International Journal of Theoretical Physics, Vol. 27, No. 2, 1988

Realistic Quantum Probability

Stanley Gudder[1]

Received July 13, 1987

A mathemetical framework for a realistic quantum probability theory is presented. The basic elements of this framework are measurements and amplitudes. Definitions of the various concepts are motivated by guidelines from the path integral formalism for quantum mechanics. The operational meaning of these concepts is discussed. Superpositions of amplitude functions are investigated and superselection sectors are shown to occur in a natural way. It is shown that this framework includes traditional nonrelativistic quantum mechanics as a special case. Proofs of most of the theorems will appear elsewhere.

1. INTRODUCTION

One of the important unsolved problems in modern physics is to find a unification of general relativity theory and quantum mechanics. During the past 20 years a considerable amount of research has been devoted toward attempting to modify general relativity theory to make it fit with quantum mechanics. These attempts have not been entirely successful. It has recently been suggested that perhaps we should turn the situation around and try to modify quantum mechanics to make it fit with general relativity (Penrose and Isham, 1986). In this work we outline such a possibility.

The framework that we present is not really a modification of traditional quantum mechanics, but is a generalization or extension of it. This formulation does not contradict traditional quantum mechanics, but includes it as a special case. One of the main differences between general relativity and traditional quantum mechanics is that the former is a realistic theory, while the latter is nonrealistic. It is our view that this difference must be reconciled if the two theories are to be melded. In a realistic theory, a physical system possesses its various properties or attributes independent of their measurement. That is, the properties of a system have meaning (both physical and

[1]Department of Mathematics and Computer Science, University of Denver, Denver, Colorado 80208.

0020-7748/88/0200-0193$06.00/0

mathematical) even if they are not observed. In the usual interpretations of traditional quantum mechanics, such a meaning cannot be given (for example, the unobserved position and momentum of an electron).

Although the path integral formalism (Feynman, 1948; Feynman and Hibbs, 1965) does allow a realistic interpretation for quantum mechanics, it is not mathematically rigorous and we cannot be sure that it gives consistent results (Gudder, 1988; Prugovečki, 1984). Nevertheless, this formalism is frequently employed in calculations of high-energy physics and often the results agree with experiment (Feynman, 1949; Ryder, 1985). Because of this agreement, many investigators are confident that the path integral formalism contains a germ of the truth. For these reasons, we shall employ this formalism as a guide for constructing our defining concepts. It is hoped that the resulting framework will not only provide a means of unifying general relativity and quantum mechanics, but will give a rigorous formulation of quantum field theory in terms of a mathematically well-defined path integral formalism.

2. GUIDELINES

Borrowing from ideas of the path integral formalism, we adopt the following guidelines.

1. We accept the fact that probabilities in quantum mechanics are computed in terms of (probability) amplitudes.
2. An outcome of a measurement is the result of various interfering alternatives and each of these alternatives possesses an amplitude for occurring.
3. The amplitude of an outcome is the "sum" of the amplitudes of the alternatives from which it results.
4. The probability of an outcome is the modulus squared of its amplitude.
5. The probability of an event for a measurement is the "sum" of the probabilities of the outcomes composing it.

We now amplify the meaning of the above guidelines. We first assume that a physical system (object) is in precisely one of a set of possible configurations (alternatives, potentialities) and that each configuration has a probability amplitude of occurring. When the system interacts with a measuring apparatus, an outcome results. In general, a particular outcome can result from many interfering configurations. By interfering we mean that the configurations cannot be distinguished without disturbing the system; that is, other measurements would be needed. The amplitude of an outcome is found by summing (in the discrete case) or integrating (in the continuum case) the amplitudes of the configurations that result in that outcome upon executing the measurement. The first four guidelines then complete the usual axioms of quantum probability in accordance with the "sum over histories" formalism (Feynman, 1948; Feynman and Hibbs, 1965). The fifth guidelines is based on classical probability theory. This is because we are now describing a completed measurement, so no interference is in effect. That is, the various outcomes are already distinguished by the original measurement itself. In this case, an outcome is considered to be an irredcucible (elementary) event and a general event is a union of outcomes. In classical probability theory, the probability of an event is the sum of the probabilities of the outcomes composing it (in the discrete case) or the integral of a probability density (in the continuum case).

3. MATHEMATICAL FRAMEWORK

We now present a rigorous mathematical framework based on the previous guidelines. This framework can then be used to construct a mathematical model for describing a particular physical system.

Let X be a nonempty set called a *sample space* and whose elements we call *sample points.* The sample points correspond to the possible configurations of a physical system S. In practice, it would be impossible to describe the configurations delineating all the properties of S (for example, some properties might be unknown), so configurations are limited to those properties on which we wish to focus. A *measurement* is a map F from X onto its range $Y_F = F(X)$ satisfying:

(M1) Y_F is the base space of a measure space (Y_F, Σ_F, ν_F).

(M2) For every $y \in Y_F$, $F^{-1}(y)$ is the base space of a measure space $(F^{-1}(y), \Sigma_y, \mu_y)$.

We call $F^{-1}(y)$ the *fiber over y*, the elements of Y_F are called *F-outcomes*, and the sets in Σ_F are *F-events.* Notice that

$$\mathscr{E}(F) \equiv \{F^{-1}(B): B \in \Sigma_F\}$$

is a σ-algebra of subsets of X. A measurement F corresponds to a laboratory procedure or experiment that can be performed on S. For every $x \in X$, $F(x)$ denotes the outcome resulting from executing F when S has configuration x. For $y \in Y_F$, the fiber $F^{-1}(y)$ is the set of sample points that result in outcome y, and for $B \in \Sigma_F$, $F^{-1}(B)$ is the set of sample points that result in the event B, when F is executed. The measure ν_F is an *a priori* weight for the F-events that is independent of the state of S. In case of total ignorance, ν_F is a uniform measure such as Lebesgue measure, Haar measure, or the counting measure in the discrete case. Similarly, μ_y is an *a priori* weight for the sample points in the fiber $F^{-1}(y)$.

An example of a measurement is the position measurement $F(q,p)=q$ where

$$X=\mathbb{R}^{2n}=\{(q,p)\colon q,p\in\mathbb{R}^n\}$$

is a phase space and the measures are the usual Lebesgue measures. In this case, the fiber $F^{-1}(q)$ is the "vertical" set $q\times\mathbb{R}^n$. We consider this example in detail later. As a second example, let $X=\{x(t)\colon t\in\mathbb{R}\}$ be a set of trajectories for a particle. For each time t, define the measurement $F_t(x)=x(t)$. For $y\in Y_{F_t}$, the fiber

$$F_t^{-1}(y)=\{x\in X\colon x(t)=y\}$$

consists of all trajectories that have position y at time t. Of course, care must be taken in defining the corresponding measure spaces (we shall not consider the technical details here).

A function $f\colon X\to\mathbb{C}$ is an *amplitude density* for the measurement F if the following conditions hold:

(A1) $f|F^{-1}(y)\in L^1(F^{-1}(y),\Sigma_y,\mu_y)$ for every $y\in Y_F$.

(A2) $F(f)(y)\equiv\int_{F^{-1}(y)} f\,d\mu_y\in L^2(Y_F,\Sigma_F,\nu_F)\equiv H_F$.

(A3) $\|F(f)\|=\int|F(f)|^2\,d\nu_F=1$.

We interpret $f(x)$ as the probability amplitude density at the configuration x (Guideline 2). Since $F^{-1}(y)$ is interpreted as the set of configurations that result in outcome y upon execution F, $F(f)(y)$ corresponds to "summing" the amplitudes over these alternatives. We call $F(f)$ the *F-wave function* for f. Now, $F(f)(y)$ gives the amplitude density of y (Guideline 3) and the resulting probability density is $|F(f)(y)|^2$ (Guideline 4). Axiom (A3) is motivated by Guideline 5.

We now summarize the physical interpretations of our previous concepts. The sample space X represents an underlying objective physical reality. An amplitude density f determines a model for the physical reality that enables us to compute probabilities for outcomes and events. The Hilbert space H_F gives a "projection" of physical reality resulting from executing the measurement F. If various measurements $F, G, \ldots$ are executed, the Hilbert spaces $H_F, H_G, \ldots$ each give a view of physical reality X, but in general, miss a complete description of X.

Let F be a measurement and let f be an amplitude density for F. A set $A\subseteq X$ is a *generalized* (F,f) *event* if the following conditions hold:

(E1) $A\cap F^{-1}(y)\in\Sigma_y$ for every $y\in Y_F$.

(E2) $f_F(A)(y)\equiv\int_{A\cap F^{-1}(y)} f\,d\mu_y\in H_F$.

We denote the set of generalized (F, f) events by $\mathscr{E}(F, f)$. The elements of $\mathscr{E}(F, f)$ are the subsets of X for which a reasonable amplitude density $f_F(A)$ can be defined. In fact, $f_F(A)(y)$ is the "sum" of the amplitudes over the configurations in A that result in the outcome y upon executing F. Interpreting $|f_F(A)|^2$ as the probability density of $A \in \mathscr{E}(F, f)$, it becomes reasonable to define the (F, f) *probability* of A as

$$P_{F,f}(A) = \int |f_F(A)|^2 \, d\nu_F = \|f_F(A)\|^2$$

We also define the (F, f) *pseudoprobability* of $A \in \mathscr{E}(F, f)$ as

$$P'_{F,f}(A) = \int f_F(A)\bar{f}_F(X) \, d\nu_F = \langle f_F(A), f_F(X)\rangle = \langle f_F(A), F(f)\rangle$$

A nonempty collection $\mathscr{S}$ of subsets of X is an *additive class* if $\mathscr{S}$ is closed under the formation of complements and finite disjoint unions. Moreover, if $\mathscr{S}$ is closed under the formation of countable disjoint unions, then $\mathscr{S}$ is a *σ-additive* class. We denote the complement of a set A by A^c and its characteristic function by χ_A.

Lemma 3.1. (a) $\mathscr{E}(F, f)$ is an additive class containing $\mathscr{E}(F)$.

(b) For every $B \in \Sigma_F$,

$$P_{F,f}[F^{-1}(B)] = P'_{F,f}[F^{-1}(B)] = \int_B |F(f)|^2 \, d\nu_F$$

(c) $P'_{F,f}$ is an additive, complex-valued set function on $\mathscr{E}(F, f)$ with $P'_{F,f}(X) = 1$.

We use the notation $P_{F,f}(B) \equiv P_{F,f}[F^{-1}(B)]$ for $B \in \Sigma_F$. It is clear that $P_{F,f}$ is a probability measure on Σ_F and we call it the *f-distribution* of F. Although $P'_{F,f}$ is additive on $\mathscr{E}(F, f)$, it has the disadvantage of being complex-valued, so it cannot be interpreted as a probability for an arbitrary $A \in \mathscr{E}(F, f)$. On the other hand, $P_{F,f}$ is nonnegative, but it is not necessarily additive and it may attain values larger than 1 on $\mathscr{E}(F, f)$. Hence, $P_{F,f}$ is not a probability measure on $\mathscr{E}(F, f)$ in general. We now show that these difficulties can be overcome under certain conditions.

Let $\mathscr{S} \subseteq \mathscr{E}(F, f)$ be an additive class. We say that an amplitude density f is *F-orthogonally scattered over* $\mathscr{S}$ if for $A, B \in \mathscr{S}$ with $A \cap B = \varnothing$ we have $f_F(A) \perp f_F(B)$ [i.e., $\langle f_F(A), f_F(B)\rangle = 0$]. We say that f is *F-orthogonal* at $A \in \mathscr{E}(F, f)$ if $f_F(A) \perp f_F(A^c)$. Notice that if f is F-orthogonally scattered over $\mathscr{S}$, then f is F-orthogonal at every $A \in \mathscr{S}$. Also, f is F-orthogonal at A if and only if f is F-orthogonally scattered over $\{X, \varnothing, A, A^c\}$. When we say that μ is a probability measure on an additive class $\mathscr{S}$, we mean that

μ is σ-additive. That is, if $A_i \in \mathcal{S}$, $i = 1, 2, \ldots$, are mutually disjoint and $\bigcup A_i \in \mathcal{S}$, then $\mu(\bigcup A_i) = \Sigma_i \mu(A_i)$.

Theorem 3.2. Let f be an amplitude density for the measurement F.

(a) f is F-orthogonally scattered over $\mathscr{E}(F)$.

(b) For $A \in \mathscr{E}(F, f)$, $P_{F,f}(A) = P'_{F,f}(A)$ if and only if f is F-orthogonal at A.

(c) If $\mathcal{S} \subseteq \mathscr{E}(F, f)$ is an additive class, the following statements are equivalent:

1. f is F-orthogonal at every $A \in \mathcal{S}$.
2. $P_{F,f}(A) = P'_{F,f}(A)$ for every $A \in \mathcal{S}$.
3. $P_{F,f}$ and $P'_{F,f}$ are probability measures on $\mathcal{S}$.

(d) If f is F-orthogonal at every $A \in \mathscr{E}(F, f)$, then $\mathscr{E}(F, f)$ is a σ-additive class and $P_{F,f} = P'_{F,f}$ is a probability measure on $\mathscr{E}(F, f)$.

If $A \in \mathscr{E}(F, f)$ with $f_F(A) \neq 0$ and $B \in \Sigma_F$, we define the *conditional* (F, f) *probability of* B *given* A *by*

$$P_{F,f}(B|A) \equiv \int_B |f_F(A)|^2 \, d\nu_F / P_{F,f}(A) = \|\chi_B f_F(A)\|^2 / \|f_F(A)\|^2$$

Notice that $P_{F,f}(\cdot|A)$ is a probability measure on Σ_F. We now show that $P_{F,f}(B|A)$ has the usual properties of a conditional probability.

Lemma 3.3. (a) $P_{F,f}(B|X) = P_{F,f}(B)$ for all $B \in \Sigma_F$.

(b) If $B \in \Sigma_F$, $A \in \mathscr{E}(F, f)$, then $F^{-1}(B) \cap A \in \mathscr{E}(F, f)$, and

$$P_{F,f}(B|A) = P_{F,f}[F^{-1}(B) \cap A] / P_{F,f}(A)$$

(c) If $B, C \in \Sigma_F$, then

$$P_{F,f}(B|F^{-1}(C)) = P_{F,f}(B \cap C) / P_{F,f}(C)$$

We denote by $L^2(F, f)$ the set of functions $g : X \to \mathbb{R}$ satisfying the following conditions:

(F1) $g^{-1}(B) \in \mathscr{E}(F, f)$ for every $B \in B(\mathbb{R})$.

(F2) $gf|F^{-1}(y) \in L^1(F^{-1}(y), \Sigma_y, \mu_y)$ for every $y \in Y_F$.

(F3) $f_F(g)(y) \equiv \int_{F^{-1}(y)} gf \, d\mu_F \in H_F$.

Notice that for $A \in \mathscr{E}(F, f)$ we have $\chi_A \in L^2(F, f)$ and $f_F(\chi_A) = f_F(A)$, so $f_F(g)$ generalizes the amplitude density of generalized events. We denote by $L^2_W(F, f)$ the set of functions $g : X \to \mathbb{R}$ satisfying (F2) and (F3). It is clear that $L^2_W(F, f)$ is a linear space. However, $L^2(F, f)$ need not be linear, since it is not necessarily closed under summation (Gudder, 1984).

For $g \in L^2_W(F, f)$ we define the (F, f) *pseudoexpectation* of g by

$$E_{F,f}(g) = \langle f_F(g), F(f) \rangle = \langle f_F(g), f_F(1) \rangle$$

where $1 = \chi_X$. It is clear that $E_{F,f}(g)$ is linear on $L^2_W(F, f)$ and $E_{F,f}(1) = 1$. Since $E_{F,f}(\chi_A) = P'_{F,f}(A)$ for all $A \in \mathscr{E}(F, f)$, we see that $E_{F,f}$ is the natural linear extension of $P'_{F,f}$. Unfortunately, $E_{F,f}(g)$ need not be real and $E_{F,f}(g)$ need not be nonnegative when g is nonnegative. We now show that these difficulties do not occur under certain conditions. Suppose $g \in L^2(F, f)$. Then

$$\mathscr{E}_g(F, f) = \{g^{-1}(B) \colon B \in B(\mathbb{R})\}$$

is a σ-subalgebra of $\mathscr{E}(F, f)$. If f is F-orthogonal at every $A \in \mathscr{E}_g(F, f)$, then by Theorem 3.2(c), $P_{F,f}$ is a probability measure on $\mathscr{E}_g(F, f)$, so $(g^{-1}(\mathbb{R}), \mathscr{E}_g(F, f), P_{F,f})$ becomes a probability space. We then define the Lebesgue integral $\int g(x)\, P_{F,f}(dx)$ in the usual way.

Theorem 3.4. (a) If $g \in L^2(F, f)$ and f is F-orthogonal at every $A \in \mathscr{E}_g(F, f)$, then

$$E_{F,f}(g) = \int g(x)\, P_{F,f}(dx)$$

(b) If $g \colon Y_F \to \mathbb{R}$ is Σ_F-measurable and $gF(f) \in H_F$, then $g \circ F$ satisfies the conditions of (a) and

$$E_{F,f}(g) \equiv E_{F,f}(g \circ F) = \int g(y)\, P_{F,f}(dy) = \langle gF(f), F(f) \rangle$$

Under the conditions of Theorem 3.4, we see that $E_{F,f}$ has the desirable properties of an expectation. We can also define pseudo-conditional expectations in a natural way. Let $A \in \mathscr{E}(F, f)$, $g \in L^2_W(F, f)$, and define

$$f_F(g \mid A)(y) = \int_{F^{-1}(y) \cap A} gf\, d\mu\, y$$

The (F, f) *pseudo-conditional expectation of* g *given* A is defined as

$$E_{F,f}(g \mid A) = \frac{\langle f_F(g \mid A), f_F(1 \mid A) \rangle}{P_{F,f}(A)} = \frac{\langle f_F(g \mid A), f_F(A) \rangle}{\|f_F(A)\|^2}$$

Notice that if $g = \chi_{F^{-1}(B)}$ for $B \in \Sigma_F$, then

$$f_F(\chi_{F^{-1}(B)} \mid A) = \chi_{F^{-1}(B)} f_F(A)$$

and

$$E_{F,f}(\chi_{F^{-1}(B)} \mid A) = P_{F,f}(B \mid A)$$

4. OPERATIONAL MEANING

In Section 3 we defined probabilities, conditional probabilities, and expectations. We now discuss their operational meaning; that is, the laboratory procedures that are modeled by the mathematical definitions. For brevity, we consider the probability $P_{F,f}(A)$ of a generalized event A, since the other definitions have similar meanings.

For simplicity, suppose that the sample space X is a finite set and that $F: X \to Y_F$ is a measurement with outcome set $Y_F = \{y_1, \ldots, y_n\}$. We assume that the measures ν_F, μ_{y_i}, $i = 1, \ldots, n$, are counting measures. Thus, for example, $\nu_F(y_i) = 1$, $i = 1, \ldots, n$. We interpret $P_{F,f}(y_i)$ as the long-run relative frequency of y_i. That is, we execute the measurement F a large number of times N, and each time the system is reconstructed in accordance with the amplitude density f. If y_i results n_i times, then $P_{F,f}(y_i) \cong n_i/N$. Now let $A \subseteq X$ be a laboratory event; that is, an event that can be prepared in the laboratory. Then A^c is also a laboratory event, since A^c occurs if and only if A does not occur. We interpret $P_{F,f}(y_i \mid A)$ as the long-run relative frequency of y_i when A is prepared. Operationally, this is similar to $P_{F,f}(y_i)$ except that we prepare the event A each time before executing F. We thus have the operationally significant quantities $P_{F,f}(y_i)$, $P_{F,f}(y_i \mid A)$, $P_{F,f}(y_i \mid A^c)$, $i = 1, \ldots, n$, and any other quantities that can be defined in terms of these are themselves operationally significant.

We interpret $P_{F,f}(A)$ as the probability of A as viewed by the measurement F when the system is described by the amplitude density f. Thus, executions of F provide information about the likelihood of an A occurrence and this information is contained in $P_{F,f}(A)$. A different measurement G would, in general, provide a different view of the system and give its own version $P_{G,f}(A)$ for the probability of A. We interpret $|f_F(A)(y_i)|^2$ as the probability that A occurs *and* the outcome y_i results upon executing F, $i = 1, \ldots, n$. We must now give an operational meaning for $|f_F(A)(y_i)|^2$; that is, we must define this quantity in terms of our previously defined operationally significant quantities. Once this is done, we shall have an operational meaning for

$$P_{F,f}(A) = \sum_{i=1}^{n} |f_F(A)(y_i)|^2$$

For $i = 1, \ldots, n$, we write

$$|f_F(A)(y_i)|^2 = P_{F,f}(y_i) P(A; y_i)$$

where we interpret $P(A; y_i)$ as the probability of A as viewed by y_i; that is, the probability that A would have occurred if an execution of F results in the outcome y_i. We now give an operational meaning for $P(A; y_i)$,

$i=1,\ldots,n$. The operational content of $P(A;y_i)$ is contained in the equation

$$P_{F,f}(y_i)=P_{F,f}(A;y_i)P_{F,f}(y_i\,|\,A)+P_{F,f}(A^c;y_i)P_{F,f}(y_i\,|\,A^c) \tag{4.1}$$

Notice that (4.1) is a type of Bayesian condition. It is also operationally reasonable to assume that

$$P_{F,f}(A;y_i)+P_{F,f}(A^c;y_i)=1 \tag{4.2}$$

For simplicity, we shall assume that

$$P_{F,f}(y_i\,|\,A)\neq P_{F,f}(y_i\,|\,A^c) \tag{4.3}$$

for every $i=1,\ldots,n$. Equation (4.3) essentially states that A and F are stochastically dependent. In fact, Lemma 4.1 will show that in classical probability theory equation (4.3) is equivalent to stochastic dependence. If (4.3) does not hold, then A and the outcome y_i would be stochastically independent and y_i would give no additional information concerning A. Under the assumption of (4.3) we can apply (4.1) and (4.2) and solve for $P_{F,f}(A;y_i)$ to obtain

$$P_{F,f}(A;y_i)=\frac{P_{F,f}(y_i)-P_{F,f}(y_i\,|\,A^c)}{P_{F,f}(y_i\,|\,A)-P_{F,f}(y_i\,|\,A^c)} \tag{4.4}$$

We thus see that $P_{F,f}(A;y_i)$, $i=1,\ldots,n$, are operationally significant and hence $P_{F,f}(A)$ has an operational meaning given by

$$P_{F,f}(A)=\sum_{i=1}^{n}P_{F,f}(y_i)P(A;y_i)$$

The next lemma shows that (4.4) holds in classical probability theory, and in this case for every i,j we have

$$P_{F,f}(A;y_i)=P_{F,f}(A;y_j)$$

Lemma 4.1. Let A, B be events in a probability space (Ω,Σ,P) satisfying $P(A)$, $P(A^c)\neq 0$.

(a) $P(B\,|\,A)=P(B\,|\,A^c)$ if and only if A and B are stochastically independent.

(b) If $P(B\,|\,A)\neq P(B\,|\,A^c)$, then

$$P(A)=\frac{P(B)-P(B\,|\,A^c)}{P(B\,|\,A)-P(B\,|\,A^c)}$$

Proof. (a) If A and B are stochastically independent, then A^c and B are stochastically independent, so we have

$$P(B\,|\,A)=P(B)=P(B\,|\,A^c)$$

Conversely, if $P(B\,|\,A)=P(B\,|\,A^c)$, then

$$\frac{P(B\cap A)}{P(A)}=\frac{P(B\cap A^c)}{P(A^c)}=\frac{P(B)-P(B\cap A)}{1-P(A)} \tag{4.5}$$

Solving (4.5) gives $P(B \cap A) = P(A)P(B)$ and hence A and B are stochastically independent.

(b) It follows from Bayes' law that

$$\begin{aligned} P(B) &= P(A)P(B|A) + P(A^c)P(B|A^c) \\ &= P(A)P(B|A) + [1 - P(A)]P(B|A^c) \\ &= P(A)[P(B|A) - P(B|A^c)] + P(B|A^c) \end{aligned}$$

and the result follows. ■

5. CATALOGS AND AMPLITUDE SPACES

Section 3 considered single measurements on a sample space X. We now consider a *catalog* $\mathscr{A}(X)$, which is a nonempty collection of measurements on X. A function $f: X \to \mathbb{C}$ is an *amplitude density* for $\mathscr{A}(X)$ if f is an amplitude density for every $F \in \mathscr{A}(X)$. Denote the set of all amplitude densities on $\mathscr{A}(X)$ by $\mathscr{F}(\mathscr{A})$. We say that a catalog $\mathscr{A}(X)$ is *complete* if for any $x \neq y$ in X there is an $F \in \mathscr{A}(X)$ such that $F(x) \neq F(y)$.

Let $\mathscr{A}(X)$ be a catalog. The set of functions $f: X \to \mathbb{C}$ satisfying (A1), (A2) of Section 3 for all $F \in \mathscr{A}(X)$ and

$$\text{(A3}') \quad \|f\| \equiv \|F(f)\|_{H_F} = \|G(f)\|_{H_G} \quad \text{for all} \quad F, G \in \mathscr{A}(X)$$

is called the *amplitude space* of $\mathscr{A}$ and is denoted $H(\mathscr{A})$. The elements of $H(\mathscr{A})$ are called *amplitude functions.* Notice that

$$H(\mathscr{A}) = \{\alpha f: \alpha \in \mathbb{C}, f \in \mathscr{F}(\mathscr{A})\}$$

and if $f \in H(\mathscr{A})$ with $\|f\| \neq 0$, then $f/\|f\| \in \mathscr{F}(\mathscr{A})$. Of course, if $f \in H(\mathscr{A})$, then $\alpha f \in H(\mathscr{A})$ for every $\alpha \in \mathbb{C}$. However, if $f, g \in H(\mathscr{A})$, we need not have $f + g \in H(\mathscr{A})$. If $\alpha f + \beta g \in H(\mathscr{A})$, $f, g \in H(\mathscr{A})$, and $\alpha, \beta \in \mathbb{C}$, we call $\alpha f + \beta g$ a *superposition* of f and g. We now characterize pairs $f, g \in H(\mathscr{A})$ for which superpositions are possible. For $f, g \in H(\mathscr{A})$ we write fsg if for every $F, G \in \mathscr{A}(X)$ we have

$$\langle F(f), F(g)\rangle_{H_F} = \langle G(f), G(g)\rangle_{H_G}$$

Notice that s is a reflexive, symmetric relation and fsg implies $(\alpha f)s(\beta g)$ for all $\alpha, \beta \in \mathbb{C}$.

Lemma 5.1. Let $f, g \in H(\mathscr{A})$. Then fsg if and only if $f + g, f + ig \in H(\mathscr{A})$.

It follows from this lemma that fsg if and only if $\alpha f+\beta g\in H(\mathscr{A})$ for every $\alpha, \beta\in\mathbb{C}$. For $A\subseteq H(\mathscr{A})$ we write

$$A^s=\{g\in H(\mathscr{A}): gsf \text{ for all } f\in A\}$$

A set $A\subseteq H(\mathscr{A})$ is a *superposition set* (or *s-set*) if $A\subseteq A^s$. Every s-set is contained in a maximal s-set. Moreover, A is a maximal s-set if and only if $A=A^s$ (Gudder, 1986). Denote the set of maximal s-sets by $M(H)$. Let $A\in M(H)$. If $f\in A$, $\alpha\in\mathbb{C}$, then $(\alpha f)sg$ for every $g\in A$. Since A is maximal, $\alpha f\in A$. If $f, g\in A$, then $f+g\in H(\mathscr{A})$ and $(f+g)sh$ for every $h\in A$. Again, by maximality, $f+g\in A$. Hence, A is a linear space. For $f, g\in A$, define $\langle f, g\rangle=\langle F(f), F(g)\rangle$ for every $F\in\mathscr{A}$. Then it is clear that $\langle f, g\rangle$ is an inner product on A. Thus, each $A\in M(H)$ forms an inner product space. Moreover, it is clear that for every $A\in M(H)$ and $F\in\mathscr{A}$, $F: A\to H_F$ is a linear isometry. The pair $(H(\mathscr{A}), s)$ gives a *partial inner product space* (Gudder, 1986). One can complete $(H(\mathscr{A}), s)$ in a natural way to obtain a partial Hilbert space. Such structures generalize direct sums of Hilbert spaces (Gudder, 1986). We call the sets $A\in M(H)$ *superselection sectors.*

We can apply our work in Section 3 and compute probabilities and expectations for a catalog $\mathscr{A}=\mathscr{A}(X)$. Let $F, G\in\mathscr{A}$ and let $f\in\mathscr{F}(\mathscr{A})$. If $A\in\Sigma_G$ and $G^{-1}(A)\in\mathscr{E}(F,f)$, then $P_{F,f}[G^{-1}(A)]$ is interpreted as the "probability" of the G-event A upon execution of an F-measurement. Operationally, this means that if a G-measurement is executed followed by an F-measurement, then $P_{F,f}[G^{-1}(A)]$ is the "probability" that G results in an outcome in A. If $A\in\Sigma_G$, $B\in\Sigma_F$, and $G^{-1}(A)\in\mathscr{E}(F,f)$, then $P_{F,f}(B\,|\,G^{-1}(A))$ is the probability that an F-measurement results in an outcome in B given that a previous G-measurement resulted in an outcome in A. If G is a real-valued measurement with $G\in L^2_W(F,f)$, then $E_{F,f}(G)$ is interpreted as the "expectation" of G determined by an F-measurement. In general this differs from the ordinary expectation $E_{G,f}(G)$ of G.

6. PHASE SPACE MODEL

This section presents an amplitude phase space model for a simple quantum system. The system consists of a single nonrelativistic, spinless particle constrained to one dimension. (The model can easily be generalized to three dimensions.) We take for our sample space the two-dimensional phase space

$$X=\mathbb{R}^2=\{(q,p): q, p\in\mathbb{R}\}$$

Define the measurements $Q: X\to\mathbb{R}$, $P: X\to\mathbb{R}$ by $Q(q,p)=q$, $P(q,p)=p$ and let $\Sigma_Q=\Sigma_P=B(\mathbb{R})$, $d\nu_Q=dq$, $d\nu_P=dp$. On the fiber $Q^{-1}(q)=q\times\mathbb{R}$ we let $\Sigma_q=q\times B(\mathbb{R})$, $d\mu_q=dp$ and on $P^{-1}(p)=\mathbb{R}\times p$ we let $\Sigma_p=B(\mathbb{R})\times p$,

$d\mu_p = dq$. Of course, Q and P correspond to position and momentum measurements, respectively. Then $\mathscr{A} = \{Q, P\}$ is a complete catalog on X. If $f \in \mathscr{F}(\mathscr{A})$, we have

$$Q(f)(q) = \int f(q, p)\, dp \in L^2(\mathbb{R}, dq)$$

$$P(f)(p) = \int f(q, p)\, dq \in L^2(\mathbb{R}, dp)$$

$$\|Q(f)\| = \|P(f)\| = 1$$

We now construct a class of physical amplitude densities that correspond to the traditional quantum states. For

$$\psi \in L^2(\mathbb{R}, dq) \cap L^1(\mathbb{R}, dq) \tag{6.1}$$

we denote the Fourier transform by

$$\hat{\psi}(p) = (2\pi\hbar)^{-1/2} \int \psi(q)\, e^{-iqp/\hbar}\, dq$$

and the inverse Fourier transform by

$$\check{\psi}(q) = (2\pi\hbar)^{-1/2} \int \psi(p)\, e^{iqp/\hbar}\, dp$$

We say that $f \in \mathscr{F}(\mathscr{A})$ is *regular* if $f_Q(A \times \mathbb{R})^\wedge = f_P(A \times \mathbb{R})$ and $f_Q(\mathbb{R} \times A)^\wedge = f_P(\mathbb{R} \times A)$ for every $A \in B(\mathbb{R})$. It is shown in Gudder (1985) that f is regular if and only if (1) for every $p \in \mathbb{R}$

$$f(q, p) = (2\pi\hbar)^{-1/2} Q(f)(q)\, e^{-iqp/\hbar} \qquad \text{a.e. } [q]$$

and (2) for every $q \in \mathbb{R}$

$$f(q, p) = (2\pi\hbar)^{-1/2} Q(f)^\wedge(p)\, e^{iqp/\hbar} \qquad \text{a.e. } [p]$$

It is not clear that regular amplitude densities exist, and from conditions 1 and 2 we see that if they exist, they must be nonmeasurable. Nevertheless, it is shown in Gudder (1984) that for every ψ satisfying (6.1) there exists a regular f such that $Q(f) = \psi$. Intuitively, the regular f are those for which the Fourier transform of position is momentum. Moreover, $\psi = Q(f)$ is a traditional quantum state.

Let f be regular with $Q(f) = \psi$. We then have

$$P(f)(p) = \int f(q, p)\, dq = (2\pi\hbar)^{-1/2} \int \psi(q)\, e^{-iqp/\hbar}\, dq = \hat{\psi}(p)$$

Hence, for every $A \in B(\mathbb{R})$ we have

$$P_{Q,f}(A) = \int_A |\psi(q)|^2 \, dq, \qquad P_{P,f}(A) = \int_A |\hat{\psi}(p)|^2 \, dp$$

which are the usual quantum mechanical formulas. Moreover, we have

$$f_Q[P^{-1}(A)](q) = \int_A f(q,p)\, dp = (2\pi\hbar)^{-1/2} \int_A \hat{\psi}(p)\, e^{iqp/\hbar}\, dp = (\chi_A \hat{\psi})^\vee(q)$$

Hence,

$$P_{Q,f}[P^{-1}(A)] = \|(\chi_A \hat{\psi})^\vee\|^2 = \|\chi_A \hat{\psi}\|^2 = \int_A |\hat{\psi}(p)|^2\, dp$$

We can thus get information about P by measuring Q. Notice that $P_{Q,f}[P^{-1}(\cdot)]$ is a probability measure. It follows that f is Q-orthogonal on $\mathscr{E}(P)$. In fact, f is Q-orthogonally scattered on $\mathscr{E}(P)$. Indeed, suppose $A, B \in B(\mathbb{R})$ and $A \cap B = \varnothing$. Then $(\mathbb{R} \times A) \cap (\mathbb{R} \times B) = \varnothing$ and we have

$$\langle f_Q[P^{-1}(A)], f_Q[P^{-1}(B)]\rangle = \langle (\chi_A \hat{\psi})^\vee, (\chi_B \hat{\psi})^\vee \rangle = \langle \chi_A \hat{\psi}, \chi_B \hat{\psi}\rangle = 0$$

Similarly,

$$P_{P,f}[Q^{-1}(A)] = \int_A |\psi(q)|^2\, dq$$

and f is P-orthogonally scattered on $\mathscr{E}(Q)$.

If $A, B \in B(\mathbb{R})$, then

$$\begin{aligned} P_{Q,f}(A \mid P^{-1}(B)) &= \chi_A f_Q[P^{-1}(B)]\|^2 / \|f_Q[P^{-1}(B)]\|^2 \\ &= \|\chi_A (\chi_B \hat{\psi})^\vee\|^2 / \|\chi_B \hat{\psi}\|^2 \end{aligned}$$

Similarly,

$$P_{P,f}(A \mid Q^{-1}(B)) = \|\chi_A (\chi_B \psi)^\wedge\|^2 / \|\chi_B \psi\|^2$$

It is shown in Gudder (1984) that these reduce to the traditional von Neumann–Lüders formulas

$$P_{Q,f}(A \mid P^{-1}(B)) = \mathrm{tr}[E^Q(A) E^P(B) P_\psi E^P(B)] / [\mathrm{tr}\, E^P(B) P_\psi]$$

$$P_{P,f}(A \mid Q^{-1}(B)) = \mathrm{tr}[E^P(A) E^Q(B) P_\psi E^Q(B)] / [\mathrm{tr}\, E^Q(B) P_\psi]$$

where P_ψ is the one-dimensional projection onto ψ and E^Q, E^P are the spectral measures for Q and P, respectively.

In the sequel, we shall assume that ψ is a Schwartz test function and that f is regular with $Q(f)=\psi$. We then obtain

$$\begin{aligned} E_{P,f}(P \mid Q^{-1}(A)) &= \langle f_P(P \mid Q^{-1}(A)), f_P(Q^{-1}(A))\rangle / P_{P,f}[Q^{-1}(A)] \\ &= \langle p f_P(Q^{-1}(A)), f_P(Q^{-1}(A))\rangle / \|\chi_A \psi\|^2 \\ &= \int p |(\chi_A\psi)\hat{}|^2 \, dp \Big/ \int_A |\psi|^2 \, dq \end{aligned}$$

In particular,

$$E_{P,f}(P) = \int p |\hat{\psi}(p)|^2 \, dp$$

Similarly,

$$E_{Q,f}(Q) = \int q |\psi(q)|^2 \, dq$$

We also have

$$\begin{aligned} f_Q(P)(q) &= \int p f(q,p) \, dp \\ &= (2\pi\hbar)^{-1/2} \int p \hat{\psi}(p) \, e^{iqp/\hbar} \, dp \\ &= (2\pi\hbar)^{-1/2} \left(-i\hbar \frac{d}{dq}\right) \int \hat{\psi}(p) \, e^{iqp/\hbar} \, dp = -i\hbar \frac{d\psi}{dq}(q) \end{aligned}$$

Let $R(\mathscr{A}) \subseteq H(\mathscr{A})$ be the set of all scalar multiples of regular amplitude functions. Then it is clear that $R(\mathscr{A}) \subseteq R(\mathscr{A})^s$, so $R(\mathscr{A})$ is a linear subspace of a superselection sector of $H(\mathscr{A})$. We can then represent Q and P as operators $\tilde{Q}$ and $\tilde{P}$ on $H_Q = L^2(\mathbb{R}, dq)$ as follows:

$$\tilde{Q}\psi(q) = q\psi(q)$$

$$\tilde{P}\psi(q) = \tilde{P}Q(f)(q) = f_Q(P)(q) = (-i\hbar \, d/dq)\psi(q)$$

If we define $\tilde{\tilde{P}}\hat{\psi}(p) = p\hat{\psi}(p)$, the following theorem holds (Gudder, 1985; Prugovečki, 1984):

Theorem 6.1. If $G(q,p) = \sum a_{mn} q^m p^n$ is a polynomial, then

$$f_Q(G)(q) = \sum a_{mn} \tilde{Q}^m \tilde{P}^n \psi(q)$$

$$f_P(G)(p) = \sum a_{mn} \tilde{\tilde{P}}^n (\tilde{Q}^m \psi)\hat{}(p)$$

It follows that any polynomial $G(q,p)=\sum a_{mn}q^m p^n$ can be represented on $L^2(\mathbb{R}, dq)$ by the operator $\tilde{G}=\sum a_{mn}\tilde{Q}^m\tilde{P}^n$. Moreover,

$$E_{Q,f}(G)=\langle \sum a_{mn}\tilde{Q}^m\tilde{P}^n\psi, \psi\rangle$$

$$E_{P,f}(G)=\langle \sum a_{mn}\tilde{P}^n\tilde{Q}^m\psi, \psi\rangle$$

Thus, the pseudoexpectation of G depends on which measurement is executed. For example, let $G(q,p)=qp$. We then obtain the following version of the Heisenberg commutation relation:

$$E_{Q,f}(G)-E_{P,f}(G)=\langle[\tilde{Q},\tilde{P}]\psi,\psi\rangle = i\hbar$$

The reason that $E_{Q,f}(G)$ and $E_{P,f}(G)$ are complex-valued is that f is not Q-orthogonal or P-orthogonal on the σ-algebra $\{G^{-1}(A)\colon A\in B(\mathbb{R})\}$. The usual uncertainty relation can be obtained by taking variances.

We now consider Schrödinger's equation. Suppose the classical Hamiltonian is

$$H(q,p)=\frac{p^2}{2m}+V(q)$$

If the system is closed, then we have conservation of energy

$$H(q,p)=E \tag{6.2}$$

Now supppose $f\in\mathcal{F}(\mathcal{A})$ is regular with $\psi=Q(f)$. Taking the Q-amplitude average of (6.2) gives $f_Q[H(q,p)]=f_Q(E)$. Using the linearity of f_Q and Theorem 6.1, we obtain

$$H(\tilde{Q},\tilde{P})\psi=(\tilde{P}^2/2m)\psi+V(\tilde{Q})\psi=E\psi$$

Of course, this is the time-independent Schrödinger equation and E, ψ form an eigenpair for the operator $H(\tilde{Q},\tilde{P})$.

Now suppose the classical dynamics is generated by the Hamilton equation

$$dp/dt=-\partial H/\partial q \tag{6.3}$$

We assume that for any time $t\in\mathbb{R}$ the system is described by a regular amplitude density $f(q,p,t)$ with corresponding vector $\psi(q,t)=Q(f)(q,t)$ and, moreover, f and ψ are differentiable with respect to t. Suppose (6.3) holds in the Q-amplitude average in the sense that

$$\frac{d}{dt}\int pf(q,p,t)\,dp=-\frac{\partial}{\partial q}\int H(q,p)f(q,p,t)\,dp \tag{6.4}$$

Then (6.4) has the form

$$\frac{d}{dt} f_Q(p) = -\frac{\partial}{\partial q} f_Q(H)$$

Applying Theorem 6.1 gives

$$\frac{d}{dt}\left(-i\hbar \frac{d\psi}{dq}\right) = -\frac{\partial}{\partial q} H(\tilde{Q}, \tilde{P})\psi \tag{6.5}$$

Interchanging the order of differentiation in (6.5) gives

$$\frac{\partial}{\partial q}\left(-i\hbar \frac{\partial\psi}{\partial t}\right) = -\frac{\partial}{\partial q} H(\tilde{Q}, \tilde{P})\psi \tag{6.6}$$

Integrating both sides of (6.6), we obtain (except for a constant that we can set equal to zero)

$$i\hbar\, \partial\psi/\partial t = H(\tilde{Q}, \tilde{P})\psi$$

This, of course, is the time-dependent Schrödinger equation. We conclude that Schrödinger's equation is an amplitude-averaged version of the Hamilton equation of classical mechanics.

If we use the other Hamilton equation $dq/dt = \partial H/\partial p$ and take the P-amplitude average, we obtain

$$\frac{d}{dt} f_P(q) = \frac{\partial}{\partial p} f_P(H)$$

Proceeding in a similar way, we obtain the Fourier-transformed time-dependent Schrödinger equation

$$i\hbar \frac{\partial\hat{\psi}}{\partial t} = \frac{p^2}{2m}\hat{\psi} + (V\psi)^{\wedge}$$

REFERENCES

Feynman, R. (1948). Space-time approach to non-relativistic quantum mechanics, *Reviews of Modern Physics*, **20**, 367–398.

Feynman, R. (1949). Space-time approach to quantum electrodynamics, *Physical Review*, **76**, 769–789.

Feynman, R., and Hibbs, A. (1965). *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York.

Gudder, S. (1984). Probability manifolds, *Journal of Mathematical Physics*, **25**, 2397–2401.

Gudder, S. (1985). Amplitude phase space model for quantum mechanics, *International Journal of Theoretical Physics*, **24**, 343–353.

Gudder, S. (1986). Partial Hilbert spaces and amplitude functions, *Annales de l'Institut Henri Poincaré*, **45**, 311-326.

Gudder, S. (1988). *Quantum Probability*, Academic Press, Orlando, Florida.

Penrose, R., and Isham, C. (1986). *Quantum Concepts in Space and Time*, Oxford University Press, Oxford, England.

Pitowski, I. (1984). A phase-space model for quantum mechanics in which all operators commute, in *Fundamental Problems in Quantum Mechanics Conference Proceedings*, CUNY, Albany, New York.

Prugovečki, E. (1984). *Stochastic Quantum Mechanics and Quantum Spacetime*, Reidel, Dordrecht, Holland.

Ryder, L. (1985). *Quantum Field Theory*, Cambridge University Press, Cambridge, England.