PROBABILITY IN THE FORMALISM OF QUANTUM MECHANICS ON PHASE SPACE

FRANKLIN E. SCHROECK, JR.

Abstract. The methods of Born and Einstein are used to obtain the probability density in the formalism of quantum mechanics on phase space. The resulting probability leads to a contextual measurement scheme. We give ramifications for paradoxes in standard quantum mechanics.

1. Introduction

In 1926, Max Born [1][2] analyzed the wave function, ψ, that was an (improper) eigenfunction for a Schroedinger equation with various potentials. In modern parlance, he came up with a candidate for $(\rho, \mathbf{j})$, with $\frac{\partial}{\partial t}\rho + \nabla \cdot \mathbf{j} = W\rho + \nabla \cdot \mathbf{j} = 0$ where

$$\rho(q) = |\psi(q)|^2, \tag{1.1}$$

$$\mathbf{j}(q) = c[\psi^*(q)\nabla\psi(q) - (\nabla\psi^*)(q)\psi(q)], \tag{1.2}$$

where c is a constant, and W equals an (improper) eigenvalue of the Hamiltonian. Note that $\nabla \cdot \mathbf{j} = c[\psi^* \triangle\psi - (\triangle\psi^*)\psi] = c[\psi^*(\triangle + V)\psi - ((\triangle + V)\psi^*)\,\psi]$; so that one may put in any potential one wishes as long as it is just a function of the variables $\{q_1, q_2, q_3\}$ of ψ. Since ∇ was the momentum operator up to a constant, one had a conserved current if $\rho(q)$ was the probability density of being at q. (However, the actual identification was made by the referee.) Thus $\rho(q)$ as in (1.1) was taken to be the probability density for standard non-relativistic quantum mechanics.

But this was non-relativistic and dealt with only eigenvalues of the Hamiltonian. The same process would not work for even special relativity. In 1988, a paper was published [3] in which the conserved current was given for cases of both Galilean and special relativity. We shall follow Born to obtain the form for the probability there. But this was predicated on the fact that measurement of ψ was to be described contextually by use of transitions to a second quantum mechanical particle described by the wave function η. We shall take that formalism and describe the conserved current equation for the symmetry group being the Heisenberg, Galilei or Poincaré groups, for massive, spinless (and spining) particles and a general Hamiltonian. This will necessarily bring us to the formalism of quantum mechanics on phase space. [4] Then we shall take these results, generalize them, follow Einstein in making a general basis for phase space(s), and obtain the probability density by means of the quantum mechanical transition probabilities. Finally, we shall see what the ramifications of this contextuality is in terms of the paradoxes of standard quantum mechanics.

Date: June 24, 2011.
1991 *Mathematics Subject Classification.* Primary 81P16, 81Q35; Secondary 81P13, 81P15.
Key words and phrases. quantum probability, phase space, contextuality.

2. The Phase Space Formalism

We are given a group (of symmetries), G, and look for the phase spaces on it so that the relevant equations, etc., on it are G-invariant. (We must have G a locally compact Lie group with a finite dimensional Lie algebra for this to work. One example is that G equals the Galilei group, but we have many other groups in mind as well.) Now, by a phase space we mean a space on which we can obtain a Poisson bracket, apply Hamilton's principle, have a G-invariant measure μ, etc. Such objects are also refered to as symplectic spaces in mathematics. In [4] we have given a complete description of such spaces and they are almost all of the form of G/H where H is one of the subgroups of G of a certain class that we need not specify here. They also may be disjoint unions of such symplectic spaces, but we shall presume there is only one such subgroup H for this paper. We emphasize that the phase spaces we obtain are just those we usually think about!

We shall use the following in the sequel: Let μ be the G-invariant measure on G/H, which exists because G/H is a phase space. On G/H, there is a natural left action of G given by $g \circ (g_1 \circ H) = (g \circ g_1) \circ H$ for each $g, g_1 \in G$. Henceforth we shall drop the $\circ$ from this notation. Also, we let $\sigma : G/H \to G$ be a Borel cross section.

Given G/H, we now form $L^2_\mu(G/H)$, which has a natural representation V of G on it defined by

$$[V(g)\Psi](\boldsymbol{x}) = \Psi(g^{-1}\boldsymbol{x}),\ \Psi \in L^2_\mu(G/H),\ g \in G,\ \text{and } \boldsymbol{x} \in G/H.$$

We may also make projective representations of G from this, but again we shall not do so in this paper. The operators $V(g)$ are unitary. Furthermore, we have the operators, $A(f)$ for f a real valued Borel function of the phase space variables, defined on $L^2_\mu(G/H)$ by

$$[A(f)\Psi](\boldsymbol{x}) = f(\boldsymbol{x})\Psi(\boldsymbol{x}).$$

The $A(f)$ (i) are a commuting set, and (ii) transform covariantly under the action of the $V(g)$. (See [4] for proofs of all of this as well as for the following.)

In any Hilbert space $\mathcal{H}$, we define the projection of φ onto ψ by

$$P_\psi \varphi = \ <\psi, \varphi> \psi,\ \psi, \varphi \in \mathcal{H},\ \ \| \psi \| = 1. \tag{2.1}$$

$\{P_\psi \mid \psi \in \mathcal{H},\ \| \psi \| = 1\}$ is termed the set of vector states in quantum mechanics. "pure state" and "vector state" coincide in the absence of super selection rules (which we will assume here). A general mixed state, ρ, is a trace class, positive operator of trace 1. By a result of von Neumann, there is an orthonormal basis $\{\phi_n\}$ of $\mathcal{H}$ and mixing probabilities $\{\alpha_n \in \mathbb{R}_{\geq 0} \mid \sum \alpha_n = 1\}$ such that $\rho = \sum_n \alpha_n P_{\phi_n}$. Consistent with the notation in quantum mechanics, we will take

$$\{Tr(P_\psi P_\varphi) = \ |<\psi, \varphi>|^2 \text{ such that } \psi, \varphi \in \mathcal{H},\ \ \| \psi \| = 1 = \| \varphi \|\} \tag{2.2}$$

to be the set of transition probabilities from pure state P_ψ to pure state P_φ. Similarly for mixed states by taking convex combinations of the P_ψ and P_φ for the ψ and φ extended to orthonormal bases. That these transition probabilities are indeed probabilities is immediate, from which we will obtain the quantum probabilities in phase space.

Theorem 1. *The transition probabilities are probabilities on the set of states in* $L^2_\mu(G/H)$. [5]

So we have that for any density operator ρ on $L^2_\mu(G/H)$,

$$\rho = \sum_j r_j P_{\Psi_j} = \sum_j r_j \mid \Psi_j >< \Psi_j \mid$$

for some orthonormal basis $\{\Psi_j\}$ of $L^2_\mu(G/H)$ and $\{r_j \mid j \text{ is countable}\}$ are mixing coefficients such that $r_j \geq 0$, $\sum_j r_j = 1$. Then we have

$$\begin{aligned} < \quad \Phi, \rho\Theta > \quad &= \sum_j r_j < \Phi \mid \Psi_j >< \Psi_j \mid \Theta > \\ = \quad & \sum_j r_j \int_{G/H} \overline{\Phi(\boldsymbol{y})}\Psi_j(\boldsymbol{y})d\mu(\boldsymbol{y}) \int_{G/H} \overline{\Psi_j(\boldsymbol{x})}\Theta(\boldsymbol{x})d\mu(\boldsymbol{x}) \\ = \quad & \int_{G/H} d\mu(\boldsymbol{y}) \int_{G/H} d\mu(\boldsymbol{x})\overline{\Phi(\boldsymbol{y})}\rho(\boldsymbol{y},\boldsymbol{x})\Theta(\boldsymbol{x}), \end{aligned}$$

defining the kernel $\rho(\boldsymbol{y},\boldsymbol{x})$. Then we have

$$\begin{aligned} Tr(A(f)\rho) &= \sum_j r_j Tr(A(f) \mid \Psi_j >< \Psi_j \mid) \\ &= \sum_j r_j < \Psi_j \mid A(f)\Psi_j > \\ &= \sum_j r_j \int_{G/H} d\mu(\boldsymbol{x}) f(\boldsymbol{x})\overline{\Psi_j(\boldsymbol{x})}\Psi_j(\boldsymbol{x}) \\ &= \int_{G/H} d\mu(\boldsymbol{x}) f(\boldsymbol{x})\rho(\boldsymbol{x},\boldsymbol{x}) \end{aligned}$$

which is equal to the classical expectation value

$$Exp(f;\rho_{CM}) = \int_{G/H} d\mu(\boldsymbol{x}) f(\boldsymbol{x})\rho_{CM}(\boldsymbol{x})$$

iff

$$\rho_{CM}(\boldsymbol{x}) \mapsto \rho(\boldsymbol{x},\boldsymbol{x}) \text{ for } a.e.\boldsymbol{x} \in G/H.$$

We recall that ρ_{CM} is defined as a Kolmogorov probability in classical statistical mechanics.

But the diagonal values of the kernel of ρ do not completely determine ρ. Presuming that the functions f, and thus the operators $A(f)$, are all that we have for determining any state ρ, we may then define the equivalence relation

$$\rho \approx \rho' \text{ iff } Tr(A(f)\rho) = Tr(A(f)\rho').$$

Thus

$$\rho \approx \rho' \text{ iff } \rho(\boldsymbol{x},\boldsymbol{x}) = \rho'(\boldsymbol{x},\boldsymbol{x}) \text{ for } a.e.\boldsymbol{x} \in G/H.$$

On the other hand, we see that setting $f = \chi_\Delta$ for any $\Delta \in Borel(G/H)$, χ_Δ the characteristic function, gives us a bijection from the Borel sets Δ to $Tr(A(\chi_\Delta)\rho)$ for all ρ states in $L^2_\mu(G/H)$. Thus we have another proof that, for $\rho(\boldsymbol{x},\boldsymbol{x})$, we have here a Kolmogorov probability and an effect valued measure in the notation of Beltrametti and Bugajski. [6]

Now, a quantum mechanical particle is given by an irreducible representation of G, and in particular, a separable Hilbert space, $\mathcal{H}$, which houses an irreducible representation of G. The following displays the way that these irreducible representations appear in physics.

Let $\mathcal{H}$ be a quantum mechanical separable Hilbert space over the manifold X where X is a space with an operation of G on the left, and with U being a representation of G on $\mathcal{H}$. Thus, for $\psi \in \mathcal{H}$ and $\boldsymbol{x} \in X$, $[U(g)\psi](\boldsymbol{x}) = \psi(g^{-1}\boldsymbol{x})$. In particular, for $\eta \in \mathcal{H}$, $\| \eta \|= 1$, and $\boldsymbol{x} \in X$, define

$$[W^{\eta}\psi](\boldsymbol{x}) = \; < U(\sigma(\boldsymbol{x}))\eta, \psi > .$$

An amazing fact is that every irreducible representation space $\mathcal{H}$ has such vectors, η, in it so that W^{η} is an isometry onto a closed subspace $P^{\eta}L^2(G/H)$ of $L^2(G/H)$. The conditions on η for this to happen are called "the α-admissibility conditions"; these conditions are easily achieved. Furthermore, the W^{η} intertwine the U's and V's. Define operators $A^{\eta}(f)$ on $\mathcal{H}$ by

$$A^{\eta}(f) \equiv [W^{\eta}]^{-1} P^{\eta} A(f) W^{\eta}.$$

From this definition, we automatically have that the $A^{\eta}(f)$ are covariant under the action of U. We may think of the operators $A^{\eta}(f)$ as the quantization of the classical statistical observables f. Note: This is a case of "prime quantization" and does not suffer from the obstructions present in other forms of quantization.

Now, in general, the set, $\mathcal{O}$, of operators on $\mathcal{H}$ is informationally complete iff, for $\rho, \rho' \in$ states in $\mathcal{H}$,

$$\begin{aligned} Tr(A\rho) &= Tr(A\rho') \text{ for all } A \in \mathcal{O} \\ &\Longleftrightarrow \rho = \rho'. \end{aligned}$$

Using this definition, we obtain that the set of the $A^{\eta}(f)$ and the $A^{\eta}(f)A^{\eta}(h)$, f and $h \in$ the set of real valued, measurable functions, is informationally complete when

$$< U(\sigma(\boldsymbol{x}))\eta, \eta > \; \neq 0.$$

We may simply say this as "the set of the $A^{\eta}(f)A^{\eta}(h)$ is informationally complete" since $A^{\eta}(\chi_{G/H}) = 1$. By a theorem of P. Busch [13], the set $\{A^{\eta}(f)A^{\eta}(h) \mid f$ and h bounded, measurable, and real valued$\}$ is dense in the bounded operators on $\mathcal{H}$ in a topology given in terms of the $Tr(A\rho)$, $A \in \mathcal{O}$. Consequently, the set $\{Tr(A^{\eta}(f)A^{\eta}(h)\rho)\}$ determines ρ uniquely. Notice the difference between this and the condition for $\approx$ defined on $L^2_{\mu}(G/H)$.

By the way, this condition for informational completeness has as a result, that ρ is spread everywhere in G/H no matter what H is. In section 4 we shall see that this means that the set of $A^{\eta}(f)$ are maximally non-local operators.

Moreover, we have a mapping from Δ to $A^{\eta}(\chi_{\Delta}), \chi_{\Delta}$ the characteristic function for the Borel set Δ of G/H. Just as before, we obtain the result that the set of $Tr(A^{\eta}(\chi_{\Delta})\rho)$ is a Kolmogorov probability for all states ρ. However, because of the projection, P^{η}, appearing in the definition of $A^{\eta}(f)$, we no longer have a one-to-one mapping. It is an effect valued measure, none-the-less. [6]

There is an important corollary to this:

Theorem 2. *Let $\mathcal{H}$ and $A^{\eta}(f)$ be as before and let $\{A^{\eta}(f)A^{\eta}(h) \mid f$ and h bounded, measurable, and real valued, with $0 \leq f, h \leq 1\} \equiv \mathcal{F}$. By the informational completeness of $\mathcal{F}$, the values $\{Tr(A^{\eta}(f)A^{\eta}(h)\rho)\}$ for any state ρ are sufficient to uniquely characterize ρ. But $\mathcal{F}$ has no non-trivial projections in it.*

The proof of this is in [7].

This has four consequences which are fundamental because of the following:

(1) The projections on non-trivial closed subspaces of $\mathcal{H}$ are excluded from $\mathcal{F}$; on the basis of the values $Tr(A^\eta(f)A^\eta(h)\rho)$, it would be impossible to conclude that ρ was supported in any closed subspace of $\mathcal{H}$. None-the-less, $\mathcal{F}$ is informationally complete.

(2,3) As we shall see, η is a wave function for a particle describing the experimental apparatus used to measure ρ in a truely quantum mechanical measurement. Thus the scheme we have devised makes no use of Gleason's Theorem [8], or any collapse of ρ.

(4) The fact that non-trivial projections on intervals in the phase space are nowhere to be found in $\mathcal{F}$ makes the interpretation of the phase space problematical. Only the contextual (the η) interpretation of the phase space is available to us by means of the average values of the variables of $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s})$ in η This is of key importance in Section 6.

We shall use the following in the sequel: If we write $\boldsymbol{x} = (\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s})$, where $\boldsymbol{p}$ is the usual momentum variable, $\boldsymbol{q}$ is the usual configuration variable, and $\boldsymbol{s}$ includes any other variables we may have such as the spin variable, we have for $f(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s}) = p^j$, then $A^\eta(f) = P^j$, and for $f(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s}) = q^j$, then $A^\eta(f) = Q^j$. Here the P^j and Q^j are the operators for the jth component of the momentum and position. [4, page 514]

3. Born's Approach and Quantum Mechanics on Phase Space

We now may review the results of [3]: In the spin zero non-relativistic case it was proved that a conserved currant was given as follows: Let

$$\begin{aligned}\widehat{\rho}(\eta; \boldsymbol{q}, t) &\equiv \int_{\mathbb{R}^3} e^{iHt/\hbar} \mid \eta(\boldsymbol{p}, \boldsymbol{q}, 0) >< \eta(\boldsymbol{p}, \boldsymbol{q}, 0) \mid e^{-iHt/\hbar} d\boldsymbol{p}, \\ \widehat{\boldsymbol{j}}(\eta; \boldsymbol{q}, t) &\equiv \int_{\mathbb{R}^3} \frac{\boldsymbol{p}}{m} e^{iHt/\hbar} \mid \eta(\boldsymbol{p}, \boldsymbol{q}, 0) >< \eta(\boldsymbol{p}, \boldsymbol{q}, 0) \mid e^{-iHt/\hbar} d\boldsymbol{p},\end{aligned}$$

where $\eta(\boldsymbol{p}, \boldsymbol{q}, 0)$ is the value of η at time $= 0$ and $H = H_0 + V$ is the Hamiltonian, V depending on $\boldsymbol{q}$ only. We have

$$\frac{\partial \widehat{\rho}}{\partial t} + \nabla_{\boldsymbol{q}} \cdot \widehat{\boldsymbol{j}} = 0.$$

Then, the conserved current for ψ was given by

$$\rho(\psi; \boldsymbol{q}, t) = \; < \psi, \widehat{\rho}(\eta; \boldsymbol{q}, t)\psi >; \; \boldsymbol{j}(\psi; \boldsymbol{q}, t) = \; < \psi, \widehat{\boldsymbol{j}}(\eta; \boldsymbol{q}, t)\psi > .$$

That is, the probability was given by

$$\begin{aligned}\rho(\psi; \boldsymbol{q}, t) &= \; < \psi, \widehat{\rho}(\eta; \boldsymbol{q}, t)\psi > \\ &= \int_{\mathbb{R}^3} < \psi, e^{iHt/\hbar}\eta(\boldsymbol{p}, \boldsymbol{q}, 0) >< \eta(\boldsymbol{p}, \boldsymbol{q}, 0)e^{-iHt/\hbar}, \psi > d\boldsymbol{p} \\ &= \int_{\mathbb{R}^3} < \psi, U(\boldsymbol{p}, \boldsymbol{q}, t)\eta >< U(\boldsymbol{p}, \boldsymbol{q}, t)\eta, \psi > d\boldsymbol{p} \\ &= \int_{\mathbb{R}^3} [W^\eta\psi]^*(\boldsymbol{p}, \boldsymbol{q}, t)[W^\eta\psi](\boldsymbol{p}, \boldsymbol{q}, t) d\boldsymbol{p}, \qquad (3.1)\end{aligned}$$

where U is the representation of the Heisenberg group. Similarly for $\boldsymbol{j}(\psi; \boldsymbol{q}, t)$. We also note that we could extend this to any mixed state. The upshot of this is that we may define the probabilities in $\boldsymbol{q}$-space as in equation 3.1.

Next, define the projections $T^\eta(\boldsymbol{x})$ on any $\mathcal{H}$ by

$$T^\eta(\boldsymbol{x})\varphi = \ < U(\sigma(\boldsymbol{x}))\eta, \varphi > U(\sigma(\boldsymbol{x}))\eta,\ \varphi \in \mathcal{H}.$$

Then we may write (3.1) as

$$\rho(\psi; \boldsymbol{q}, t) = \int_{\mathbb{R}^3} < \psi, T^\eta(\boldsymbol{p}, \boldsymbol{q}, t)\psi > d\boldsymbol{p}, \tag{3.2}$$

and we may define the probabilities in $\boldsymbol{q}$-space as in equation (3.2).

Similarly in the case of the Galilean group, showing that we have a definition more general than for just the Heisenberg group! Moreover, in the relativistic case (Poincaré group) we have essentially the same situation, except we have p^μ instead of p_i, j^μ instead of j_i, $\mu \in \{0, 1, 2, 3\}$, $i \in \{1, 2, 3\}$, and we integrate over $V_m^+ = \{p^\mu \in \mathbb{R}^4 = \mathfrak{M} \mid p_0 > 0, p_\mu p^\mu = m^2c^2\}$. [Recall that $\mathfrak{M}$ is the Minkowski space $\mathbb{R}^4$ with the metric $diag(1, -1, -1, -1)$. The tangent space to $\mathfrak{M}$ is denoted $T\mathfrak{M}$ and the dual to $T\mathfrak{M}$ is denoted $T^*\mathfrak{M}$.] The case of relativistic massive particles with spin may be found in [9]. We obtain a conserved current given by

$$\rho(\psi; \boldsymbol{q}) = \int_{V_m^+} d\boldsymbol{p} \int_{\mathfrak{S}^2(p)} d\boldsymbol{s} [W^\eta\psi]^*(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s})[W^\eta\psi](\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s}), \tag{3.3}$$

$$\boldsymbol{j}(\psi; \boldsymbol{q}) = \int_{V_m^+} d\boldsymbol{p} \int_{\mathfrak{S}^2(p)} d\boldsymbol{s} \frac{\boldsymbol{p}}{m} [W^\eta\psi]^*(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s})[W^\eta\psi](\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s}), \tag{3.4}$$

for $\boldsymbol{p} \in V_m^+$ and $\boldsymbol{s} \in \mathfrak{S}^2(\boldsymbol{p}) = \{s^\mu \in T^*\mathfrak{M},\ \boldsymbol{p} \cdot \boldsymbol{s} = 0,\ \boldsymbol{s} \cdot \boldsymbol{s} = -S^2\}$, whenever there is a conserved current to be had. [There is no integral over spin when the spin is zero. Also, see [10] for relativistic particles without spin.] Again, (3.3) defines a probability density for ψ; so, we may take (3.3), (3.2), or (3.1) as the general expression for the probability density. Note that the integral over $\boldsymbol{s}$ has to be taken first; this is a reflection of the property that you do not have a "joint distribution" in the sense of A. Fine. [11] Otherwise, the spin and the momentum would commute which is contrary to the Lie algebra of the Poincaré group.

There is a similar result in signal analysis in which time and frequency form the coordinates of the phase space. See [12].

These results for (ρ, j) may also be put in the forms given by the expressions in terms of the T^η's, suitably generalized. Furthermore, they have a conserved current (and not a conserved quasi-probability current) defined for any group that is relevant, including the Poincaré case. It seems that there is an advantage to defining wave functions that are square integrable over phase space!

4. Einstein's Approach Applied to Quantum Mechanics on Phase Space

This section is (also) a review. We enter it for those readers not familiar with quantum mechanics on phase space.

There are many cases where there is no conserved current; for example, in the cases where momentum and position are not variables. There is a probability density, however. This will be a remedy for those cases, and in fact for all cases. Furthermore, this probability density is in terms of the $\boldsymbol{p}$'s, $\boldsymbol{q}$'s and $\boldsymbol{s}$'s; we may obtain the probability density over any subset of the $(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{s})$'s by marginality. In the case of the Heisenberg group, the probability density in terms of the momentum is obtained by means of the Fourier transform of $\psi(\boldsymbol{q})$; one has no possibility of using the Fourier transform for a more general group. In fact, for the Poincaré group, with zero mass particles there is no $\psi(\boldsymbol{q})$ by this route! Moreover, we will obtain a formulation in which we have a contextual interpretation.

The fiirst step is to obtain a basis for quantum mechanics over the phase space G/H. Let $\mathcal{H}$ denote the space on which quantum mechanics is carried out. Let U denote the representation of G on $\mathcal{H}$. We have from before that for any η which is α-admissible, $W^\eta : \mathcal{H} \hookrightarrow L^2_\mu(G/H)$ is an injection; we may thus consider $\mathcal{H}$ as being "over the phase space G/H."

An overcomplete basis for $\mathcal{H}$ is obtained by taking such a wave function, $\eta \in \mathcal{H}$, and transporting it to all points of the phase space by means of the $U(g)$. This is done in a manner similar to the rods and clocks argument of Einstein. (For example, take the η to describe one rods-clocks-mass spectronomer-Stern-Gerlach-screen device at "the origin" in the massive, spinning case, and then transport it everywhere with the group. [14],[9]) Thus we have an η and then a set of $U(g)\eta$ for all $g \in G$. Then we take any $\phi \in \mathcal{H}$ and compare it to $U(g)\eta$ by means of the transition probability $|< U(g)\eta, \phi >|^2$. In fact, we may employ the set of $|< U(\sigma(\boldsymbol{x}))\eta, \phi >|^2$ for all $\boldsymbol{x} \in G/H$ which has a degeneracy equal to $\dim(H)$ in size, as $\eta \in \mathcal{H} \hookrightarrow L^2(G/H)$. [This is very similar to what Einstein did with rods and clocks for Minkowski space instead of G/H, and there was no opportunity to put in an η, since classically you would sample a point as a point. We shall return to this.]

Now, in general we have the result:

$$A^\eta(f) = \int_{G/H} f(\boldsymbol{x}) T^\eta(\boldsymbol{x}) d\mu(\boldsymbol{x}).$$

With this and for ρ a state on the space $\mathcal{H}$, we may form the Hilbert space expectation functional $Tr(A^\eta(f)\rho)$. We have $\rho = \sum r_n T_{\phi_n}$ for $\{\phi_n\}$ some orthonormal basis for $\mathcal{H}$. (Note that we have used the convention that lower case Greek letters denote vectors in $\mathcal{H}$, while upper case Greek letters denote vectors in $L^2_\mu(G/H)$.) Consequently we have

$$\begin{aligned}
(4.1) \qquad Tr(A^\eta(f)\rho) &= \int_{G/H} f(\boldsymbol{x}) Tr(T^\eta(\boldsymbol{x})\rho) d\mu(\boldsymbol{x}) \\
&= \sum_n \int_{G/H} f(\boldsymbol{x}) Tr(T^\eta(\boldsymbol{x}) T_{\phi_n}) d\mu(\boldsymbol{x}) \\
(4.2) \qquad &= \sum_n \int_{G/H} f(\boldsymbol{x}) \ |< U(\sigma(\boldsymbol{x}))\eta, \phi_n >|^2 \ d\mu(\boldsymbol{x}),
\end{aligned}$$

of which (4.1) is an integral over f of the transition probability of $T^\eta(\boldsymbol{x})$ with respect to ρ, and (4.2) is a sum and integral over transition probabilities. Expression (4.1) also tells us that we may consider

$$\rho \mapsto Tr(T^\eta(\boldsymbol{x})\rho)$$

as the connection between the density operators, ρ, on $\mathcal{H}$ and the standard classical statistical states on G/H. Then, (4.1) expresses the equivalence of the Hilbert space, $\mathcal{H}$, expectation values and the standard classical statistical expectations.

With this, we have obtained an operational interpretation for how we measure. In particular, the expression (4.2) is experimentally interpretable when f is a Borel function between 0 and 1 describing a fuzzy set. Furthermore, we have an immediate interpretation of the η in the expressions: It represents the wave function by which we will determine approximately where ρ or ϕ_n is! For example, suppose we are looking at an experiment in which we have an electron that is captured "at" a certain site on a screen. Then, η is the wave function for the electron with expectation values at that site, and f is a fuzzy function describing the location and range of the momentum sensitivity of the screen, while ρ or ϕ_n is arbitrary. See [4] for more on this.

We conclude that the probability density for finding P_ψ (or ρ) at almost every $\boldsymbol{x} \in G/H$ using wave function η for the instrument is

$$Tr(T^\eta(\boldsymbol{x})P_\psi) = \ |< U(\sigma(\boldsymbol{x}))\eta, \psi >|^2 \ .$$

This formula may be generalized using convex combinations of the P_ψs and of the T^ηs without any difficulty. It is a generalization of (3.1), (3.2), and (3.3).

Also, we have that the condition on η for the A^η, or equivalently the T^η, to be informationally complete is now a condition for maximal non-locality.

We add that defining everything in terms of transition probabilities is very natural, since the symmetries of our system are defined as the transformations which leave the transition probabilities invariant.

5. The Quantum Expectation and Variance on Phase Space

Now that we have a quantum probability on phase space depending on which vector η we choose with which to experimentally measure, we will interpret the form that the quantum expectation and variance take. Moreover, we will compare with the definitions in classical statistical mechanics, in standard quantum mechanics, in $L^2_\mu(G/H)$, and in phase space quantum mechanics.

Classically, we have for $f =$ a classical observable (a Borel function that is real valued) and ρ_{CM} a classical state,

$$Exp_{\rho_{CM}}(f) \equiv \ < f >_\rho \ = \int_{G/H} f(\boldsymbol{x})\rho_{CM}(\boldsymbol{x})d\mu(\boldsymbol{x}), \tag{5.1}$$

$$Var_{\rho_{CM}} = \ < (f - \ < f >_\rho)^2 >_\rho \ = \ < f^2 >_\rho - \ < f >^2_\rho \ . \tag{5.2}$$

We see that these expectations and variances are not contextual, but rather ideal.

For the standard quantum mechanics, we have $A =$ a single self-adjoint operator on a Hilbert space, $\{E_\lambda\}$ the spectral measure for A, h almost any function, and ρ a density operator,

$$Exp_\rho(A) = Tr(A\rho) = \int_{\mathbb{R}} \lambda d(Tr(E_\lambda \rho)), \tag{5.3}$$

$$Exp_\rho(h(A)) = Tr(h(A)\rho) = \int_{\mathbb{R}} h(\lambda) d(Tr(E_\lambda \rho)), \tag{5.4}$$

$$Var_\rho(A) = Exp_\rho(A^2) - [Exp_\rho(A)]^2. \tag{5.5}$$

The measure here depends on the spectral projection for which operator you take. For multiple commuting operators, the results are similar but you have to integrate over $\mathbb{R}^n$, n the number of independent operators. But for multiple non-commuting operators, there is no expression available. None-the-less, these expressions are ideal rather than contextual.

To obtain a measure which arises from the geometry of the phase space, G/H, we take the measure μ which is G-invariant and work on $L^2_\mu(G/H)$. For each measurable function f which is a bounded classical statistical observable, we obtain the bounded operator $A(f) : L^2_\mu(G/H) \to L^2_\mu(G/H)$. Then for state ρ,

$$\rho = \sum_j r_j \mid \Psi_j >< \Psi_j \mid$$

for an orthonormal basis $\{\Psi_j\}$ of $L^2_\mu(G/H)$ and the r_j as mixing coefficients,

$$\begin{aligned} Exp_\rho(A(f)) &= Tr(A(f)\rho) = \sum_j r_j Tr(A(f) \mid \Psi_j >< \Psi_j \mid) \\ &= \sum_j r_j \int_{G/H} f(\boldsymbol{x}) \mid \Psi_j(\boldsymbol{x}) \mid^2 d\mu(\boldsymbol{x}), \end{aligned} \tag{5.6}$$

$$\begin{aligned} Var_\rho(A(f)) &= \sum_j r_j \int_{G/H} \mid f(\boldsymbol{x}) \mid^2 \mid \Psi_j(\boldsymbol{x}) \mid^2 d\mu(\boldsymbol{x}) \\ & \quad -[Exp_\rho(A(f))]^2. \end{aligned} \tag{5.7}$$

Also, we may take the f to be unbounded, work on domains, etc. (Recall that we have here a commuting set of the $A(f)$s.) The summary of this is that these expressions are not contextual, but has the invariant measure on the phase space.

When we pass to an irreducible representation space (the usual space of quantum mechanics), we will have both invariance of the measure and contextuality. Explicitely, since $A^\eta(f) \equiv [W^\eta]^{-1} P^\eta A(f) W^\eta$, we have

$$\begin{aligned} Exp_\rho(A^\eta(f)) &= Tr(A^\eta(f)\rho) \\ &= \int_{G/H} f(\boldsymbol{x}) Tr(T^\eta(\boldsymbol{x})\rho) d\mu(\boldsymbol{x}) \\ &= \int_{G/H} f(\boldsymbol{x}) < U(\sigma(\boldsymbol{x}))\eta, \rho U(\sigma(\boldsymbol{x}))\eta > d\mu(\boldsymbol{x}), \end{aligned} \tag{5.8}$$

$$Var_\rho(A^\eta(f)) = Tr([A^\eta(f)]^2\rho) - [Tr(A^\eta(f)\rho)]^2. \tag{5.9}$$

For the variance, we note that

$$[A^\eta(f)]^2 = [W^\eta]^{-1} P^\eta A(f) P^\eta A(f) W^\eta \neq A^\eta(f^2).$$

We may use P^η in the form [4, p 333, eqn (55)] or compute directly to obtain

$$\begin{aligned} & Tr([A^\eta(f)]^2\rho) \\ = & \int_{G/H} \int_{G/H} f(\boldsymbol{x}) f(\boldsymbol{y}) K^\eta(\boldsymbol{y};\boldsymbol{x}) < U(\sigma(\boldsymbol{x}))\eta, \rho U(\sigma(\boldsymbol{y}))\eta > d\mu(\boldsymbol{x}) d\mu(\boldsymbol{y}) \end{aligned} \tag{5.10}$$

where

$$K^\eta(\boldsymbol{y};\boldsymbol{x}) = \ < U(\sigma(\boldsymbol{y}))\eta, \rho U(\sigma(\boldsymbol{x}))\eta >,$$

which is a "reproducing kernel." [4, pp. 544-547] We notice that we have expressions (5.8) and (5.9)-(5.10) which are analogous to (5.6) and (5.7), except that they are spread out by the presence of η.

If we could take the limit as $U(\sigma(\boldsymbol{x}))\eta$ becomes sharp in all its variables of $< U(\sigma(\boldsymbol{x}))\eta, \rho U(\sigma(\boldsymbol{y}))\eta >$, then $K^\eta(\boldsymbol{y};\boldsymbol{x})$ would become a delta function and $K^\eta(\boldsymbol{y};\boldsymbol{x}) < U(\sigma(\boldsymbol{x}))\eta, \rho U(\sigma(\boldsymbol{y}))\eta >$ would "tend to $< \boldsymbol{x}, \rho\boldsymbol{x} >$," which is just (5.7). Unfortunately, one may not do this as this would then violate all the uncertainty relations that are present in $U(\sigma(\boldsymbol{x}))\eta$!!! That K^η is "almost a delta function" is a feature of most experiments, but not all. Thus, we have a contextual measurement scheme that is not "almost non-contextual" in many experiments and certainly many prominent ones.

We last note that we may put $Tr(A^{\eta}(f)\rho) = \int_{G/H} f(\boldsymbol{x})Tr(T^{\eta}(\boldsymbol{x})\rho)d\mu(\boldsymbol{x})$ in terms of the derivatives of the transition probabilities $Tr(T^{\eta}(\boldsymbol{x})\rho) \equiv F(\boldsymbol{x})$ for suitable fs and for η informationally complete. In particular, for f being one component X_k of the momentum or position, for

$$\rho = \sum_j r_j \mid \psi_j >< \psi_j \mid, \ \{\psi_j\} \text{ an orthonormal basis of } \mathcal{H},$$

and

$$U(\sigma(\boldsymbol{x})) = \exp(-\boldsymbol{x} \cdot \boldsymbol{X}),$$

we have

$$\begin{aligned}
\frac{\partial}{\partial x_k} F(\boldsymbol{x}) &= \sum_j r_j \frac{\partial}{\partial x_k} \mid < U(\sigma(\boldsymbol{x}))\eta, \psi_j > \mid^2 \\
&= \sum_j r_j \left\{ \begin{array}{c} < X_k U(\sigma(\boldsymbol{x}))\eta, \psi_j >< \psi_j, U(\sigma(\boldsymbol{x}))\eta > \\ + < U(\sigma(\boldsymbol{x}))\eta, \psi_j >< \psi_j, -X_k U(\sigma(\boldsymbol{x}))\eta > \end{array} \right\} \\
&= Tr(\{X_k\rho - \rho X_k\} \mid U(\sigma(\boldsymbol{x}))\eta >< U(\sigma(\boldsymbol{x}))\eta \mid) \\
&= Tr(\{X_k\rho - \rho X_k\} T^{\eta}(\boldsymbol{x})) \\
&= Tr([X_k, \rho] T^{\eta}(\boldsymbol{x})),
\end{aligned}$$

and

$$\frac{\partial^2}{\partial x_k \partial x_l} F(\boldsymbol{x}) = Tr([X_k, [X_l, \rho]] T^{\eta}(\boldsymbol{x})),$$

etc. By using informational completeness of the $\{T^{\eta}(\boldsymbol{x})\}$ when we have it, and obtaining these expressions for all $\boldsymbol{x}$, we obtain $[X_k, \rho]$ and $[X_k, [X_l, \rho]]$, etc. Thus on multiplying by $f(\boldsymbol{x})$ and integrating, we have a method of mapping from ρ to either certain commutators involving ρ, or to certain derivatives of $F(\boldsymbol{x})$.

6. On Conditional Probabilities

You may have noticed that we have not used any concept of "collapse of the wave function" here. We can handle the conditional probability as follows: You have a system in a vector state P_{ψ} (or a state ρ, but we won't treat that explicitely as it is just a mixture of these states). You let it evolve for some time t. Then the state is in a new state $P_{\psi'}$. We could measure any classical statistical function, f, by computing $Tr(A^{\eta}(f)P_{\psi'}) = < \psi', A^{\eta}(f)\psi' > \forall f$, and using the "informational completeness of" η. Notice that this entails an infinite set of measurements on $P_{\psi'}$ in principle! It also entails that you use the Hamiltonian that you obtain by making the measuring apparatuses you will use to get these expectation values for the different $Tr(A^{\eta}(f)P_{\psi'})$. This is very contextual! It also does not cause any collapse of the wave function since you have assumed you may reproduce an infinity of identical states P_{ψ} and hence $P_{\psi'}$. Now you have an infinite set of $P_{\psi'}$ for measuring one $A^{\eta}(f)$. Then you let the system evolve further and then at some further time $t + t'$, in which you have in mind measuring $A^{\eta}(f')$, or even $A^{\eta'}(f')$, and you get $P_{\psi"}$. You get $Tr(A^{\eta'}(f')P_{\psi"})$, which involves another infinite set of measurments. This is the conditional probability of measuring $A^{\eta'}(f')P_{\psi"}$ at $t + t'$ when you have measured $A^{\eta}(f)$ at t.

Let us give an example. You have an oven which gives you a beam of atoms which you pass through a collumator and then a suitable magnetic field letting the beam fall on some region of a screen. You measure some particle in that region with η as the wave function for the particle as captured on the screen. Then you take the entire system and instead of capturing the particles at the particular region, you place a hole at that region. Then you place another (Stern-Gerlach) device beyond the hole in the screen and obtain the distribution you get for the double Stern-Gerlach device.

This gives you a different idea of what conditional probability means. It certainly is more complicated than the usual "conditional probability" but it is closest to the experimental situation.

A different scheme is to measure $A^\eta(f)$ in $P_{\psi'}$ by means of $Tr(A^\eta(f)P_{\psi'})$. You obtain a distribution for this. Then take the state $(A^\eta(f))^{\frac{1}{2}}P_{\psi'}(A^\eta(f))^{\frac{1}{2}}/N$, where $N = Tr(A^\eta(f)P_{\psi'})$, as the state which you evolve to the time $t + t'$, measure, etc. This has the property of being contextual but is not quite the experimental situation. It is the "statistical collapse" scheme, which we have also used in some discussions.

7. Escaping Some of the Paradoxes of Quantum Mechanics

We would like to point out that many of the so called paradoxes of quantum mechanics are resolved by working in the formalism of quantum mechanics on phase space. For example, the many problems that arise when taking a "measurement at $\boldsymbol{q}$" which takes you out of the Hilbert space $\mathcal{H}$ for ψ by "$\psi \mapsto \mid \boldsymbol{q} >$." Or problems of "measurement at $\boldsymbol{q}$" leading to Giancarlo Ghirardi's non-linear reformulation of quantum mechanics. [15] Or problems with the interpretation of relativity which do not consider the points in space as being only contextually defined. [16] Or problems with the Wigner quasi-probability density not being a proper probability distribution. [17] Or problems with the energy not being bounded below and zitterbewegung for a single electron [18][19], leading to the the concept of the Dirac sea [20] - which means that there was no single particle electron theory by this route. Or the problems of local quantum field theory, of which there are many. Or $\cdots$.

We would like to deal explicitely with one of these paradoxes, namely the results of the Einstein-Podolski-Rosen experiment [21] as modified by Bohm [22], and experimentally performed by Aspect, Grangier and Roger [23] with others. This is important as it deals explicitely with probability! It also is one that is almost non-contextual.

One has the following set-up: An atom has electrons in excited states which drop to lower energy states by giving off two photons, "1" and "2." The resulting photons are in a combined state having a total spin zero and fly off in the directions toward the two detectors at causally separated locations D1 and D2. Because the atom's electron states are in $\mathcal{H}$, i.e., in states that are not eigenfunctions of position and/or momentum, they have state functions that are square integrable in both position and momentum. They give off photons that are also not eigenfunctions of position and momentum. The resulting wave functions for the photons thus have helicities that are spread out in helicity space. Only if they weren't spread out in momentum space could they possibly be eigenfunctions of the helicity projections. (Helicity is parallel to the direction of the momentum. Also, the fact that the photons are not in wave functions "$\exp\{i\boldsymbol{x}\cdot\boldsymbol{k}\}$" but are distributed in momentum space is well-documented. It is the reason there is a self-interaction of the photon with itself and also the reason that photons in lasers have wave functions that are distributed in momentum space.) Thus, the wave functions ψ_1 and ψ_2 are not eigenfunctions of the helicity projection operators $Ps_\pm$: Photons are nesessarily relativistic particles, and in view of what we have said so far, they have wave functions [24] that are of the form

$$\psi(\boldsymbol{q},\boldsymbol{p},\boldsymbol{s})=\phi(\boldsymbol{q},\boldsymbol{p})s(\boldsymbol{p})$$

where $s(\boldsymbol{p})$ is the helicity function, and ϕ is the momentum and configuation space function. These $(\boldsymbol{q},\boldsymbol{p},s)$ form cosets where s is a null vector $s(\boldsymbol{p})=\lambda\boldsymbol{p}$, $\boldsymbol{p}\in V_0^+$, and $\boldsymbol{q}=\boldsymbol{q}+\mathbb{R}\boldsymbol{p}$. From the first of these, there are just two values for the helicity at each $\boldsymbol{p}$. Therefore, we may write the variables $(\boldsymbol{q},\boldsymbol{p},s)$ as $(\boldsymbol{q},\boldsymbol{p},\pm)$. We also have the measure on the phase space as $d\mu(\boldsymbol{q}_j,\boldsymbol{p}_j,s_j)=d\mu(\boldsymbol{q}_j,\boldsymbol{p}_j)$ where we have surpressed the $\pm$.

Then we have the $\boldsymbol{q}$ dependence, which may also be described in the familiar terms of an E-field and a B-field perpendicular to the direction $\boldsymbol{p}$. But, we have the $\boldsymbol{p}$ varying! For the set up of the EPRB experiment, we have that $\boldsymbol{p}$ is almost in one direction at D1 and the opposite direction at D2. Almost, but not exactly. So, we have to use these wave functions to obtain the probabilities that govern the results. This is what we shall do next.

We have that the wave function for the two photon state is symmetric in the indices 1 and 2, and that the symmetry is in the total momentum-position-spin space. The following arguement reduces the choice of photons so that they are symmetric in the helicity space: The total helicity of the system is zero. Thus the total helicity is not ± 2, ruling out the two states in which the helicities add. The two remaining states may have either both the ϕs and $\boldsymbol{s}$s symmetric or both antisymmetric. Because of the way that the statistical data is is accumulated, only the spatially symmetric case contributes. Thus we have the helicities in a symmetric state as well. We also have that $Ps_+ + Ps_- = \mathbf{1}$ in helicity space for any direction of $+$. (This is the equivalent of $P_n=\frac{1}{2}(1\pm n\cdot\sigma)$ for particles with spin 1/2. In fact these projections might be replaced by

$$\begin{aligned}E(n,\lambda) &\equiv \frac{1}{2}(1+\lambda n\cdot\sigma)\\ &= \frac{1}{2}(1+\lambda)P_n+\frac{1}{2}(1-\lambda)P_{-n}\end{aligned}$$

for unsharp measurements of spin. Measuring these operators is tantamount to stating the measurement process in statistical terms. See [25][26].) This fact, whether for photons or spin 1/2 particles, gives rise to the prediction that the results, when measured at D1 say, is independent of the angular direction in which D1 is oriented.

The detectors have the property that they measure the helicity in approximate directions d1 and d2 that may be chosen independently and at the last second so that the causally separated nature is not violated. One obtains the results that the two detectors register the photons as being (almost) parallel, consistent with the square integrability of quantum mechanics and without any regard to the causal relationship. If we take the measurement as being $A^\eta(f_1)$ "at" D1 and $A^\eta(f_2)$ "at" D2, where the functions f_1 and f_2 describe D1 and D2 fuzzily, then the probability, *Prob*, of the measurement has the form $Tr([A^\eta(f_1)\otimes A^\eta(f_2)]P_{\mathfrak{S}\psi_1\otimes\psi_2})$, where $\mathfrak{S}$ is the symmetrizer, and where we have propagated the wave function so that it is "at" D1 and/or D2. The f_j have variables $(\boldsymbol{q}_j, \boldsymbol{p}_j)$ as well. As a reflection of the fuzziness of Dj, we have that $f_j(\boldsymbol{q}_j, \boldsymbol{p}_j) \geq 0$ a.e.$(\boldsymbol{q}_j, \boldsymbol{p}_j)$, and

$$\int_{G/H} f_j(\boldsymbol{q}_j, \boldsymbol{p}_j) d\mu(\boldsymbol{q}_j, \boldsymbol{p}_j) = 1.$$

Since, for any θ_1 and θ_2,

$$\mathfrak{S}(\theta_1 \otimes \theta_2) = \frac{1}{\sqrt{2}}(\theta_1 \otimes \theta_2 + \theta_2 \otimes \theta_1),$$

we now have,

$$Prob = (\mathfrak{S}(\psi_1 \otimes \psi_2), [A^\eta(f_1) \otimes A^\eta(f_2)]\mathfrak{S}(\psi_1 \otimes \psi_2)).$$

This term, using the wave functions in the form $\psi_j = \phi_j s_j$ and using the argument of the symmetric helicities, is of the form

$$\begin{aligned} & \text{Prob} \\ = \; & \int_{G/H \times G/H} d\mu(\boldsymbol{q}_1, \boldsymbol{p}_1) d\mu(\boldsymbol{q}_2, \boldsymbol{p}_2) f_1(\boldsymbol{q}_1, \boldsymbol{p}_1) f_2(\boldsymbol{q}_2, \boldsymbol{p}_2) \\ & \times(\mathfrak{S}(\phi_1, \phi_2)\mathfrak{S}(s_1, s_2), T^\eta(\boldsymbol{q}_1, \boldsymbol{p}_1) T^\eta(\boldsymbol{q}_2, \boldsymbol{p}_2)\mathfrak{S}(\phi_1, \phi_2)\mathfrak{S}(s_1, s_2)). \end{aligned}$$

The terms $\mathfrak{S}(s_1, s_2)$ are each equal to $\cos\theta_{12} = \widehat{\boldsymbol{p}_1} \cdot \widehat{\boldsymbol{p}_2}$. The remaining integrals over the $\boldsymbol{q}$s and $\boldsymbol{p}$s is where the form of η enters the game. If η is quite spread out in either the p or q space or both, then we obtain approximately zero for the overall value for *Prob*. Likewise, if η is fairly sharp in both, then we get appromimately $\cos^2\theta_{12}$ for the value. Edward Prugovečki has computed this in concept for the η being a function that describes a particle confined to approximately one elementary particle's average width. One obtains an answer of approximately $\cos^2\theta_{12}$, the difference being so small that the world would have to wait for the experimentalists to design an experiment sensitive enough. [27] It seems that the probability in standard quantum mechanics is a good approximation to the one in the quantum mechanics on phase space in this experiment.

We also have the result that for any unsharp measurement of spin in the EPR "experiment" (which has not been done yet), the Bell inequalities are satisfied for POVMs if enough unsharpness is introduced as in the discussion of the $E(n, \lambda)$. [25][26] It seems that we do not have "sufficient unsharpness" in the EPRB experiment for photons, presuming a similar result.

References

[1] Max Born, Zeits. Phys. 37 (1926), 863-867.
[2] Max Born, Zeits. Phys. 38 (1926), 803-826.
[3] S. Twareque Ali, J. A. Brooke, P. Busch, R. Gagnon, and F. E. Schroeck, Jr., Can. J. Phys. 66 (1988), 238-244.
[4] Franklin E. Schroeck, Jr., Quantum Mechanics on Phase Space, Kluwer Acad. Pubs., Dordrecht, 1996.
[5] A. S. Holevo, Stastical Structure of Quantum Theory, Lecture Notes in Physics, m. 67, Springer, 2001, Prop. 2.1.1, p 41.
[6] E. C. Beltrametti and S. Bugajski, J. Phys. A: Math. Gen. 28 (1995), 3329-3343.
[7] Franklin E. Schroeck, Jr., "The C* Axioms and the Phase Space Formalism of Quantum Mechanics," in Contributions in Mathematical Physics, a Tribute to Gerard G. Emch, S. Twareque Ali and Kalyan B. Sinha, eds., Hindustan Book Agency, 2007, pp. 197-212; and Int. J. Theor. Phys. 47 (2008), 175-184.
[8] A. M. Gleason, J. Mathematics and Mechanics 6 (1957), 885-893.

[9] J. A. Brooke and F. E. Schroeck, Jr., "Phase Space Representations of Relativistic Massive Spinning Particles," in development.

[10] E. Prugovečki, Stochastic Quantum Mechanics and Quantum Spacetime, D. Reidel Pub. Co., Dordrecht, 1984, section 2.8.

[11] A. Fine, J. Math. Phys. 23 (1982), 1306-1310.

[12] F. E. Schroeck, Jr., Int. J. Theor. Phys. 28 (1989), 247-262, eqns. (5)-(7).

[13] P. Busch, Int. J. Theor. Phys. 30 (1991), 1217-1227.

[14] J. A. Brooke and F. E. Schroeck, Int. J. Theor. Phys. 44 (2005), 1889-1904.

[15] Giancarlo Ghirardi, Erkenntnis 45 (1997), 349-365.

[16] Franklin E. Schroeck, Jr., J. Phys. A: Math. Theor. 42 (2009), 155301.

[17] R. L. Hudson, Rpt. on Math. Phys. 6 (1974), 249-252.

[18] P. A. M. Dirac, Proc. Roy. Soc. (London) 117A (1928), 610-624; 118A (1928), 351-361.

[19] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. 78 (1950), 29-36.

[20] P. A. M. Dirac, Roy. Soc. Proc. 126 (1930), 360-365.

[21] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. 47 (1935), 777-780.

[22] D. J. Bohm, Quantum Theory, Prentice-Hall, Englewood Cliffs, N. J., (1951), pp. 611-623.

[23] A. Aspect, P. Grangier and G. Roger, Phys. Rev. Lett. 47 (1981), 460-463; and 49 (1982), 91-94.

[24] J. A. Brooke and F. E. Schroeck, Jr., J. Math. Phys. 37 (1996), 5958-5986.

[25] P. Busch, in Symposium on the Foundations of Modern Physics, P. Lahti and P. Mittelstaedt, eds., World Scientific Pub. Co., Singapore, 1985, pp. 343-357.

[26] P. Busch, nato.tex; 15/01/2002; 0:13; pp. 167-185.

[27] E. Prugovečki, in Symposium on the Foundations of Modern Physics, P. Lahti and P. Mittelstaedt, eds., World Scientific Pub. Co., Singapore, 1985, pp. 525-539.

University of Denver, Florida Atlantic University
E-mail address: `fschroec@du.edu`