omega$.

Theorem 6. *Let $\mathscr{E}$ be a factorisable instrument and let $a(\cdot)$ be its associated observable. Then there exists a positive linear map $T: \mathscr{B}(\Omega) \to \mathscr{B}(\Omega)$ satisfying*

(i) $T(1) = 1$;

(ii) *if φ_n is an increasing bounded sequence in $\mathscr{B}(\Omega)$ converging pointwise to φ then $T(\varphi_n)$ converges pointwise to $T(\varphi)$. $\mathscr{E}$ is determined by the equation*

$$\langle \varphi, \mathscr{E}(E) \mu \rangle = \langle a(E), T(\varphi), \mu \rangle$$

valid for all $\mu \in V$, $\mu \in \mathscr{A}$ and $\varphi \in \mathscr{B}(\Omega)$.

Proof. For each $E \in \mathscr{A}$ we define $a(E) \in \mathscr{B}(\Omega)$ by $a(E)(w) = A(E, w)$. Then for all $\mu \in V$ and $E \in \mathscr{A}$

$$
\begin{aligned}
\langle \tau, \mathscr{E}(E) \mu \rangle &= \int_{\Omega} P(\Omega, E, w)\, \mu(dw) \\
&= \int_{\Omega} A(E, w)\, B(\Omega, w)\, \mu(dw) \\
&= \int_{\Omega} A(E, w)\, \mu(dw) \\
&= \langle a(E), \mu \rangle
\end{aligned}
$$

so $a(\cdot)$ is the observable determined by $\mathscr{E}$. For $\varphi \in \mathscr{B}(\Omega)$ we define $T\varphi \in \mathscr{B}(\Omega)$ by

$$(T\varphi)(w) = \int_{\Omega} \varphi(\sigma)\, B(d\sigma, w)$$

so that $T(\chi_M)(w) = B(M, w)$. It is easy to show that T is a linear map from $\mathscr{B}(\Omega)$ to $\mathscr{B}(\Omega)$ satisfying (i) and (ii). Finally

$$
\begin{aligned}
\langle \varphi, \mathscr{E}(E) \mu \rangle &= \int_{\Omega} \varphi(\sigma) \left\{ \int_{\Omega} P(d\sigma, E, w)\, \mu(dw) \right\} \\
&= \int_{\Omega} \int_{\Omega} A(E, w)\, B(d\sigma, w)\, \varphi(\sigma)\, \mu(dw) \\
&= \int_{\Omega} (T\varphi)(w)\, A(E, w)\, \mu(dw) \\
&= \int_{\Omega} (T\varphi)(w)\, a(E)(w)\, \mu(dw) \\
&= \langle a(E) \cdot T(\varphi), \mu \rangle
\end{aligned}
$$

In accordance with Moy [20] we now say that a linear map $T : \mathscr{B}(\Omega) \to \mathscr{B}(\Omega)$ is a *conditional expectation* if it satisfies (i) and (ii) above and also (iii) for all $\varphi, \psi \in \mathscr{B}(\Omega)$, $T(\varphi \, T(\psi)) = (T\varphi)(T\psi)$. Moy shows that it is possible to construct a conditional expectation given a sub-σ-field of $\mathscr{M}$ and also a particular probability measure on Ω. The following theorem shows that random variables and conditional expectations are closely related to repeatable decomposible instruments.

Theorem 7. *Let $\mathscr{E}$ be a decomposible instrument on the standard Borel space $(X, \mathscr{A})$. Then $\mathscr{E}$ is weakly repeatable if and only if its observable $a(\cdot)$ is determined by a random variable $\alpha : \Omega \to X$ and $T(\chi_{\alpha^{-1}E}) = \chi_{\alpha^{-1}E}$ for all $E \in \mathscr{A}$. $\mathscr{E}$ is repeatable if and only if its observable is determined by a random variable $\alpha : \Omega \to X$ and*

$$T\{\chi_{\alpha^{-1}E} T(\varphi)\} = \chi_{\alpha^{-1}E} T(\varphi)$$

for all $\varphi \in \mathscr{B}(\Omega)$ and all $E \in \mathscr{A}$.

Proof. $\langle \tau, \mathscr{E}(E) \, \mathscr{E}(F) \, \mu \rangle = \langle a(E), \mathscr{E}(F) \, \mu \rangle$

$$= \langle a(F) \cdot T(a(E)), \mu \rangle$$

while $\langle \tau, \mathscr{E}(E \cap F) \, \mu \rangle = \langle a(E \cap F), \mu \rangle$

so $\mathscr{E}$ is weakly repeatable if and only if

$$a(F) \cdot T(a(E)) = a(E \cap F)$$

for all $E, F \in \mathscr{A}$. This equation is equivalent to

$$T(a(E)) = a(E) \quad \text{and} \quad a(E \cap F) = a(E) \cdot a(F)$$

for all $E, F \in \mathscr{A}$. Since X is a standard Borel space the second condition is equivalent to $a(\cdot)$ being determined by a random variable $\alpha : \Omega \to X$. Similarly repeatability is equivalent to the condition

$$a(E \cap F) \cdot T(\varphi) = a(F) \cdot T\{a(E) \, T(\varphi)\}$$

for all $E, F \in \mathscr{A}$ and all $\varphi \in \mathscr{B}(\Omega)$, which is itself equivalent to the pair of conditions

$$a(E \cap F) = a(E) \, a(F)$$

$$T\{a(E) \cdot T(\varphi)\} = a(E) \cdot T(\varphi)$$

for all $E, F \in \mathscr{A}$ and $\varphi \in \mathscr{B}(\Omega)$.

Corollary 8. $\mathscr{E}$ *is repeatable if a is determined by a random variable $\alpha: \Omega \to X$ and T is a conditional expectation associated with a sub-σ-field of $\mathscr{M}$ containing the σ-field $\alpha^{-1}(\mathscr{A})$.*

§ 5. Conditional Expectations on von Neumann Algebras

Conditional expectations on von Neumann algebras have been studied extensively in [5, 6, 13] as generalisations of Moy's conditional expectation and it is evident that they are closely connected with the probabilistic interpretation of quantum mechanics. We now show how to make this precise in the appropriate probability theory.

Let $\mathscr{A}$ be a von Neumann algebra on a Hilbert space $\mathscr{H}$ and let V be the space of self-adjoint ultraweakly continuous linear functionals on $\mathscr{A}$. Let K be the cone of positive linear functionals in V and let τ be the linear functional $\tau(v) = v(1)$. Then (V, K, τ) is a state space, K is just the set of normal positive linear functionals on $\mathscr{A}$ and V^* may be identified with the set of self-adjoint elements of $\mathscr{A}$, [21].

A normal conditional expectation on $\mathscr{A}$ is a self-adjoint positive linear map $T: \mathscr{A} \to \mathscr{A}$ such that

(i) $T(1) = 1$.

(ii) If $\{B_n\}$ is a monotone net of elements of $\mathscr{A}$ converging ultra-weakly to $B \in \mathscr{A}$ then $\{T(B)\}$ converges monotonely and ultraweakly to $T(B)$.

(iii) For all $B, C \in \mathscr{A}$, $T(T(B) C) = T(B) T(C) = T(B T(C))$.

If T is a normal conditional expectation then $T^2 = T$ and $\mathscr{B} = T(\mathscr{A})$ is equal to $\{B \in \mathscr{A} : T(B) = B\}$ and is a von Neumann subalgebra of $\mathscr{A}$. Let P be a projection-valued measure on any Borel space X whose range is contained in $\mathscr{A}$ and call

$$\{B \in \mathscr{A} : B P(E) = P(E) B \quad \text{for all Borel sets } E\}$$

the relative commutant of $P(\cdot)$ in $\mathscr{A}$.

Theorem 9. *Let T be a normal conditional expectation on $\mathscr{A}$ whose range is the relative commutant of $P(\cdot)$ in $\mathscr{A}$. Then the equation*

$$\langle \mathscr{E}(E) \varphi, B \rangle = \langle P(E) T(B), \varphi \rangle$$

where $\varphi \in V$ and $B \in \mathscr{A}$ defines an instrument $\mathscr{E}$ on X. $\mathscr{E}$ is repeatable, $P(\cdot)$ is the observable determined by $\mathscr{E}$, and T is the adjoint of the linear map $\mathscr{E}(X): V \to V$.

Proof. For each $\varphi \in K$ and $E \subseteq X$ the right hand side defines a normal positive linear functional of B so there exists $\psi \in K$ such that

$$\langle P(E)\, T(B), \varphi \rangle \quad \langle \psi, B \rangle$$

for all $B \in \mathscr{A}$. Defining $\mathscr{E}(E)\, \varphi = \psi$ one easily verifies that $\mathscr{E}$ is an instrument. The observable A determined by $\mathscr{E}$ is given by

$$\langle A(E), \varphi \rangle = \tau [\mathscr{E}(E)\, \varphi] = \langle \mathscr{E}(E)\, \varphi, I \rangle = \langle P(E), \varphi \rangle$$

for all $\varphi \in V$ so $A(E) = P(E)$ for all $E \subseteq X$. If $E, F \subseteq X$ and $\varphi \in V$ and $B \in \mathscr{A}$

$$\begin{aligned}
\langle \mathscr{E}(E)\, \mathscr{E}(F)\, \varphi, B \rangle &= \langle \mathscr{E}(F)\, \varphi, P(E)\, T(B) \rangle \\
&= \langle \varphi, P(F)\, T\{P(E)\, T(B)\} \rangle \\
&= \langle \varphi, P(F)\, P(E)\, T(B) \rangle \\
&= \langle \varphi, P(E \cap F)\, T(B) \rangle \\
&= \langle \mathscr{E}(E \cap F)\, \varphi, B \rangle
\end{aligned}$$

so $\mathscr{E}(E)\, \mathscr{E}(F) = \mathscr{E}(E \cap F)$. Finally for all $\varphi \in V$ and $B \in \mathscr{A}$

$$\begin{aligned}
\langle \mathscr{E}(X)^* B, \varphi \rangle &= \langle \mathscr{E}(X)\, \varphi, B \rangle = \langle P(X)\, T(B), \varphi \rangle \\
&= \langle T(B), \varphi \rangle
\end{aligned}$$

so $\mathscr{E}(X)^*(B) = T(B)$.

In the special case $\mathscr{A} = \mathscr{L}(\mathscr{H})$, V is the space of self-adjoint trace class operators on $\mathscr{H}$. if P is a projection-valued measure on the discrete space X then the formula

$$T(B) = \sum_{x \in X} P_x B P_x$$

defines a normal conditional expectation on $\mathscr{L}(\mathscr{H})$ whose range is the commutant of $\{P_x : x \in X\}$. The corresponding instrument $\mathscr{E}$, defined by

$$\mathscr{E}(E, v) = \sum_{x \in E} P_x v P_x$$

is easily seen to satisfy the following axioms.

(SR 1) $\mathscr{E}_x \mathscr{E}_y v = \delta_{xy} \mathscr{E}_y v$ for all $x, y \in X$ and $v \in V$.

(SR 2)   If $\langle \tau, \mathscr{E}_x v \rangle = \langle \tau, v \rangle$ then $\mathscr{E}_x v = v$.

(SR 3)   If $B \in K^*$ and $\langle B, \mathscr{E}_x v \rangle = 0$ for all $x \in X$ and all $v \in K$ then $B = 0$.

A discrete instrument which satisfies these conditions is said to be *strongly repeatable*, and we note that, unlike the axioms for conditional expectations, each of these axioms has a direct physical interpretation. (SR 1) is the repeatability hypothesis. (SR 2) is a principle of minimum disturbance, similar to that discussed by Lüders and others [10–12]. (SR 3) is a non-degeneracy condition and would be violated if for example the range of the instrument lay in a super-selection sector.

Theorem 10. *Let V be the space of self-adjoint trace class operators on a Hilbert space $\mathcal{H}$ and let X be a discrete value space. Then the formula*

$$\mathscr{E}(E)\, v = \sum_{x \in E} P_x v P_x$$

sets up a one-one correspondence between the discrete projection-valued measures P on X and the strongly repeatable instruments $\mathscr{E}$ on V. In particular a strongly repeatable instrument is uniquely determined by its observable, which is a projection-valued measure.

Proof. Given $x \in X$ let $F_x \subseteq K$ be the set $\{v \in K : \mathscr{E}_x v = v\}$. Then F_x is convex and closed in the trace-norm topology and we show that it is extreme in the sense that if $v_1, v_2 \in K$ and there exists α such that $0 < \alpha < 1$ and $\alpha v_1 + (1 - \alpha) v_2 \in F_x$ then $v_1, v_2 \in F_x$.

Since $\mathscr{E}$ is an instrument

$$0 \leqq \langle \tau, \mathscr{E}_x v_i \rangle \leqq \langle \tau, v_i \rangle, \quad (i = 1, 2),$$

and as

$$\alpha \langle \tau, \mathscr{E}_x v_1 \rangle + (1 - \alpha) \langle \tau, \mathscr{E}_x v_2 \rangle = \langle \tau, \mathscr{E}_x v \rangle = \langle \tau, v \rangle$$

$$= \alpha \langle \tau, v_1 \rangle + (1 - \alpha) \langle \tau, v_2 \rangle$$

we have

$$\langle \tau, \mathscr{E}_x v_i \rangle = \langle \tau, v_i \rangle, \quad (i = 1, 2),$$

and by (SR 2)

$$\mathscr{E}_x v_i = v_i, \quad (i = 1, 2).$$

If $\sum \alpha_n \xi_n \otimes \bar{\xi}_n \in F_x$ then $\xi_n \otimes \bar{\xi}_n \in F_x$ for all n since F_x is extremal. Also if $\xi_i \otimes \bar{\xi}_i \in F_x (i = 1, 2)$ then $(\xi_1 + \xi_2) \otimes (\xi_1 + \xi_2)^- \in F_x$ since

$$(\xi_1 + \xi_2) \otimes (\xi_1 + \xi_2)^- + (\xi_1 - \xi_2) \otimes (\xi_1 - \xi_2)^- = 2\xi_1 \otimes \bar{\xi}_1 + 2\xi_2 \otimes \bar{\xi}_2 \in F_x.$$

It follows that the set $L_x = \{\xi \in \mathcal{H} : \xi \otimes \bar{\xi} \in F_x\}$ is a closed subspace of $\mathcal{H}$ and $F_x = \{v \in K : P_x v = v P_x = v\}$.

If x, y are distinct points of X then $L_x \perp L_y$. For let η_x, η_y be two unit vectors, $\eta_x \in L_x$ and $\eta_y \in L_y$, and let

$$\xi, \eta \to \langle \mathscr{E}_x(\xi \otimes \bar{\eta}) \eta_x, \eta_y \rangle$$

be the bounded sesquilinear form on $\mathscr{H} \times \mathscr{H}$. This form determines a bounded self-adjoint operator S on $\mathscr{H}$; in fact S satisfies

$$0 \leqq S \leqq 1 .$$

Since

$$\mathscr{E}_x(\eta_x \otimes \bar{\eta}_x) = \eta_x \otimes \bar{\eta}_x$$

we have

$$(S\eta_x, \eta_x) = 1 .$$

On the other hand,

$$\langle \tau, \eta_y \otimes \bar{\eta}_y \rangle = \langle \tau, \mathscr{E}_y(\eta_y \otimes \bar{\eta}_y) \rangle \leqq \langle \tau, \mathscr{E}_x(\eta_y \otimes \bar{\eta}_y) \rangle$$

$$+ \langle \tau, \mathscr{E}_y(\eta_y \otimes \bar{\eta}_y) \rangle$$

$$\leqq \langle \tau, \mathscr{E}(X)(\eta_y \otimes \bar{\eta}_y) \rangle = \langle \tau, \eta_y \otimes \bar{\eta}_y \rangle$$

so that $\langle \tau, \mathscr{E}_x(\eta_y \otimes \bar{\eta}_y) \rangle = 0$ and $\mathscr{E}_x(\eta_y \otimes \bar{\eta}_y) = 0$ which entails $(S\eta_y, \eta_y) = 0$. By the spectral properties of S it follows that $(\eta_x, \eta_y) = 0$.

Let Q be the projection onto the orthogonal complement of $\bigoplus_{x \in X} L_x$; then if v is any state in K and $x \in X$ we have by (SR 1)

$$\mathscr{E}_x(\mathscr{E}_x v) = \mathscr{E}_x v$$

so $\mathscr{E}_x v \in F_x$ and $P_x(\mathscr{E}_x v) = (\mathscr{E}_x v) P_x = \mathscr{E}_x v$. It follows that $Q\mathscr{E}_x v = 0$ and so $\langle Q, \mathscr{E}_x v \rangle = \operatorname{trace}(Q\mathscr{E}_x v) = 0$ for all $x \in X, v \in K$ and from (SR 3) it follows that $Q = 0$. Hence

$$\mathscr{H} = \bigoplus_{x \in X} L_x .$$

It remains to determine explicitly the form of the map $v \to \mathscr{E}_x v$. The set $G_x \subseteq K$ given by $G_x = \{v \in K : \mathscr{E}_x v = 0\}$ is again a closed convex extremal set and so there exists a closed subspace K_x of $\mathscr{H}$ given by $K_x = \{\xi \in \mathscr{H} : \mathscr{E}_x(\xi \otimes \bar{\xi}) = 0\}$. Let Q_x be the projection onto K_x; then $G_x = \{v \in K : Q_x v = v Q_x = v\}$. We showed above that $K_x \supseteq L_y$ for all $y \neq x$, and as $K_x \cap L_x = 0$ so $K_x = L_x^\perp$. Now let ξ_1, ξ_2 be unit vectors, $\xi_1 \in L_x$

and $\xi_2 \in L_x^{\perp}$; for any α_1, α_2 in $\mathbb{C}$ we have

$$\mathscr{E}_x\{(\alpha_1 \xi_1 + \alpha_2 \xi_2) \otimes (\alpha_1 \xi + \alpha_2 \xi_2)^-\} + \mathscr{E}_x\{(\alpha_1 \xi_1 - \alpha_2 \xi_2) \otimes (\alpha_1 \xi_1 - \alpha_2 \xi_2)^-\}$$
$$= 2\alpha_1 \bar{\alpha}_1 \mathscr{E}_x(\xi_1 \otimes \bar{\xi}_1) + 2\alpha_2 \bar{\alpha}_2 \mathscr{E}_x(\xi_2 \otimes \bar{\xi}_2) = 2\alpha_1 \bar{\alpha}_1 \xi_1 \otimes \bar{\xi}_1$$

and so

$$0 \leq \mathscr{E}_x\{(\alpha_1 \xi_1 + \alpha_2 \xi_2) \otimes (\alpha_1 \xi_1 + \alpha_2 \xi_2)^-\}$$
$$\leq 2\alpha_1 \bar{\alpha}_1 \xi_1 \otimes \bar{\xi}_1 .$$

Hence $\mathscr{E}_x\{(\alpha_1 \xi_1 + \alpha_2 \xi_2) \otimes (\alpha_1 \xi_1 + \alpha_2 \xi_2)^-\} = \gamma(\xi_1 \otimes \bar{\xi}_1)$ for some $\gamma \in \mathbb{R}$. It follows by polarization that for each $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ there exists $\gamma \in \mathbb{C}$ such that

$$\mathscr{E}_x\{(\alpha_1 \xi_1 + \alpha_2 \xi_2) \otimes (\beta_1 \xi_1 + \beta_2 \xi_2)^-\} = \gamma(\xi_1 \otimes \bar{\xi}_1)$$

where $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \rightarrow \gamma$ is given by a self-adjoint operator T on $\mathbb{C}^2$. Since $0 \leq \mathrm{trace}(\mathscr{E}_x v) \leq \mathrm{trace}\, v$ for all $v \in K$, so $0 \leq T \leq 1$. Since $(T(1, 0), (1, 0)) = 1$ and $(T(0, 1), (0, 1)) = 0$, so $(1, 0)$ and $(0, 1)$ are eigenvectors of T and

$$\mathscr{E}_x\{(\alpha_1 \xi_1 + \alpha_2 \xi_2) \otimes (\beta_1 \xi_1 + \beta_2 \xi_2)^-\} = \alpha_1 \bar{\beta}_1 (\xi_1 \otimes \bar{\xi}_1) .$$

Thus for all $\xi \in \mathscr{H}$

$$\mathscr{E}_x(\xi \otimes \bar{\xi}) = (P_x \xi) \otimes (P_x \xi)^- = P_x(\xi \otimes \bar{\xi}) P_x$$

and hence for all states $v \in K$

$$\mathscr{E}_x v = P_x v P_x .$$

It follows that for all states v and sets $E \subseteq X$

$$\mathscr{E}_E v = \sum_{x \in E} \mathscr{E}_x v = \sum_{x \in E} P_x v P_x .$$

We now consider instruments whose value space is not discrete. Since it is known that if $\mathscr{B}$ is a maximal Abelian von Neuman algebra on a Hilbert space $\mathscr{H}$ with no discrete part, there does not exist any normal conditional expectation on $\mathscr{L}(\mathscr{H})$ whose range is the commutant of $\mathscr{B}$, we cannot expect to be able to construct repeatable observables with continuous value space in this model. As all discussions of this difficulty have implicitly assumed repeatability most of the results obtained have been essentially negative. We will show elsewhere [14] that by abandoning strict repeatability it is possible to construct explicitly interesting families of instruments on continuous value spaces which have strong properties close to the repeatability property.

§ 6. Material Implication

At various times von Neumann considered alternative verbal formulations of the relation of material implication in a physical setting. In this section we translate three such formulations into mathematical terms and examine some of their properties.

Birkhoff and von Neumann (p. 109 of [22]) pointed out that in a theory in which the mathematical representative of an experimental proposition is a closed subspace of a Hilbert space the following conditions on two experimental propositions are equivalent:

(i) For any statistical ensemble of systems, the probability of P is at most the probability of Q.

(ii) P implies Q – that is whenever one can predict P with certainty, one can predict Q with certainty.

In an unfinished manuscript written about 1937 and reviewed by Taub in [23] von Neumann described the relation of "implication" $(P \leqq Q)$ as follows:

(iii) $P \leqq Q$ means this: If a measurement of P on S (the physical system) has shown P to be true, then an immediately subsequent measurement of Q on S will certainly show Q to be true.

The three notions of material implication can be formulated in our notation as follows:

Let $\mathscr{E}, \mathscr{F}, \ldots$ be instruments on a two-point space $X = \{0,1\}$; then (i), the most commonly used of the three, becomes

$$\mathscr{E} \to_1 \mathscr{F} \quad \text{if and only if} \quad \langle \tau, \mathscr{E}(1) v \rangle \leqq \langle \tau, \mathscr{F}(1) v \rangle \quad \text{for all} \quad v \in K$$

and (ii) becomes

$$\mathscr{E} \to_2 \mathscr{F} \quad \text{if and only if} \quad \langle \tau, \mathscr{F}(1) v \rangle = \langle \tau, v \rangle \quad \text{for all } v$$

for which $\langle \tau, \mathscr{E}(1) v \rangle = \langle \tau, v \rangle$, while (iii) becomes

$$\mathscr{E} \to_3 \mathscr{F} \quad \text{if and only if} \quad \langle \tau, \mathscr{F}(1) \mathscr{E}(1) v \rangle = \langle \tau, \xi(1) v \rangle \quad \text{for all } v \in K.$$

In the beam model of probability which we used to motivate our definitions we saw that there was a correspondence between physical qualities and observables and between pieces of apparatus and instruments. Now material implication is usually conceived of as a relation between physical qualities rather than between pieces of apparatus, and so $\to_3$ is suspect on the grounds that it refers to the instrument ξ, whereas $\to_1$ and $\to_2$ refer only to the observables which the instruments determine. There are other objections to $\to_3$ in the general situation; there are two properties of material implication as it occurs in ordinary discourse which are

extremely important:

(MT 1) $\mathscr{F} \to \mathscr{J}$

(MI 2) $((\mathscr{E} \to \mathscr{F})$ and $(\mathscr{F} \to \mathscr{G}))$ entails $(\mathscr{E} \to \mathscr{G})$.

It is clear that $\to_1$ and $\to_2$ satisfy (MI 1) and (MI 2); in general there is no reason why $\to_3$ should satisfy either condition although we do have the following

Theorem 11. *Let $\mathscr{E}, \mathscr{F}, \mathscr{G}$ be instruments on a two-point value space. Then*

     (i) *if $\mathscr{F}$ satisfies (SR 1) then $\mathscr{F} \to_3 \mathscr{F}$;*

     (ii) *if $\mathscr{F}$ satisfies (SR 2) then $\mathscr{E} \to_3 \mathscr{G}$ whenever $\mathscr{E} \to_3 \mathscr{F}$ and $\mathscr{F} \to_3 \mathscr{G}$.*

Proof. (i) Suppose $\mathscr{F}$ satisfies (SR 1) then $\mathscr{F}(1) \mathscr{F}(1) v = \mathscr{F}(1) v$ for all v so that

$$\langle \tau, \mathscr{F}(1) \mathscr{F}(1) v \rangle = \langle \tau, \mathscr{F}(1) v \rangle .$$

(ii) Suppose $\mathscr{F}$ satisfies (SR 2) and that $\mathscr{E} \to_3 \mathscr{F}$ and $\mathscr{F} \to_3 \mathscr{G}$. Then $\langle \tau, \mathscr{F}(1) \mathscr{E}(1) v \rangle = \langle \tau, \mathscr{E}(1) v \rangle$ for all $v \in K$, and by (SR 2) $\mathscr{F}(1) \mathscr{E}(1) v = \mathscr{E}(1) v$ for all $v \in K$. But $\langle \tau, \mathscr{G}(1) \mathscr{F}(1) w \rangle = \langle \tau, \mathscr{F}(1) w \rangle$ for all $w \in K$. Put $w = \mathscr{E}(1) v$ so that $\langle \tau, \mathscr{G}(1) \mathscr{F}(1) \mathscr{E}(1) v \rangle = \langle \tau, \mathscr{F}(1) \mathscr{E}(1) v \rangle$ and hence $\langle \tau, \mathscr{G}(1) \mathscr{E}(1) v \rangle = \langle \tau, \mathscr{E}(1) v \rangle$ for all v.

However, for strongly repeatable instruments we see that $\to_2$ and $\to_3$ are equivalent:

Theorem 12. *Let $\mathscr{E}, \mathscr{F}$ be instruments on a two-point space.*

     (i) *If $\mathscr{E}$ satisfies (SR 1) then $(\mathscr{E} \to_2 \mathscr{F})$ entails $(\mathscr{E} \to_3 \mathscr{F})$.*

     (ii) *If $\mathscr{E}$ satisfies (SR 2) then $(\mathscr{E} \to_3 \mathscr{F})$ entails $(\mathscr{E} \to_2 \mathscr{F})$.*

Proof. (i) Suppose $\mathscr{E}$ satisfies (SR 1) and $\mathscr{E} \to_2 \mathscr{F}$. Then, by (SR 1), for all $v \in K$ $\langle \tau, \mathscr{E}(1) \mathscr{E}(1) \rangle = \langle \tau, \mathscr{E}(1) v \rangle$ and since $\mathscr{E} \to_2 \mathscr{F}$, $\langle \tau, \mathscr{F}(1) \mathscr{E}(1) v \rangle = \langle \tau, \mathscr{E}(1) v \rangle$ for all v.

(ii) Suppose $\mathscr{E}$ satisfies (SR 2) and $\mathscr{E} \to_3 \mathscr{F}$. Then if v is such that $\langle \tau, \mathscr{E}(1) v \rangle = \langle \tau, v \rangle$ we have $\mathscr{E}(1) v = v$ by (SR 2); but since $\mathscr{E} \to_3 \mathscr{F}$ we have

$$\langle \tau, \mathscr{F}(1) \mathscr{E}(1) v \rangle = \langle \tau, \mathscr{E}(1) v \rangle \quad \text{for all } v$$

and so $\langle \tau, \mathscr{F}(1) v \rangle = \langle \tau, v \rangle$ whenever $\langle \tau, \mathscr{E}(1) v \rangle = \langle \tau, v \rangle$.

In standard quantum theory where the observables are taken to be projection-valued measures the three forms of implication which we have considered are equivalent on the strongly repeatable instruments which are uniquely determined by the 'yes-no' observables. If we admit arbitrary positive-operator-valued measures we are left with the problem of choosing between three inequivalent candidates for material implication. The verbal description of $\to_1$ fails to do justice to the conceptual

content of the relation as it occirs in ordinary discourse. We have seen that $\to_3$ is unsatisfactory except when it is equivalent to $\to_2$. We propose therefore to adopt $\to_2$. This agrees with the choice of Pool [24].

Mackey [2] assumes that the set of 'yes-no' observables is an ortho-complemented lattice. The set of 'yes-no' observables in our sense is a lattice if and only if V^* can be identified with the self-adjoint part of a commutative C^*-algebra (this follows from the work of Kakutani [25] on M-spaces). However in quantum theory this never happens. It is also interesting to note that in our setting Mackey's relation of disjointness, $\circ$, does not have the properties one would desire and seems not to be a useful notion.

References

1. Kolmogorov, A. N.: Foundations of probability theory. Trans. by N. Morrison. New York: Chelsea Publ. Comp. 1950.
2. Mackey, G. W.: Mathematical foundations of quantum mechanics. New York: W. A. Benjamin Inc. 1963.
3. Urbanik, K.: Joint probability distribution of observables in quantum mechanics. Studia Math. 21, 117—133 (1961).
4. Varadarajan, V. S.: Probability in physics and a theorem on simultaneous observability. Commun. Pure Appl. Math. 15, 189—217 (1962).
5. Umegaki, H.: Conditional expectation in an operator algebra. Tohôku Math. J. 6, 177—181 (1954).
6. Nakamura, M., Turumaru, T.: Expectations in an operator algebra. Tohôku Math. J. 6, 182—188 (1954).
7. von Neumann, J.: Mathematical foundations of quantum mechanics. Trans. by R. T. Beyer. Princeton: Unive. Press 1955.
8. Schwinger, J.: The algebra of microscopic measurement. Proc. Nat. Acad. Sci. U.S. 45, 1542—1554 (1959) and 46, 257—265 (1960) and 46, 570—579 (1960).
9. Haag, R., Kastler, D.: An algebraic approach to quantum field theory. J. Math. Phys. 5, 848—861 (1964).
10. Luders, G.: Über die Zustandsänderung durch den Meßprozeß. Ann. Physik (6) 8, 322—328 (1951).
11. Goldberger, M. L., Watson, K. M.: Measurement of time correlations for quantum-mechanical systems. Phys. Rev. (2) 134, B919—928 (1964).
12. Furry, W. H.: Some aspects of the quantum theory of measurement pp. 1—64. Lectures in Theoretical Physics vol. VIIIa. Boulder: University of Colorado Press 1966.
13. Nakamura, M. Umegaki, H.: On von Neumann's theory of measurement in quantum statistics. Math. Japon 7, 151—157 (1962).
14. Davies, E. B.: On the repeated measurement of continuous observables in quantum mechanics. J. Funct. Anal. (to appear).
15. — Quantum stochastic processes. Commun. Math. Phys. 15, 277—304 (1969).
16. Schatten, R.: Norm ideals of completely continuous operators. Erg. der Math. Berlin-Göttingen-Heidelberg: Springer: 1960.
17. Ellis, A. J.: The duality of partially ordered normed linear spaces. J. London Math. Soc. 39, 730—744 (1964).

18. Namioka, I.: Partially ordered linear topological spaces. Mem. Am. Math. Soc. **24** (1957).
19. Mackey, G. W.: Borel structure in groups and their duals. Trans. Am. Math. Soc. **85**, 134—165 (1957).
20. Moy, S.-T. C.: Characterisations of conditional expectations as a transformation on function spaces. Pacific J. Math. **4**, 47—64 (1954).
21. Dixmier, J.: Les algèbres d'operateurs dans l'espace hilbertien (algebres de von Neumann). Cahiers scientifiques 25. 2nd ed. Paris: Gauthiers-Villars 1969.
22. Birkhoff, G., von Neumann, J.: The logic of quantum mechanics. Ann. Math. **37**, 823—843 (1936).
23. von Neumann, J.: Collected works, Vol. IV. No. 17. Edited by A. H. Taub. Oxford: Pergamon Press 1962.
24. Pool, J. C. T.: Baer*-semigroups and the logic of quantum mechanics. Commun. Math. Phys. **9**, 118—141 (1968).
25. Kakutani, S.: Concrete representation of abstract (*M*)-spaces. Ann. Math. **42**, 994—1024 (1941).

E. B. Davies   J. T. Lewis
Mass. Institute of Technology   Institute for Advanced Study
Department of Mathematics   Princeton, New Jersey, USA
Room 24-418
Cambridge, Mass. 02139, USA