Study of Wigner's Theorem on Joint Probabilities

M. Mugur-Schächter

Reprinted from FOUNDATIONS OF PHYSICS

Vol. 9, Nos. 5/6, June 1979
Printed in Belgium

Reprinted from FOUNDATIONS OF PHYSICS Vol. 9, Nos. 5/6, June 1979
Printed in Belgium

Study of Wigner's Theorem on Joint Probabilities

M. Mugur-Schächter[1]

Received August 22, 1978

The exact bearing of an important theorem proved by Wigner is established. The study brings out the fact that marginal conditions as well as mean conditions of a form currently required in joint probability attempts are in fact inadequate for the determination of a relevant concept of a joint probability. New vistas are thereby opened up.

1. INTRODUCTION

Wigner has demonstrated[(1)] an important theorem concerning joint probabilities associated with the quantum mechanical state vectors. There seems to be a tendency to interpret this theorem as the expression of a final impossibility of defining for any state vector a nonnegative joint probability of a position variable and a momentum variable. It will be shown in this work that such an interpretation is erroneous.[2]

The analyses which will be performed in order to specify the exact bearing of Wigner's theorem will bring into evidence the fact that neither marginal conditions nor mean conditions of a certain apparently straightforward structure are in fact adequate for the determination of a relevant concept of a joint probability, even though up to now both these types of conditions have been required in joint probability attempts.

The criticisms leading to this conclusion, while they yield an improved insight into the problem of joint probabilities, raise at the same time the question of a satisfactory formulation of this fundamental problem. Thereby they endow Wigner's proof with an outstanding heuristic interest.

[1] Université de Reims, Laboratoire de Mécanique Quantique, Reims, France.

[2] The same conclusion was reached in a previous work,[(2)] but here we present a much improved version.

0015-9018/79/0600-0389$03.00/0

2. WIGNER'S THEOREM

We start by reproducing Wigner's work. This will be done in detail, in order to facilitate any eventual comparison.

2.1. The Demonstration

Given a one-system wave function $\psi(q)$ (in one-dimensional notation), Wigner studies a joint function $P(q, p)$ of the positional variable q and the momentum variable p, on which he imposes the following conditions:

(a) That it be a "Hermitian form of $\psi(q)$," i.e.,

$$P(q, p) = (\psi, M(q, p)\psi) \tag{1}$$

where M is a self-adjoint operator depending on q and p.

(b) That $P(q, p)$, if integrated over p, give the proper probabilities for the values of q, as

$$\int P(q, p)\, dp = |\psi(q)|^2 \tag{2a}$$

and, if integrated over q, give the proper probabilities for the momentum, as:

$$\int P(q, p)\, dq = (2\pi\hbar)^{-1} \left| \int \psi(q)\, e^{-ipq/\hbar}\, dq \right|^2 \tag{2b}$$

Condition (b) admits the somewhat milder substitute that $P(q, p)$ should give the proper expectation value for all operators that are sums of a function of p and a function of q, as

$$\iint P(q, p)(f(p) + g(q))\, dq\, dp = \left(\psi, \left(f\left(\frac{\hbar}{i}\frac{\partial}{\partial q}\right) + g(q)\right)\psi\right) \tag{2}$$

A third "very natural" condition on $P(q, p)$ would be that it is nonnegative for all values of q and p:

$$P(q, p) \geqslant 0 \tag{3}$$

But Wigner demonstrates that the conditions (a) and (b) are incompatible with (3). This is realized by showing that the assumption that a $P(q, p)$ satisfying all three conditions (a), (b), and (3) can be defined *for every* ψ leads to a contradiction.

The contradiction is obtained for wave functions $\psi(q)$ of a particular form, namely for ψ that are linear combinations $(a\psi_1 + b\psi_2)$ of any two fixed functions such that ψ_1 vanishes for all q for which ψ_2 is nonnull, and vice versa. Wigner starts with the following lemmas:

Lemma 1. If $\psi(q)$ vanishes in an interval I, and if $g(q)$ is zero outside this interval and nowhere negative therein, one has for the P corresponding to the $\psi(q)$ above

$$\int P(q, p)\, g(q)\, dq = 0 \tag{4}$$

for all p (except for a set of measure zero).

This follows trom (2) with $f = 0$: the integral of (4) with respect to p vanishes because the right side of (2) vanishes

$$\iint P(q, p)\, g(q)\, dp\, dq = (\psi, g(q)\, \psi) = 0 \tag{4a}$$

However, the integrand with respect to p, that is, the left side of (4), is nonnegative for the g postulated, as long as (3) holds for P. It follows then that the integrand with respect to p must vanish except for a set of p of measure zero. QED

Furthermore, (4) is valid for every function $g(q)$ that satisfies the conditions of Lemma 1. It can then be concluded in a similar way that:

Lemma 2. If $\psi(q)$ vanishes in an interval I, the corresponding $P(q,p)$ vanishes for all values of q in that interval (except for a set of measure zero).

Wigner's demonstration then continues as follows:

Let us consider two functions $\psi_1(q)$. and $\psi_2(q)$ which vanish outside of two nonoverlapping intervals I_1 and I_2, respectively. Because of (1), the distribution function $P_{ab}(q, p)$ which corresponds to $\psi = a\psi_1 + b\psi_2$ will have the form

$$P_{ab}(q, p) = | a |^2 P_1 + a^*bP_{12} + ab^*P_{21} + | b |^2 P_2 \tag{5}$$

Setting $b = 0$, we note that P_1 is the distribution function for ψ_1, and similarly, setting $a = 0$, P_2 is the distribution function for ψ_2. Let us consider (5) for the q outside the interval I_1. Since (according to Lemma 2) P_1 vanishes almost everywhere for such q, the distribution function (5) cannot be positive for all a and b unless both P_{12} and P_{21} vanish if q is outside I_1 (except for a set of measure zero in q and p). A similar conclusion can be drawn when q is outside I_2. Hence, we have instead of (5), almost everywhere,

$$P_{ab}(q, p) = | a |^2 P_1(q, p) + | b |^2 P_2(q, p) \tag{6}$$

This means that the distribution function P_{ab} is almost everywhere independent of the complex phase of a/b. But this is impossible if P_{ab} is to give the proper momentum distribution for $\psi = a\psi_1 + b\psi_2$, i.e., is to satisfy (2b). Indeed, let us denote the Fourier transforms of $\psi_1(q)$ and $\psi_2(q)$ by $\varphi_1(p)$ and $\varphi_2(p)$. Equation (2b) then reads

$$|a|^2 \int P_1(q, p)\, dq + |b|^2 \int P_2(q, p)\, dq$$
$$= |a|^2 |\phi_1(p)|^2 + |b|^2 |\varphi_2(p)|^2 + 2 \operatorname{Re} ab^* \varphi_1(p)\, \varphi_2{}^*(p) \qquad (7)$$

Since this must be valid for all a and b, it requires identically in p

$$\varphi_1(p)\varphi_2{}^*(p) = 0 \qquad (7a)$$

But this is impossible, since $\varphi_1(p)$ and $\varphi_2(p)$, being Fourier transforms of functions restricted to finite intervals, are analytic functions (in fact, entire functions) of their arguments, and cannot vanish over any finite interval.

2.2. The Conclusion

Wigner formulates the result of his demonstration in the following terms (Ref. 1, p. 28): "no nonnegative distribution function can fulfil both postulates (a) and (b)."

3. STUDY OF THE THEOREM

We shall first analyze the proof, and afterwards, on the basis thus acquired, we shall examine the conclusion.

3.1. Analysis of the Proof

Framework of the Proof. The framework consists of the postulates: (a) [Hermitian forms defined by (1)], (b) [the two marginal conditions (2a), (2b) and the mean condition (2) for any ψ] and (3) (the nonnegativity condition). The assumptions of nonnegativity and of Hermiticity are entailed by the significance of a probability required for the distribution $P(q, p)$, hence they cannot be dropped without disintegrating the very problem chosen for examination, which consists precisely in the possibility of a *probability* distribution $P(q, p)$. Thus only the definition (1) and/or the postulate (b) are a priori questionable. We shall examine them successively.

Examination of the Definition (1). Definition (1) is not the most general one conceivable. The distribution operator M is required self-adjoint and dependent exclusively on q and p. The second requirement entails for M

independence of ψ, and this entails $P(q, p)$ as a *sesquilnear* form of ψ. Now the functional $P(q, p)$ is researched such as to accept the significance of a probability. Then the concept of a probability requires by its definition the reality of $P(q, p)$, so that $P(q, p)$ must be indeed a Hermitian form of ψ: the condition that M be self-adjoint cannot be dropped. But the independence of M on ψ is not imposed via the probabilistic significance desired for $P(q, p)$, so that in the examined context it is an arbitrary a priori restriction. We shall now show that:

Proposition 1. In the absence of the arbitrary restriction to a sesquilinear form for $P(q, p)$, Wigner's demonstration cannot be realized.

Proof. Instead of (1) we start out with the most general definition a priori conceivable for a joint probability distribution of q and p, namely

$$P(q, p) = (\psi, M(q, p, \psi)\psi) \tag{1'}$$

where the distribution operator $M(q, p, \psi)$ is self-adjoint and depends on q, p, and ψ. All the other assumptions introduced by Wigner are left unchanged. We introduce the notations: ψ_{ab} is a state vector $a\psi_1 + b\psi_2$ where the supports of ψ_1 and ψ_2 are disjoint; P'_{ab}, P'_1, and P'_2 are respectively the distributions obtained for ψ_{ab}, ψ_1, and ψ_2 by use of Definition (1'); P'_{12} and P'_{21} are respectively the analogs of P_{12} and P_{21} from (5) obtained by use of (1'). With these notations the expression of the joint distribution for ψ_{ab} yielded by Definition (1') is

$$P'_{ab}(q, p) = |a|^2 (\psi_1, M(q, p, \psi_{ab})\psi_1) + a^*bP'_{12} + ab^*P'_{21} + |b|^2 (\psi_2, M(q, p, \psi_{ab})\psi_2) \tag{5'}$$

In Wigner's expression (5), the factor of $|a|^2$ in the first term and the factor of $|b|^2$ in the last term identify respectively with the distribution P_1 yielded for ψ_1 by Definition (1) and with the distribution P_2 yielded for ψ_2 by Definition (1). The sequel of Wigner's proof is directly founded on this fact and on Lemma 2, as can be verified by inspection. But this fact is *not* reproduced in expression (5'). Now this is so precisely because of the dependence on ψ of the distribution operator M from (1'), which introduces ψ_{ab} in the argument of M, instead of, respectively, ψ_1 in the factor of $|a|^2$ and ψ_2 in the factor of $|b|^2$. For this reason—even though Lemma 2 continues to hold in the assumed context—Wigner's proof can no longer be reproduced with the nonsesquilinear definition (1'). QED

At this stage the following question naturally arises: are nonsesquilinear joint distributions compatible with both marginal conditions (2a) and (2b) possible?

The answer is positive, as a well-known example suffices to show: the "trivial" or "correlation-free" distribution $|\psi(q)|^2 \, \varphi(p)|^2$ [where $\varphi(p)$ is the Fourier transform of $\psi(q)$] is a nonnegative Hermitian and nonsesquilinear form of ψ defined for any ψ and fulfils both marginal conditions. Therefore it can be concluded that—in consequence of the restriction to distributions sesquilinear in ψ introduced by the definition (1)—Wigner's theorem has no bearing on a nonvoid class of joint probabilities a priori possible. On mathematical grounds (considerations of continuity) it seems probable that this class is not reduced to the trivial distribution alone. It cannot be decided whether this class does or does not contain "interesting" members, as long as the structure of *all* the conditions to be imposed upon a joint probability (time evolution, mean conditions, correspondence rules between functions and operators, etc., *and the marginal conditions themselves*) has not yet been thoroughly defined and studied *as an organic whole*. The attempts made up to now in this direction are not numerous and, as far as we know, none of them is both exhaustive and guided by an explicit and coherent system of physical criteria for the choice of the mathematical conditions. A very striking illustration will be found just below, where it will be shown that in fact neither the marginal conditions nor mean conditions of the form (2) are adequate for the determination of a relevant concept of a joint probability. Obviously this will relegate to a second place the above conclusion that the definition (1) is devoid of maximal generality.

Examination of Postulate (b). The mathematical conditions which a joint probability function $P_\psi(q, p)$ can satisfy are not independent of the physical significance assigned to the symbols q and p from the argument of P_ψ. In turn the significance assignable to the symbols q and p is subjected to the obvious criterion of relevance to the aim of the attempts at defining a joint probability concept. As long as the semantical criteria convenient for a relevant joint probability concept are not taken into account in great detail and most carefully throughout the process of determination of the concept, there is no hope whatever for bringing forth a concept endowed with descriptive usefulness. In any research of a mathematical description of a given domain of reality, the choice of an adequate structure of semantical contraints is a stage that has to be accomplished prior to any analysis of syntactic properties. Therefore the examination of the postulate (b) requires preliminarily—as a referential—the explicit specification of the adequate semantical content to be demanded for the function $P_\psi(q, p)$ and for the variables q and p.

Semantical Constraints on a Relevant Joint Probability Function. The prime source of all the attempts at defining a joint probability $P_\psi(q, p)$ associable to any quantum mechanical state vector lies in the reduction problem. This problem is well known: The quantum mechanical formalism yields only a statistical prediction concerning the outcome of one individual act of measurement, while this act brings forth a unique well-defined result, thereby "reducing" the predicted spectrum to a given certitude.

The main purpose of those who desire a hidden variables version of quantum mechanics is to obtain a "causal" solution for the reduction problem. Such a solution can be researched along the following lines: It is postulated that the studied system possesses, independent of observation, certain intrinsic properties statistically describable by a virtual distribution of values of an appropriate set of "hidden" parameters (hidden to quantum mechanics but not necessarily also to observation). Let us denote by μ' this set of hidden parameters, and by $[V(\mu_t')]$ the set of possible values for μ' at the time t. For one given system at any given time, only one set $[V(\mu_t')]$ of values is conceived to be realized for μ'. Each observable "quantity w of a system" is conceived as related with a corresponding function $w(\mu', \lambda)$ depending on the hidden parameters μ' and on parameters λ characterizing the measurement device (hence, implicitly, also on the time t). One individual act of measurement of w is conceived as a process of interaction between the system and a w-measurement device, which act induces into a deterministic evolution the unique but unknown value $w(\mu'_{t_0}, \lambda_{t_0})$ possessed by the function $w(\mu', \lambda)$ at the initial moment t_0 of this act of measurement. The unique observed value w_j of the observable quantity w brought forth by one act of measurement can thus be considered to emerge as an observable result of the system–device interaction, deterministically connected with the unique preexisting initial set of values $[V(\mu'_{t_0})]$ of the hidden parameters μ' of S, via the system–device interaction evolution. It has to be stressed, however, that the assumed existence of a deterministic connection between each observed w_j with a well-defined set $[V(\mu'_{t_0})]$ of initial values of μ' does not entail a one-to-one relation between the possible sets of initial values $[V(\mu'_{t_0})]$ of the system parameters μ' and the possible observable values w_j. Different sets $[V(\mu'_{t_0})]$ might conceivably lead—determiniscally—to the same w_j, in consequence of interactions with some particular apparatus configurations, while one given set $[V(\mu'_{t_0})]$ leads in general to different w_j via interactions with different λ_{t_0}. Thus, as soon as the role of the measurement device is taken into account with maximal generality, the assumption of a one-to-one revelation $[V(\mu'_{t_0})] \leftrightarrow w_j$ is obviously not characteristic of what is named a "causal" solution to the reduction problem (that is, a solution such that it can be conceived to relate each observed value w_j to only one *total* set of initial values $\{[V(\mu'_{t_0})] + \lambda_{t_0}\}$ for the system + measurement device. Bohr's views on the role of the measurement device were very profound. A hidden variables attempt which does not integrate these views carefully is doomed to insignificance.

Now, a position variable q and a momentum variable p constitute a particular set of hidden variables μ' assignable to a microsystem S, while a joint probability function $P_\psi(q, p)$ is a distribution function defined on the set of possible values $[V(q, p)]$ of this particular set of hidden system parameters. Therefore a joint probability function is endowed with descriptive usefulness only insofar as it is able to lead to a causal solution for the reduction problem. Joint probabilities that are a priori inadequate to this aim have to be discarded as semantically irrelevant.

Semantical Constraints on the Variables q and p. If the possible significance of q and p is considered, it is immediately obvious that the significance of "pure observables" (i.e., values of some observable entities for which the designations of "position" and "momentum" are decreed, but which are defined *exclusively* by the specification of some experimental circumstances involving the system, and where these entities emerge) cannot be relevant to the reduction problem. The conceptual organization which characterizes semantically the structure of the solution sought for this problem requires for the system-parameters q and p a *definibility* independent of observation, even if, in particular, the properties designated by q and p were conceived as being *also* observable. Therefore we discard a pure-observable significance.

The Beable Significance for q, p. Any property possessed by a system independent of observation has been called by Bell a beable property. We adopt this terminology. We shall now specify in detail what definitions we assign to the two important concepts of a beable position and of a beable momentum.

Beable Position. By definition we pose that this concept consists in the assumption of beable properties of the system which possess characteristics describable with the aid of the classical quantity position, i.e., which in any referential are, at any given time, nonnegligible only inside a finite and relatively small spatial domain. Such an assumption is equivalent to a minimal *model* of the object named "system." However—by its minimality—this model by no means entails the naive atomistic, multitudinist hypothesis concerning the structure of microreality. The finiteness and the smallness of the domain inside which the conceived beable position properties are "confined" are only *relative* to some specified (and modifiable) degree of approximation chosen for the description of these properties, while their "existence" is defined only with respect to some specified but arbitrary range of spatial dimensions characterizing the chosen scale of (imagined) contemplation (passive observation). The concept of the object called system itself, to which a beable position is assigned, emerges only relative to some choices of such approximations and of such a scale. Thus the notion that a beable position is possessed by what is named system has nothing absolute in it. In particular it leaves open the problems of separability of the systems and of locality of the phenomena in which they are involved.

Beable Momentum. It is not impossible to conceive a beable posisition which does *not* perform a continuous dynamics, but which merely consists of a discontinuous juxtaposition of an uninterrupted succession of locations possessed by some properties of the system (in the sense specified above) which vanish at one place while they are engendered at another. But this sort of a beable position might reproduce the "essentially probabilistic" features which a causal solution for the reduction problem attempts to remove. Such a beable significance for q in the argument of a joint probability might therefore yield a concept irrelevant to the reduction problem, so that we discard it. If then a beable position which does perform a continuous dynamics is assumed, ipso facto some definite continuous time variation of this beable position is equally assumed. This—by definition—is what we call a beable momentum.

The Beable Individual Kinematic Relation. Thus the assumption of a continuously moving beable position of a system is interdependent with the assumption of a beable momentum of this system. These two united assumptions are equivalent to the assumption of the descriptive relevance of a position variable q and a corresponding momentum variable p tied to one another by the individual kinematic relation (in one-dimensional terms)

$$p = K\, dq/dt \tag{8}$$

where K is a factor of proportionality playing the role of an inertial mass, but different, in general, from the inertial mass assigned to the considered system as a whole. This individual kinematic relation is a nontrivial and important implication of the concept of a continuously moving beable position, because it entails statistical correlations and these can be found to be either compatible or incompatible with a given condition of consistency with quantum mechanics, envisaged for a joint distribution of q and p.

We have thus specified for the variables q and p significances which are a necessary condition of relevance to the reduction problem. We are now prepared to examine furthermore the question of whether or not the mathematical conditions expressed by the postulate (b) can be imposed upon a joint probability $P_\psi(q, p)$ of variables q and p endowed with such significances, without a priori hindering thereby the relevance of this joint probability with respect to the reduction problem. The answer, as already announced, will be negative.

Rejection of the Requirement of Both Marginal Conditions (2a) *and* (2b). The marginal conditions (2a), $\int P_\psi(q, p)\, dp = |\psi(q)|^2$ and (2b), $\int P_\psi(q, p)\, dq = (2\pi\hbar)^{-1} | \int \psi(q) e^{-ipq/\hbar}\, dq |^2$ require the observability of both statistical distributions $P_\psi(q) = \int P(q, p) dp$ and $P_\psi(p)$. This does not entail that the individual values of q and p also have to be observable. Hence two complementary hypotheses are left open for investigation: either the beable variables q and p both have observable individual values, or both these variables do not have observable individual values.

Let us suppose first that both beables q and p do have observable individual values. Then it can be shown that:

Proposition 2. A joint probability distribution $P_\psi(q, p)$ of individually observable beables q and p cannot satisfy both marginal conditions (2a) and (2b) for any state vector at any time.

Proof. We produce an example: Consider the state vector

$$\psi(\mathbf{q}) = (1/\sqrt{2})\, \phi_{\mathbf{p}_1}(\mathbf{q}) + (1/\sqrt{2})\, \phi_{\mathbf{p}_2}(\mathbf{q})$$

where $\varphi_{\mathbf{p}_1}(\mathbf{q})$ and $\varphi_{\mathbf{p}_2}(\mathbf{q})$ are eigendifferentials of the quantum mechanical observable momentum (vector) corresponding respectively to the eigenvalues $\mathbf{p}_1$ and $\mathbf{p}_2$, the directions of which make an angle $\alpha \neq 0$, the norms being equal and nonnull, $|\mathbf{p}_1| = |\mathbf{p}_2| \neq 0$. Since this state requires a two-dimensional description, we refer it to two orthogonal axes ox, oz, the axis ox being chosen parallel to the bisectrix of α. The quantum mechanical position distribution $|\psi(x, z)|^2 = |\psi(\mathbf{q})|^2$ is then uniform along ox and periodic along oz; furthermore, this quantum mechanical distribution is *stationary*. We consider now a joint probability distribution $P_\psi(\mathbf{q}, \mathbf{p})$ associated with the chosen ψ and fulfilling both marginal conditions (2a) and (2b). By hypothesis $\mathbf{q}$ and $\mathbf{p}$ in the argument of $P_\psi(\mathbf{q}, \mathbf{p})$ are individually observable beables. Then the beable character of $\mathbf{q}, \mathbf{p}$ entails that at each given time each instantaneous individual value of the momentum variable $\mathbf{p}$ possesses a kinematic definition (8), $\mathbf{p} = K\, d\mathbf{q}/dt$, according to which it is generated by the time variation of a corresponding joint $\mathbf{q}$. Via this kinematic definition and the hypothesis of observability of the individual $\mathbf{p}$ the marginal condition (2b) for the *momentum* entails consequences for the time variations of the individual values of the position variable, and these in turn entail consequences for the statistical position distribution $P_\psi(\mathbf{q}) = \int P_\psi(\mathbf{q}, \mathbf{p})\, d\mathbf{p}$. Now for the chosen state vector the consequence for $P_\psi(\mathbf{q})$ of (8) and (2b) are not compatible with the stationarity of $P_\psi(\mathbf{q})$ required by the hypothesis of observability of the individual $\mathbf{q}$ and by the marginal condition (2a) for the position. Indeed (8) and (2b) entail nonnull z components for the time variations of the (observable) $\mathbf{q}$

$$dq_z/dt = p_z/K = \pm | p_z | \neq 0$$

This entails that, if at some initial time t_0, (2a) is realized, throughout the future $t > t_0$ the location with respect to oz of the maxima and minima of $P_\psi(\mathbf{q})$ keep reversing by a continuous observable process, with a time periodicity

$$dt = K\, dq_z/| p_z | = Ki/2 | p_z |$$

where i is the distance at t_0 between two successive maxima of $P_\psi(\mathbf{q})$. But such a reversal is not compatible with the stationarity of $P_\psi(\mathbf{q})$ required by (2a). QED

This example suffices for establishing Proposition 2. It shows that a joint probability $P_\psi(q, p)$ of individually observable beable values q and p which satisfies both marginal conditions for any ψ is a self-contradictory concept.

Let us now examine the hypothesis of individually inobservable beables q and p. It can be quite trivially asserted that the marginal condition (2a) [or (2b)] entails an arbitrarily a priori restricted statistical distribution of the values of an inobservable beable q (or p).

Indeed, let us consider first the momentum beable, because it seems less strange to suppose that its individual values are not observable. Making then this supposition for some given state of the studied system S, let us denote by p' the inobservable beable values of the momentum, in order to distinguish them from the observed values produced by the acts of momentum measurement performed on S. Even though the individual values p' are not observable, the marginal condition (2b) requires that the statistical distribution $P_\psi(p')$ shall coincide with the observable quantum mechanical distribution $(2\pi\hbar)^{-1} | \int \psi(q) e^{-ipq/\hbar}\, dq |^2$ of the values p (i.e., to each unknown value p' there corresponds an observed value p which arises statistically the same number of times). This, however, is a very strong a priori restriction on the relation permitted between observed values p and inobservable values p', which obviously eliminates *maximal* generality for the allowed types of system–device interactions. In this sense this restriction is arbitrarily restrictive so that it invalidates the approach.

An analogous argument holds for q also, even if it seems much less trivial.

Since a joint probability $P_\psi(q, p)$ of beables q and p that satisfies both marginal conditions for any ψ is either self-contradictory or a priori arbitrarily restricted, while—for relevance to the reduction problem—the beable significance for q and p has to be conserved, we are led to accept that at least one of the two marginal conditions *has* to be dropped when a relevant concept of a joint probability is sought. But then the imperative of maximal a priori generality commands that in fact we begin by requiring no marginal condition at all, that is, to start out with a hypothetical joint probability $P_\psi(q, p)$ where neither the position beable q nor the momentum beable p is a priori asserted to be observable. Therefore we finally conclude that:

The marginal conditions (2a) and (2b) are not relevant conditions in an investigation of the possibility of a joint probability associated to any state vector.

In other terms, only mean conditions can be relevantly imposed a priori upon a joint probability. This leads us now to the examination of Wigner's mean condition (2).

Rejection of the Mean Condition (2). There exist major reasons for which a mean condition of the form (2) does not correspond to a relevant concept of a joint probability. These reasons can first be expressed in connection with the quantum mechanical Schrödinger law of time evolution for ψ and with the relation between quantum mechanical operators and the hypothetical beable quantities: Obviously a joint probability function $P_\psi(q', p')$ has to be sought for *any* time t. Then $P_\psi(q', p')$ must satisfy to a certain time evolution law, and this law has to be compatible with the quantum mechanical evolution law of ψ. The time variation of the function $P_\psi(q', p')$ is (by mathematical definition)

$$\frac{\partial P_\psi(q', p')}{\partial t} = \frac{\partial P_\psi(q', p')}{\partial q'} \frac{dq'}{dt} + \frac{\partial P_\psi(q', p')}{\partial p'} \frac{dp'}{dt} \tag{9}$$

This general mathematical form can be specified so as to be integrated into the physical Newtonian framework. For this it is necessary and sufficient to assign to the time variation dq'/dt of the variable q' the physical significance of a momentum by posing the kinematic definition $dq'/dt = p'/K$ [(8)] and to equate dp'/dt to a certain "force"

$$dp'/dt = F_b \tag{10}$$

accordingly to the fundamental Newtonian dynamical postulate. Indeed with (8) and (10) the time evolution of $P_\psi(q', p')$ acquires the canonical Newtonian form

$$\frac{\partial P_\psi(q', p')}{\partial t} = \frac{p'}{K} \frac{\partial P_\psi(q', p')}{\partial q'} + F_b \frac{\partial P_\psi(q', p')}{\partial p'} \tag{9'}$$

The index b on the force F_b in (10) and (9′) stresses an essential feature, namely that, via (8) and (10), this force is related *directly* with the beable positional properties q' assigned to the studied system S (while it is *not* defined in direct connection with S considered as a whole). The force F_b can be conservative, dissipative, or a sum of a conservative term and of a dissipative term. Only in the first case is it derivable from a potential function, and then (10) acquires a Hamiltonian form. Now, it is well established that, given the Schrödinger evolution of ψ determined by some macroscopic potential $V(q)$, it is in general not possible to find a Newtonian evolution (9') for an attempted joint probability $P_\psi(q', p')$ if $F_b(q')$ in (10) is a priori required to be identical with the macroscopic force $F(q) = -\text{grad } V(q)$ acting on S as a whole: proofs of this impossibility are contained implicitly, but rather obviously, in the studies of the WKB approximation as well as in Feynman's path integral approach or in de Broglie's and Bohm's hidden variable attempts. Thus $F_b(q')$ in (10) has to be conceived as an unknown force which cannot be posed, but which has to be determined consistently with the Schrödinger evolution of $\psi(q)$ as a functional of $V(q)$ *via* $\psi(V(q))$. This functional, if its form were established in a satisfactory way, would probably yield the most specific mathematical descriptor of a nonnaive model of a microsystem,[3] whereas a brutal identification $F_b(q') \equiv F(q) = -\text{grad } V(q)$ for $q' = q$ would obviously be equivalent to a naive reduction of the whole microsystem S to its beable position-like property q' alone, which would be equivalent to the postulation of an atomistic, material-point model for S. But such a model was clearly known to be insufficient already from the time of de Broglie's Thesis and its confirmation by the Davisson–Germer experiment. There is a profound unity between de Broglie's model, which assigns to S a beable positional property q' incorporated into a physical wave like phenomenon which interacts with q', and the incompatibility between the Schrödinger time evolution of $\psi(q)$ and the time evolution of $P_\psi(q', p')$ if an identification $F_b(q') \equiv F(q) = -\text{grad } V(q)$ for $q' = q$ is made.

On the basis of these remarks it will now be easy to show that:

Proposition 3. Given a macroscopic classical dynamical quantity $f(q, p)$, a corresponding beable classical dynamical quantity is not always definable; if it happens, however, that such a beable quantity can be defined, then the function $f_b(q', p')$ which describes it has in general a form $f_b \neq f$ different from that of $f(q, p)$, so that f_b cannot be found in general by simply reversing the correspondence rule which led from $f(q, p)$ to the quantum mechanical operator $f(q, (\hbar/i)\, \partial/\partial q)$.

Proof. Again we produce an example. Consider the macroscopic dynamical quantity total energy $f(q, p) = H(q, p) = p^2/2m + V(q)$. Consider also one individual microsystem S. What can be said concerning a beable total energy of S? With our previous assumptions S possesses a beable position q' and a corresponding beable momentum $p' = K\,dq'/dt$. One can then form for S a kinetic energy $(p')^2/2K$ (where a priori K is not identical to the mass m of S as a whole). But in order to preserve for a joint probability $P_\psi(q', p')$ attempted for S the possibility of a time evolution compatible with that of $\psi_s(q)$, the force $F_b(q') = dp'/dt$, which—by Newtonian postulate—is equated to dp'/dt, has to be in general different from the macroscopic force $F(q) = -\text{grad}\, V(q)$, i.e., in general $F_b(q') \neq F(q)$ for $q' = q$. If, moreover, $F_b(q')$ is not conservative, then S simply does not possess a beable Hamiltonian, notwithstanding the fact that the time evolution of $\psi_s(q)$ is expressed by a Hamiltonian (operational) formalism. If, on the contrary, $F_b(q')$ also does derive from a certain potential, this potential $V_b(q')$ is in general a function of q' different from $V(q)$, $V_b(q') \neq V(q)$ for $q' = q$. Then S does possess a beable Hamiltonian $H_b = (p')^2/2K + V_b(q')$, but this Hamiltonian is a function $H_b(q', p') \neq H(q, p)$ of q', p' different in general from the function of q, p describing the macroscopic Hamiltonian $H(q, p) = p^2/2m + V(q)$ to which there corresponds the Hamiltonian operator

$$H\left(q, \frac{\hbar}{i}\frac{\partial}{\partial q}\right) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q)$$

which governs the evolution of $\psi_s(q)$. Replacement in $H(q, (\hbar/i)\,\partial/\partial q)$ of $(\hbar/i)\,\partial/\partial q$ by p, and of the multiplicative operator $V(q)$ by the function $V(q)$, yields back $H(q, p)$ but does not yield $H_b(q', p')$. QED

On the basis of Proposition 3 it is now obvious that marginal conditions of the form

$$\iint f(q, p)\, P_\psi(q, p)\, dq\, dp = \left\langle \psi \,\middle|\, f\left(q, \left(\frac{\hbar}{i}\frac{\partial}{\partial q}\right)\right) \psi \right\rangle \qquad (2')$$

where the *same* functional form f is posed in both members, is a priori devoid of general significance. This comes out most strikingly when a condition of the form (2′) is applied in particular to the quantity potential energy:

$$\iint V(q)\, P_\psi(q, p)\, dq\, dp = \langle \psi \mid V(q)\, \psi \rangle$$

One obtains then a mathematical *definition* of an atomistic material-point postulate on the structure of microreality, namely the identification of the potential assumed to yield the forces acting on the positional property q' assigned to the microsystem S, with the macroscopic potential yielding the forces that act on S as a whole. $P_\psi(q', p')$ and $\psi(q)$ cannot be purely algorithmically treated as if they were both fit for relevantly calculating means of *any* and the *same* functions. $P_\psi(q', p')$ can yield relevant means for beable values only, while $\psi(q)$ is relevant for calculating means of observed values of quantum mechanical operators only. Park and Margenau have explicity contested—on logical grounds—the relevance of mean conditions written with the macroscopic functions $f(q, p)$.[(4)] Proposition 3 only gives a more physical reason for this irrelevance.

Since the mean condition (2) is a particular case of (2′) (corresponding precisely to a classical macroscopic functional form of the *additive* type considered in the proof of Proposition 3), it is now obvious that such a mean condition is in general irrelevant.

Comment on the Conclusion of Wigner's Proof. Syntactically Wigner's proof is not contestable. Hence it is true indeed that a nonnegative joint distribution function satisfying both postulates (a) and (b) cannot exist. This conclusion, however, by no means excludes the possibility of associating a joint probability to the quantum mechanical formalism, since the postulate (a) is devoid of maximal generality, while the postulate (b) is inadequate for determining a relevant joint probability concept. The framework of the proof is not *semantically* worked out so as to ensure relevance to the problem of joint probabilities.

3.2. Outlook

The preceding analysis does not detract from the importance of Wigner's theorem; on the contrary, it endows it with an outstanding heuristic interest. The critical knowledge which has been gained concerning the structure and the bearing of Wigner's proof yields a strongly improved insight into the joint probability problem, replacing an irrelevant impossibility by a new orientation for further investigations on relevant possibilities. Wigner's proof ceases to appear as precluding a horizon, suggesting instead questions of a fundamental constructive importance. We shall deal with these questions in subsequent work. We shall show that—rather surprisingly—the joint probability problem has never yet been *formulated* in a semantically satisfactory way. In particular, this problem has always been dealt with at the outer level of probability *measures* alone, while the probability *spaces* where these measures are necessarily integrated have been left more or less in the dark. It is then explicable that up to now no solution has been found to the fundamental problem of joint probabilities. We shall be able to fill this lacuna with the help of a new abstract concept, that of "the tree of classificatory measure spaces of a random phenomenon," unifying in it a probabilistic approach with a logical and an informational expression of the semantical contents. The conjugated roles played by observation *and by time* will find a place carefully prepared for them inside this new concept. We shall sketch out a solution to the problem of joint probabilities (and more generally to the hidden variables problem), constructing thereby a framework where a description of microsystems deeper than quantum theory becomes feasible.

ACKNOWLEDGMENTS

I am profoundly indebted to Prof. Wigner for having accepted to discuss a primitive version of this work. I am also indebted to Prof. Shimony for an improving remark on my previous formulation.

REFERENCES

1. E. P. Wigner, in *Perspectives in Quantum Theory*, W. Yourgrau and A. van der Merwe, eds. (MIT Press, 1971).
2. M. Mugur-Schächter, in *Quantum Mechanics a Half Century Later* (Reidel, 1977).
3. M. Mugur-Schächter, D. Evrard, and F. Thieffine, *Phys. Rev. D* **6**, 3397 (1972).
4. J. Park and H. Margenau, *Int. J. Theor. Phys.* **1**, 211 (1968).

Printed by the St. Catherine Press Ltd., Tempelhof 37, Bruges, Belgium