Joint Distributions of Observables

S. GUDDER

Communicated by GIAN-CARLO ROTA

I. Introduction. If f and g are random variables on a probability space $(\Omega, \mathfrak{F}, p)$, the joint distribution of f and g is defined as the Borel measure $p_{f,g}$ on R^2 where $p_{f,g}(E) = p\{\omega: (f(\omega), g(\omega)) \in E\}$. Similar definitions are given for joint distributions of any finite number of random variables and finite dimensional joint distributions may be defined for an arbitrary collection of random variables. If E, F are one-dimensional Borel sets then $p_{f,g}(E \times F) = p(f^{-1}(E) \cap g^{-1}(F))$ may be thought of as the probability that f has a value in E and g has a value in F . Alternatively one may say that f and g have a joint distribution if there exists a measure μ on the Borel sets $B(R^2)$ such that $\mu(E \times F) = p(f^{-1}(E) \cap g^{-1}(F))$ for every $E, F \in B(R)$. Using this definition it is obvious that a joint distribution μ always exists, is unique, and $\mu \equiv p_{f,g}$. The situation is quite different when one considers observables in quantum mechanics. First one must decide how to define a joint distribution of observables. One soon finds that with any reasonable definition joint distributions do not always exist for observables. Thus a given collection C of observables can be classified into one of three different categories. If C always has finite joint distributions then we call C probabilistically definite, if C has no finite joint distributions we call C probabilistically indefinite, and otherwise we call C probabilistically mixed. In this paper two different types of joint distributions are considered and some of their properties are derived.

2. Joint observables. In this paper we shall assume that all observables are defined on a sum logic L (see [4] for definitions and notation.) If $(\Omega, \mathfrak{F})$ is a measurable space, a σ -homomorphism h from Ω into L is a map which satisfies

- (i) $h(\Omega) = 1$; (ii) $h(E) \perp h(F)$ if $E \cap F = \emptyset, E, F \in \mathfrak{F}$;
- (iii) $h(\cup E_i) = \sum h(E_i), E_i \cap E_j = \emptyset, i \neq j, E_i \in \mathfrak{F}$.

Let x and y be observables on L . If there is a σ -homomorphism z from $B(R^2)$ into L which satisfies $z(E \times R) = x(E), z(R \times F) = y(F)$ for all $E, F \in B(R)$, then z is called a *joint observable* for x and y . We define joint observables for an arbitrary family of observables in an analogous fashion. To simplify matters in the

sequel all our results will be given for pairs of observables. However, all the definitions and theorems may be easily generalized to arbitrary families.

Lemma 2.1. (i) A σ -homomorphism $z: B(R^2) \rightarrow L$ is a joint observable for x and y if and only if $z(E \times F) = x(E) \wedge y(F)$ for all $E, F \in B(R)$. (ii) If z is a joint observable for x and y , then z is unique.

Proof. (i) Sufficiency is obvious. To prove necessity note that $x(E) = z(E \times R) = z((E \times F) \cup (E \times F')) = z(E \times F) + z(E \times F')$ and similarly $y(F) = z(E \times F) + z(E' \times F)$. Applying the distributive law it follows that $x(E) \wedge y(F) = z(E \times F)$. (ii) Since z is determined on measurable rectangles it follows from a standard argument that z is unique.

It is obvious that if x and y have a joint observable, then $x \leftrightarrow y$. Conversely, suppose $x \leftrightarrow y$. Let A be the algebra of subsets of $B(R^2)$ consisting of finite unions of disjoint measurable rectangles and define $z^*: A \rightarrow L$ by

$$z^*\left(\bigcup_{i=1}^n E_i \times F_i\right) = \sum_{i=1}^n x(E_i) \wedge y(F_i)$$

where $(E_i \times F_i) \cap (E_j \times F_j) = \emptyset$, $i \neq j$. Then z^* is countably additive on A since if $E \times F = \bigcup (E_i \times F_i)$, $(E_i \times F_i) \cap (E_j \times F_j) = \emptyset$, $i \neq j$, we have by the distributive law

$$\begin{aligned} \sum z^*(E_i \times F_i) &= \sum (x(E_i) \wedge y(F_i)) = [\sum x(E_i)] \wedge [\sum y(F_i)] \\ &= x(\bigcup E_i) \wedge y(\bigcup F_i) = x(E) \wedge y(F) = z^*(E \times F). \end{aligned}$$

Using a procedure analogous to the standard method of extending measures from an algebra to the generated σ -algebra one can show that z^* has a unique extension to a σ -homomorphism z on $B(R^2)$. Thus x and y have a joint observable. (This result was proved by Varadarajan [9] using different methods.)

Following [9] we can use joint observables to define functions of several simultaneous observables. Let $x_1, \dots, x_n$ be simultaneous observables, let z be their joint observable and let ψ be a Borel function on R^n . We define $\psi(x_1, \dots, x_n)$ as the observable $E \rightarrow z(\psi^{-1}(E))$. We now show that this definition of a function of several simultaneous observables has the properties one would expect of a functional calculus.

Theorem 2.2. If $x_i = u_i(x)$, $i = 1, 2, \dots, n$ and ψ is a Borel function on R^n , then $\psi(x_1, \dots, x_n) = \psi(u_1, \dots, u_n)(x)$.

Proof. We shall prove the theorem for the case $n = 2$. Applying Theorem 3.4 [9] there is a measurable space $(\Omega, \mathfrak{F})$, a σ -homomorphism $h: \mathfrak{F} \rightarrow L$ and an $\mathfrak{F}$ -measurable function f such that $x(E) = h(f^{-1}(E))$. Defining $f_i = u_i \circ f$ we have

$$x_i(E) = x(u_i^{-1}(E)) = h(f^{-1}u_i^{-1}(E)) = h(f_i^{-1}(E)) \quad i = 1, 2.$$

Defining $\phi: \Omega \rightarrow R^2$ by $\phi(\omega) = (f_1(\omega), f_2(\omega))$ and $z: B(R^2) \rightarrow L$ by $z(E) = h(\phi^{-1}(E))$

it is easy to check that z is the joint observable for x_1, x_2 . Hence

$$\begin{aligned}\psi(x_1, x_2)(E) &= z(\psi^{-1}(E)) = h[\phi^{-1}(\psi^{-1}(E))] \\ &= h[(u_1 \circ f, u_2 \circ f)^{-1}(\psi^{-1}(E))] = h[f^{-1}((u_1, u_2)^{-1}(\psi^{-1}(E)))] \\ &= x[(u_1, u_2)^{-1}(\psi^{-1}(E))] = x[\psi(u_1, u_2)^{-1}(E)] = \psi(u_1, u_2)(x)(E).\end{aligned}$$

We thus have, for example, that $u_1(x) + u_2(x) = (u_1 + u_2)(x)$.

3. Type 1 joint distributions. If x and y have a joint observable and m is a state, we define the *joint distribution* $m_{x,y}$ of x and y in the state m to be the two-dimensional Borel measure $m_{x,y}(E) = m(z(E)), E \in B(R^2)$. Notice $m_{x,y}(E \times F) = m(x(E) \wedge y(F)), E, F \in B(R)$. Now it is possible that x and y do not have a joint observable (equivalently $x \nleftrightarrow y$) and yet we can still define joint distributions in certain states. That is, we say that x and y have a *type 1 joint distribution in the state m* if there is a two-dimensional Borel measure $m_{x,y}$ such that $m_{x,y}(E \times F) = m(x(E) \wedge y(F))$ for all $E, F \in B(R)$ and we call $m_{x,y}$ the *type 1 joint distribution of x and y in the state m* . We say that x and y are *type 1 probabilistically definite, indefinite or mixed* depending upon whether x and y have a type 1 joint distribution in every state, no states, or some states and not others, respectively.

Theorem 3.1. *The following statements are equivalent.*

- (i) $x \leftrightarrow y$.
- (ii) x and y have a joint observable.
- (iii) x and y are type 1 probabilistically definite.

Proof. The equivalence of (i) and (ii) have already been shown and it is obvious that (ii) implies (iii). For the rest of the proof we need the following lemma.

Lemma 3.2. *If a and b are events, then $a \leftrightarrow b$ if and only if $m(a) = 1$ and $m(a \wedge b) = 0$ implies $m(b) = 0$ for any state m .*

Proof. For necessity we have $a = (a - a \wedge b) + (a \wedge b)$, $b = (b - a \wedge b) + (a \wedge b)$. If $m(a) = 1$ and $m(a \wedge b) = 0$, then $m(b) = m(b - a \wedge b)$. Now $b - a \wedge b \leq (a - a \wedge b)'$ and hence $m(b - a \wedge b) \leq 1 - m(a) = 0$. Thus $m(b) = 0$. For sufficiency, suppose $m(a - a \wedge b) = 1$. Then $m(a) = 1$ and $m(a \wedge b) = 0$. Hence $m(b) = 0$ and $m(b - a \wedge b) = 0$. Thus $m((b - a \wedge b)') = 1$ and $a - a \wedge b \leq (b - a \wedge b)'$ and $(a - a \wedge b) \perp (b - a \wedge b)$. Thus $a \leftrightarrow b$.

We now continue the proof of Theorem 3.1. Suppose (iii) holds and $E, F \in B(R)$. Let m be a state which satisfies $m[x(E)] = 1$ and $m[x(E) \wedge y(F)] = 0$. Then $m_{x,y}(E \times R) = m[x(E)] = 1$ and $m_{x,y}(E \times F) = m[x(E) \wedge y(F)] = 0$. Hence $0 = m_{x,y}(R \times F) = m[y(F)]$. Applying Lemma 3.2, $x(E) \leftrightarrow y(F)$ and hence (i) holds.

We now obtain some conditions under which two observables have a type 1 joint distribution in a state m .

Theorem 3.3. *The following statements are equivalent. (i) x and y have a type 1 joint distribution in the state m . (ii) If $E \times F = \bigcup_{i=1}^{\infty} E_i \times F_i$ where $E, F \in B(R)$ and $E_i \times F_i$ are disjoint measurable rectangles, $i = 1, 2, \dots$, then*

$$m[\vee (x(E_i) \wedge y(F_i))] = m[(\vee x(E_i)) \wedge (\vee y(F_i))].$$

(iii) *There is a measure μ on $B(R^2)$ such that*

$$\mu\{(\omega_1, \omega_2): \chi_1(\omega_1) + \chi_2(\omega_2) = 2\} = m[(\chi_1(x) + \chi_2(y))\{2\}]$$

for any measurable characteristic functions χ_1 and χ_2 .

Proof. Suppose (i) holds and $E \times F$ satisfies the conditions of (ii). Then since $E \times F = (\cup E_i) \times (\cup F_i)$ we have

$$\begin{aligned} m[\vee (x(E_i) \wedge y(F_i))] &= \sum m(x(E_i) \wedge y(F_i)) \\ &= \sum m_{x,y}(E_i \times F_i) = m_{x,y}(\cup E_i \times F_i) = m_{x,y}(E \times F) \\ &= m_{x,y}((\cup E_i) \times (\cup F_i)) = m(x(\cup E_i) \wedge y(\cup F_i)) \\ &= m[(\vee x(E_i)) \wedge (\vee y(F_i))]. \end{aligned}$$

Now suppose (ii) holds and A is the algebra consisting of finite unions of disjoint measurable rectangles. Define μ_m on A by $\mu_m(\cup E_i \times F_i) = \sum m(x(E_i) \wedge y(F_i))$. Letting $E \times F$ be as in condition (ii) we have

$$\begin{aligned} \sum \mu_m(E_i \times F_i) &= \sum m(x(E_i) \wedge y(F_i)) = m[\vee (x(E_i) \wedge y(F_i))] \\ &= m[(\vee x(E_i)) \wedge (\vee y(F_i))] = m[x(\cup E_i) \wedge y(\cup F_i)] \\ &= m[x(E) \wedge y(F)] = \mu_m(E \times F). \end{aligned}$$

By the Hahn extension theorem μ_m has a unique extension to a measure $m_{x,y}$ on $B(R^2)$. Now suppose (iii) holds. Then

$$\begin{aligned} \mu(E \times F) &= \mu\{(\omega_1, \omega_2) : \chi_E(\omega_1) + \chi_F(\omega_2) = 2\} \\ &= m[(\chi_E(x) + \chi_F(y))\{2\}] = m(x(E) \wedge y(F)) \end{aligned}$$

where the third equality follows from Lemma 6.2 [3]. Thus μ is the type 1 joint distribution for x and y . Conversely if a type 1 joint distribution $m_{x,y}$ exists we have

$$\begin{aligned} m_{x,y}\{(\omega_1, \omega_2) : \chi_E(\omega_1) + \chi_F(\omega_2) = 2\} &= m_{x,y}(E \times F) \\ &= m(x(E) \wedge y(F)) = m[(\chi_E(x) + \chi_F(y))\{2\}]. \end{aligned}$$

This completes the proof.

Given two observables x and y one can now ask in what states do x and y have a type 1 joint distribution? Before we discuss this question we need some definitions. A state m is an *eigenstate* for an observable x if there is a real number λ such that $m(x\{\lambda\}) = 1$. In this case λ is called the *eigenvalue* corresponding to the eigenstate m . Although an eigenvalue may correspond to many eigen-

states, it is clear that an eigenstate can have only one eigenvalue. If there are real numbers λ, μ such that $m(x\{\lambda\}) = m(y\{\mu\}) = 1$, then m is a *common eigenstate* and (λ, μ) is a *common eigenpair* for x and y . If a state m has the form $m = \sum r_i m_i$, $r_i \geq 0$, $\sum r_i = 1$, then m is called a *mixture* of the states m_i , $i = 1, 2, \dots$. It is clear that a mixture of eigenstates with the same eigenvalue is again an eigenstate with that eigenvalue. An observable x has pure point spectrum if there is a sequence of real numbers λ_i such that $\sum x\{\lambda_i\} = 1$.

Theorem 3.4. (i) If m is a mixture of common eigenstates for x and y , then x and y have a type 1 joint distribution in the state m . (ii) Let x and y have pure point spectrum and let (λ_i, μ_i) , $i = 1, 2, \dots$, be the common eigenpairs. If x and y have a type 1 joint distribution $m_{x,y}$ in the state m , then $\sum m(x\{\lambda_i\} \wedge y\{\mu_i\}) = 1$. In particular if x and y have no common eigenpair then x and y are type 1 probabilistically indefinite; if x and y have one common eigenpair, then m is a corresponding common eigenstate.

Proof. (i) Suppose $m = \sum r_i m_i$, $r_i \geq 0$, $\sum r_i = 1$ and $m_i(x\{\lambda_i\}) = m_i(y\{\mu_i\}) = 1$, $i = 1, 2, \dots$. Define the set function μ_m by $\mu_m(G) = \sum \{r_i : (\lambda_i, \mu_i) \in G\}$, $G \in B(R^2)$. Then μ_m is a probability measure on $B(R^2)$ and

$$\begin{aligned} m(x(E) \wedge y(F)) &= m[(\sum_i x\{\lambda_i : \lambda_i \in E\}) \wedge (\sum_j y\{\mu_j : \mu_j \in F\})] \\ &= \sum_k r_k m_k[(\sum_i x\{\lambda_i : \lambda_i \in E\}) \wedge (\sum_j y\{\mu_j : \mu_j \in F\})]. \end{aligned}$$

Now $m_k[(\sum_i x\{\lambda_i : \lambda_i \in E\}) \wedge (\sum_j y\{\mu_j : \mu_j \in F\})] = 0$ unless $(\lambda_k, \mu_k) \in E \times F$ and it follows from Corollary 6.4 [3] that this expression equals 1 if $(\lambda_k, \mu_k) \in E \times F$. Thus $m(x(E) \wedge y(F)) = \sum_k \{r_k : (\lambda_k, \mu_k) \in E \times F\} = \mu_m(E \times F)$ and hence x and y have a type 1 joint distribution in the state m . (ii) Since $m_{x,y}$ is concentrated on $\cup \{(\lambda_i, \mu_i)\}$ we have

$$1 = m_{x,y}[\cup \{(\lambda_i, \mu_i)\}] = \sum_{i,j} m(x\{\lambda_i\} \wedge y\{\mu_j\}).$$

Now if $i \neq j$, $\lambda_i \neq \lambda_j$, $\mu_i \neq \mu_j$, then $x\{\lambda_i\} \wedge y\{\mu_j\} = 0$ since otherwise there is a state m_0 such that $m_0(x\{\lambda_i\} \wedge y\{\mu_j\}) = 1$. Then $1 = m_0(x\{\lambda_i\}) = m_0(y\{\mu_j\})$ and (λ_i, μ_j) is a common eigenpair which is a contradiction. Then (ii) follows.

Notice that it follows from (i) that two non-simultaneous observables with a common eigenstate are type 1 probabilistically mixed.

We now consider a special case in which we can characterize the states in which two observables have a type 1 joint distribution. This is the case which is most important in quantum mechanics. Let H be a Hilbert space and let L_H be the logic consisting of all closed subspaces of H . In this case observables may be identified with self-adjoint operators (not necessarily bounded) which we also denote by x, y , etc. We denote the orthogonal projections in the resolution of identity for x by $E \rightarrow x(E)$ and we identify closed subspaces with their unique orthogonal projections. It is proved in [2] that every state on L_H has the form $m(a) = \sum r_i \langle \phi_i, a \phi_i \rangle$ where $r_i \geq 0$, $\sum r_i = 1$, and $\{\phi_i\}$ is an ortho-

normal set of vectors, $a \in L_H$. Thus every state is a mixture of *pure states* $a \rightarrow \langle \phi_i, a \phi_i \rangle$. We begin with a pair of lemmas the second of which has an interest of its own.

Lemma 3.5. *Let $\{\phi_i\}$ be an orthonormal set. If a vector ϕ satisfies $\|\phi\|^2 = \sum |\langle \phi, \phi_i \rangle|^2$, then $\phi = \sum \langle \phi, \phi_i \rangle \phi_i$.*

Proof. $\|\phi - \sum \langle \phi, \phi_i \rangle \phi_i\|^2 = \|\phi\|^2 - \sum |\langle \phi, \phi_i \rangle|^2 = 0$.

Lemma 3.6. *Let x and y be self-adjoint operators with pure point spectrum. Then $x(E)y(F)\phi = y(F)x(E)\phi$ for all $E, F \in B(R)$ if and only if ϕ is a linear combination of common eigenvectors of x and y .*

Proof. Sufficiency is straightforward and is left to the reader. To prove necessity let x and y have eigenvalues $\{\lambda_i\}$, $\{\mu_i\}$ respectively and let

$$P_i = x\{\lambda_i\} \quad Q_i = y\{\mu_i\}.$$

Then

$$Q_i = \sum_j P_j \wedge Q_i + (Q_i - \sum_j P_j \wedge Q_i)$$

and

$$P_i = \sum_j P_i \wedge Q_j + (P_i - \sum_j P_i \wedge Q_j).$$

Since $H = \sum Q_i = \sum P_i$ we can find a complete orthonormal set of eigenvectors of y , ϕ_{ij}^k, ψ_j^k where the ϕ_{ij}^k are common eigenvectors in $P_i \wedge Q_i$ and $\psi_j^k \in (Q_i - \sum_j P_i \wedge Q_j)$, $k = 1, 2, \dots$; and also a complete orthonormal set of eigenvectors of x , ϕ_{ij}^k, χ_i^k where $\phi_{ij}^k \in P_i \wedge Q_i$ and $\chi_i^k \in (P_i - \sum_j P_i \wedge Q_j)$. Now since $P_i Q_i \phi = Q_i P_i \phi$ we have

$$P_i \sum_{i,k} \langle \phi, \phi_{ij}^k \rangle \phi_{ij}^k + P_i \sum_k \langle \phi, \psi_j^k \rangle \psi_j^k = Q_i \sum_{i,k} \langle \phi, \phi_{ij}^k \rangle \phi_{ij}^k + Q_i \sum_k \langle \phi, \chi_i^k \rangle \chi_i^k.$$

We then obtain

$$\psi = \sum_{k,s} \langle \phi, \psi_j^k \rangle \langle \psi_j^k, \chi_i^s \rangle \chi_i^s = \sum_{k,s} \langle \phi, \chi_i^k \rangle \langle \chi_i^k, \psi_j^s \rangle \psi_j^s.$$

If $\psi \neq 0$ then $\psi \in P_i - \sum_j P_i \wedge Q_j$ and from the right equality in the above expression $\psi \in Q_i$ which is a contradiction. Thus $\psi = 0$ and since the χ_i^s are linearly independent, $\sum_k \langle \phi, \psi_j^k \rangle \langle \psi_j^k, \chi_i^s \rangle = 0$. Hence $\langle \phi, \chi_i^s \rangle = \sum_{j,k} \langle \phi, \psi_j^k \rangle \langle \psi_j^k, \chi_i^s \rangle = 0$ and since this holds for all i and s , $\phi = \sum_{i,j,k} \langle \phi, \phi_{ij}^k \rangle \phi_{ij}^k$.

As a corollary to this lemma we have the well known result that two commuting self-adjoint operators with pure point spectrum have a complete orthonormal set of common eigenvectors.

Theorem 3.7. *Let x and y be observables in L_H . Then x and y have a type 1 joint distribution in the state $m = \sum r_i \langle \phi_i, \cdot \phi_i \rangle$ if and only if $x(E)y(F)\phi_i = y(F)x(E)\phi_i$ for every $E, F \in B(R)$, $i = 1, 2, \dots$.*

Proof. We first prove sufficiency. It follows from Theorem 13.7 [7] that

$x(E) \wedge y(F)\phi_i = x(E)y(F)\phi_i = x(E)y(F)\phi_i$, $i = 1, 2, \dots$, for all $E, F \in B(R)$. Letting $E \times F = \bigcup E_i \times F_i$ satisfy condition (ii) of Theorem 3.3 we have

$$\begin{aligned} m[\vee(x(E_i) \wedge y(F_i))] &= \sum_i r_i \langle \phi_i, \vee(x(E_i) \wedge y(F_i))\phi_i \rangle \\ &= \sum_i r_i \langle \phi_i, \sum_j x(E_j)y(F_j)\phi_i \rangle = \sum_i r_i \langle \phi_i, x(\bigcup E_j)y(\bigcup F_j)\phi_i \rangle \\ &= \sum_i r_i \langle \phi_i, (\vee x(E_j)) \wedge (\vee y(F_j))\phi_i \rangle = m[(\vee x(E_j)) \wedge (\vee y(F_j))]. \end{aligned}$$

Applying Theorem 3.3 sufficiency is proved. Conversely, if μ is the type 1 joint distribution for x and y we have

$$\mu(E \times F) = \sum r_i \langle \phi_i, x(E) \wedge y(F)\phi_i \rangle, \quad E, F \in B(R).$$

Since the sets $E \times F$, $E' \times F'$, $E' \times F$, and $E \times F'$ are disjoint and have union equal to R^2 we obtain

$$\begin{aligned} \sum r_i [\langle \phi_i, x(E) \wedge y(F)\phi_i \rangle + \langle \phi_i, x(E') \wedge y(F')\phi_i \rangle + \langle \phi_i, x(E') \wedge y(F)\phi_i \rangle \\ + \langle \phi_i, x(E) \wedge y(F')\phi_i \rangle] = 1. \end{aligned}$$

Thus

$$\begin{aligned} \|\phi_i\|^2 = 1 &= \langle \phi_i, x(E) \wedge y(F)\phi_i \rangle + \langle \phi_i, x(E') \wedge y(F')\phi_i \rangle \\ &+ \langle \phi_i, x(E') \wedge y(F)\phi_i \rangle + \langle \phi_i, x(E) \wedge y(F')\phi_i \rangle \quad i = 1, 2, \dots \end{aligned}$$

It follows from Lemma 3.5 that ϕ_i is a linear combination of common eigenvectors of $x(E)$ and $y(F)$ and applying Lemma 3.6 the proof is complete.

Corollary 3.8. *If x and y are observables in L_H with pure point spectrum, then x and y have a type 1 joint distribution in the state $m = \sum r_i \langle \phi_i, \cdot \phi_i \rangle$ if and only if ϕ_i is a linear combination of common eigenvectors for x and y , $i = 1, 2, \dots$.*

4. Type 2 joint distributions. In this section we consider a second type of joint distribution which was originally defined by Urbanik [8] in the L_H case. We need a few preliminaries before we can define this type of joint distribution. A collection of states M on L is *full* if $a \leq b$ whenever $m(a) = 1$ implies $m(b) = 1$ for $m \in M$. We say that the *sum of two unbounded observables x, y exists* if there is a full set of states M and an observable z such that $m(x)$, $m(y)$ exists and are finite and $m(z) = m(x) + m(y)$ for all $m \in M$. If μ is a Borel measure on $B(R^2)$ and $(r, s) \in R^2$ we define $\mu^{r,s}$ on $B(R)$ as

$$\mu^{r,s}(E) = \mu\{(\omega_1, \omega_2) : r\omega_1 + s\omega_2 \in E\}, \quad E \in B(R).$$

Finally we say that x and y have a *type 2 joint distribution in the state m* if $rx + sy$ exists for all $(r, s) \in R^2$ and if there is a Borel measure μ_m on $B(R^2)$ such that $\mu_m^{r,s}(E) = m((rx + sy)(E))$ for all $(r, s) \in R^2$, and $E \in B(R)$. That is,

$$\mu_m\{(\omega_1, \omega_2) : r\omega_1 + s\omega_2 \in E\} = m((rx + sy)(E)).$$

It is shown in [1] that if a type 2 joint distribution exists it is unique.

The author feels that the type 1 is a more natural concept than the type 2 joint distribution. However, we shall briefly discuss this type since it has already been introduced in the literature. Note that the concept of type 1 joint distribution can be defined for any pair of observables x and y (although, of course, it need not exist) while the type 2 joint distribution can only be considered if $rx + sy$ exists for all $(r, s) \in R^2$. Notice further that x and y may have a type 1 joint distribution (e.g. x and y might have a common eigenstate.) without having a type 2 joint distribution. (e.g. in L_H x and y might not have a common dense domain.) We define type 2 probabilistically definite, indefinite, and mixed systems in an analogous fashion to the type 1 definitions. Urbanik [8] has given examples of all three of these type 2 systems and has derived some interesting necessary and sufficient conditions for observables in L_H to have type 2 joint distributions.

Theorem 4.1. *If $x \leftrightarrow y$ then x and y are type 2 probabilistically definite and their type 2 joint distribution is equal to their type 1 joint distribution.*

Proof. Let z be the joint observable for x and y , let m be a state and let μ_m be the type 1 joint distribution for x and y in the state m . If $f: R^2 \rightarrow R$ is a Borel function we define the observable $f(z)$ as $f(z)(E) = z(f^{-1}(E))$ for $E \in B(R)$. As in Theorem 2.2 we can prove that $f_1(z) + f_2(z) = (f_1 + f_2)(z)$ for any two-dimensional Borel functions f_1, f_2 . Letting $\pi_i: R^2 \rightarrow R$ be the Borel functions defined by $\pi_i(\omega_1, \omega_2) = \omega_i$, $i = 1, 2$, we see that $x = \pi_1(z)$, $y = \pi_2(z)$. If $(r, s) \in R^2$ and $E \in B(R)$ we have

$$\begin{aligned} m((rx + sy)(E)) &= m((r\pi_1(z) + s\pi_2(z))(E)) \\ &= m[(r\pi_1 + s\pi_2)(z)(E)] = m[z(r\pi_1 + s\pi_2)^{-1}(E)] \\ &= \mu_m((r\pi_1 + s\pi_2)^{-1}(E)) = \mu_m^{r,s}(E). \end{aligned}$$

Theorem 4.2. *Suppose x and y have pure point spectrum. If x and y are type 2 probabilistically definite, then $x \leftrightarrow y$.*

Proof. Let the eigenvalues of x and y be $\{\lambda_i\}$, $\{\mu_i\}$ respectively. Since $(\mu_j - \mu_r)/(\lambda_i - \lambda_s)$ is a countable set there is a number λ such that $\lambda \neq (\mu_j - \mu_r)/(\lambda_i - \lambda_s)$ for any $j, r, i \neq s$. Let m be a state and μ_m the corresponding type 2 joint distribution. Then by definition $m((\lambda x + y)(E)) = \mu_m\{(\omega_1, \omega_2): \lambda\omega_1 + \omega_2 \in E\}$. Since μ_m is concentrated on the points (λ_i, μ_i) , $i, j = 1, 2, \dots$, we have $m((\lambda x + y)(\lambda\lambda_i + \mu_i)) = \mu_m\{\lambda_i, \mu_i\}$. Let $c = (\lambda x + y)(\lambda\lambda_i + \mu_i)$. Since

$$m(\{\lambda_i\}) = \mu_m(\{\lambda_i\} \times R) \geq \mu_m\{\lambda_i, \mu_i\} = m(c)$$

we have $c \leq x\{\lambda_i\}$ and similarly $c \leq y\{\mu_i\}$. Now suppose $m(x\{\lambda_i\}) = 1$ and $m(c) = 0$. It then follows that $\mu_m(\{\lambda_i\} \times R) = 1$ and $\mu_m\{\lambda_i\} = 0$ and hence $0 = \mu_m(R \times \{\mu_i\}) = m(y\{\mu_i\})$. It then follows that $(x\{\lambda_i\} - c) \perp (y\{\mu_i\} - c)$. Since $x\{\lambda_i\} = (x\{\lambda_i\} - c) + c$ and $y\{\mu_i\} = (y\{\mu_i\} - c) + c$ we have $x\{\lambda_i\} \leftrightarrow y\{\mu_i\}$. It now follows that $x \leftrightarrow y$.

This last theorem is a generalization of a theorem due to Urbanik [8] who obtained this result in the Hilbert space case. Varadarajan [9] proved this result for arbitrary self-adjoint operators in the Hilbert space case. The author does not know if this theorem holds for arbitrary observables.

Theorem 4.3. *Suppose x and y are self-adjoint operators with pure point spectrum. (i) If x and y have a type 2 joint distribution in the state $m = \sum r_i \langle \phi_i, \cdot \phi_i \rangle$ then ϕ_i is a linear combination of eigenvectors of $\lambda x + y$ (λ as in the proof of Theorem 4.2) corresponding to eigenvalues $\lambda \lambda_i + \mu_i$. (ii) If ϕ_i is a linear combination of common eigenvectors of x and y and if $rx + sy$ exists for all $(r, s) \in R$, then x and y have a type 2 joint distribution in the state $m = \sum \phi_i \langle \phi_i, \cdot \phi_i \rangle$ which equals the type 1 joint distribution in the state m .*

Proof. (i) Let the eigenvalues of x and y be $\{\lambda_i\}$, $\{\mu_i\}$ respectively and let λ be as in the proof of Theorem 4.2. Since the type 2 joint distribution μ is concentrated at $\{(\lambda_i, \mu_i) : i, j = 1, 2, \dots\}$ and $\mu\{(\lambda_i, \mu_i)\} = m((\lambda x + y)(\lambda \lambda_i + \mu_i))$ we have

$$1 = \sum_{i,j} \mu\{(\lambda_i, \mu_i)\} = \sum_{i,j,k} r_k \sum_s |\langle \phi_k, \phi_{ij}^* \rangle|^2 = \sum_k r_k \sum_{i,j,s} |\langle \phi_k, \phi_{ij}^* \rangle|^2$$

where $(\lambda x + y)\phi_{ij}^k = (\lambda \lambda_i + \mu_j)\phi_{ij}^k$, $k = 1, 2, \dots$. We thus have $\sum_{i,j,s} |\langle \phi_k, \phi_{ij}^* \rangle|^2 = 1 = \|\phi_k\|^2$, $k = 1, 2, \dots$. Applying Lemma 3.5. $\phi_k = \sum_{i,j,s} \langle \phi_k, \phi_{ij}^* \rangle \phi_{ij}^*$. (ii) is left to the reader.

We now show that not all states of the form given in Theorem 4.3 (i) admit type 2 joint distributions.

Theorem 4.4. *Let x and y be as in Theorem 4.3 and suppose ϕ is a unit eigenvector of $\lambda x + y$ with eigenvector $\lambda \lambda_i + \mu_j$ which is not a common eigenvector of x and y corresponding to eigenvalues λ_i, μ_j respectively. Then x and y do not have a type 2 joint distribution in the state $m = \langle \phi, \cdot \phi \rangle$.*

Proof. Suppose x and y have a type 2 joint distribution μ in the state m . Let $\omega \neq \lambda$ satisfy $\omega \neq (\mu_j - \mu_r)/(\lambda_i - \lambda_s)$, $j, r, i \neq s$. Since ϕ is not a common eigenvector of x and y corresponding to λ_i and μ_j respectively it follows $(\omega x + y)\phi \neq (\omega \lambda_i + \mu_j)\phi$. Indeed, if equality did occur, then subtracting it from $(\lambda x + y)\phi = (\lambda \lambda_i + \mu_j)\phi$ we obtain $x\phi = \lambda_i \phi$ and hence $y\phi = \mu_j \phi$, a contradiction. We then have

$$\begin{aligned} 1 &= m((\lambda x + y)(\lambda \lambda_i + \mu_j)) = \mu\{(\lambda_i, \mu_j)\} \\ &= m((\omega x + y)(\omega \lambda_i + \mu_j)) \neq 1, \end{aligned}$$

a contradiction.

5. Remarks. We say that two observables x and y are type 1 independent in the state m if their type 1 joint distribution $m_{x,y}$ exists in the state m and is the product of their marginal distributions; i.e., $m_{x,y}(E \times F) = m(x(E))m(y(F))$, for all $E, F \in B(R)$. We now show that this independence reduces to the independence considered in [4].

Lemma 5.1. *Two observables x, y are type 1 independent in the state m if and only if $m(x(E) \wedge y(F)) = m(x(E))m(y(F))$ for all $E, F \in B(R)$.*

Proof. For sufficiency define $m_{x,y}$ at the product of the marginal distributions. Then $m_{x,y}$ is a probability measure on $B(R^2)$ satisfying

$$m_{x,y}(E \times F) = m_{x,y}(x(E))m_{x,y}(y(F)) = m(x(E) \wedge y(F))$$

and is thus the type 1 joint distribution. Conversely, letting $m_{x,y}$ be the type 1 joint distribution we have

$$m(x(E) \wedge y(F)) = m_{x,y}(E \times F) = m(x(E))m(y(F)).$$

In an analogous way we can define type 2 independence. This has been done in [8] for the Hilbert space case and also some interesting and important examples are considered.

Other types of joint distributions have been considered in the literature. For example, in [6] the joint distribution μ of self-adjoint operators x and y is defined in a state $m = \langle \phi, \cdot \phi \rangle$ by $\mu(E \times F) = \langle \phi, x(E)y(F)\phi \rangle$. Unfortunately this may not be a real number although it is real if and only if

$$\langle \phi, x(E)y(F)\phi \rangle = \langle \phi, y(F)x(E)\phi \rangle \quad \text{for all } E, F \in B(R).$$

Thus a sufficient condition for this quantity to be real is that x and y have a type 1 joint distribution. In [5] the joint distribution μ of observables x and y in state m is defined by $\mu(E \times F) = m(x(E))m(y(F))$. This seems unsatisfactory since it implicitly assumes that x and y are independent.

Let x and y be observables on the Hilbert space logic L_H . We now briefly consider what one might define as an "order" probability for x and y . In conventional probability theory it does not matter in what order measurements are made; however, this is not the case in quantum mechanics. Suppose we are in a pure state defined by the unit vector ϕ and suppose we measure the observable x and find it has a value in the Borel set E . According to quantum mechanical theory the state given by ϕ then abruptly changes to the state given by the unit vector $x(E)\phi/||x(E)\phi||$ assuming $x(E)\phi \neq 0$. We thus define the *first order probability of y given $x(E)$ in the state $m = \langle \phi, \cdot \phi \rangle$ as the Borel measure*

$$F \rightarrow m'(x(E), y(F)) = \langle \phi, x(E)\phi \rangle \langle x(E)\phi, y(F)x(E)\phi \rangle / ||x(E)\phi||^2$$

assuming $x(E)\phi \neq 0$. Notice that

$$m'(x(E), y(F)) = \langle \phi, x(E)y(F)x(E)\phi \rangle.$$

This may be interpreted as the probability of first obtaining a value in $E \in B(R)$ for a measurement of x and then a value in $F \in B(R)$ for a measurement of y . Notice that $m'(x(E), y(F))$ need not equal $m'(y(F), x(E))$ and thus this probability depends on the order. Alternatively the first order probability of y given $x(E)$ in the state m may be defined as the random variable $f'_{x(E),y}$ which satisfies

$$\int_F f'_{x(E),y}(\lambda) m(y(d\lambda)) = m'(x(E), y(F))$$

for all $F \in B(R)$. We may now define a second order probability which results from a measurement of x and y and then x again. Thus the *second order probability of y given $x(E)$ in the state m* is

$$F \rightarrow m^2(x(E), y(F)) = \langle \phi, x(E)y(F)x(E)y(F)x(E)\phi \rangle.$$

Notice m^2 need not equal m' . Continuing in this way we may define n th order probabilities m^n , $n = 1, 2, \dots$. It is interesting to note that if the type 1 joint probability $m_{x,y}$ of x and y exists in the state m then the order probabilities are all equal to $\langle \phi, x(E)y(F)\phi \rangle = \langle \phi, y(F)x(E)\phi \rangle = m_{x,y}(E \times F)$ and thus the measurements can be made in any order. Also it is interesting to note that $\lim_{n \rightarrow \infty} m^n(x(E), y(F)) = \lim_{n \rightarrow \infty} m^n(y(F), x(E)) = \langle \phi, x(E) \wedge y(F)\phi \rangle$. Thus the probability that x has a value in E and y has a value in F may be found operationally by computing successive order probabilities.

REFERENCES

- [1] H. CRAMER & H. WOLD, Some theorems on distribution functions, *J. London Soc.*, **11** (1936) 290–294.
- [2] A. GLEASON, Measures on closed subspaces of a Hilbert space, *J. Rat. Mech. Anal.*, **6** (1957) 885–893.
- [3] S. GUDDER, Uniqueness and existence properties of bounded observables, *Pac. J. Math.*, **19** (1966) 81–93.
- [4] ———, Hilbert space, independence and generalized probability, *J. Math. Anal. Appl.*, **20** (1967) 48–61.
- [5] H. MORGENTHAU, Measurements in quantum mechanics, *Annals Phys.*, **23** (1963) 469–485.
- [6] J. MOYAL, Quantum mechanics as a statistical theory, *Proc. Camb. Phil. Soc.*, **45** (1949) 99–124.
- [7] J. VON NEUMANN, *Functional Operators Vol. II*, Princeton University Press, 1950.
- [8] K. URBANIK, Joint distributions of observables in quantum mechanics, *Stud. Math.*, **21** (1961) 117–131.
- [9] V. VARADARAJAN, Probability in physics and a theorem on simultaneous observability, *Comm. Pure Appl. Math.*, **15** (1962) 189–217.

University of Wisconsin

Date communicated: JANUARY 4, 1968