International Journal of Theoretical Physics, Vol. 35, No. 11, 1996

Fundamentals of Fuzzy Probability Theory

Sławomir Bugajski[1]

Received June 11, 1996

The canonical classical extension of quantum mechanics studied recently by E. G. Beltrametti and S. Bugajski opens a new way toward generalizing the standard probability theory. The emerging fuzzy probability theory is able to give a full account of both classical and quantal probabilities, and—like the standard probability theory—could be of universal use, far outside the borders of physics. A specific feature of this hypothetical theory of probability is its mixed, classical–quanta character: classical as well as quantal random variables are described on an equal footing in a unified framework. Some new features of the fuzzy probability theory are shown on simple examples.

1. THE RATIONAL VERSUS THE INTUITIVE

The carefully prepared prospectus of this conference spotlights a confrontation between the conscious ordering of the physical and social world, symbolized by "Einstein," and the imaginary world of "Magritte." This deep opposition, which dates back to the ancients, has many faces; its epistemological aspect can be verbalized as the opposition between "rational" and "intuitive" ways of cognition. Needless to say, the former is the paradigm for science, whereas the latter is typical for everyday life as well as for the fine arts, literature, and the humanities. Rather than attempt to discuss this topic in great detail, I stress some characteristic features of these competing ways of acquiring knowledge.

Rationalism is based on a well-defined, though oversimplified, model of reality. The rational universe is categorical, fragmented, and deterministic: things can be objectively classified as belonging or not to a given category, the same cause in the same circumstances produces the same effect, etc. The basic features of the rational model are imitated by the laws of classical logic.

[1] Institute of Physics, University of Silesa, PL 40-007 Katowice, Poland.

0020-7748/96/1100-2229$09.50/0

The intuitive approach puts into question the axioms of the rationalist myth. The intuitive view of the world allows it to be fuzzy, holistic, and indeterministic: properties are usually subjective and often not sharp, free will is acausal, etc. No wonder that irrational cognition is also "illogical," as it does not respect the rules of classical logic. Note that, contrary to the rationalist ideal, there is no borderline between rational and intuitive cognitions. Moreover, it seems that the latter in a sense "contains" the former: a person who appreciates intuition can nevertheless make use of the rational rules of thinking.

It is ironic that rational science proves in a rational way that nature is "irrational." I will show below that the probabilistic background of quantum mechanics is formed by what I call fuzzy probability theory, which—contrary to the standard (Fréchet–Kolmogorov) probability theory—goes well beyond rational limitations, but nevertheless contains the standard, rationalism-based, theory. Perhaps fuzzy probability theory with its rational description of irrationality would be one of places where "Einstein" meets "Magritte."

2. QUANTUM MECHANICS SUGGESTS FUZZY PROBABILITY THEORY

Approaching the main topic of this essay, we observe first that the standard probability theory is designed according to the rules of rationalism. A standard random variable (a standard property) attaches a well-defined value to any elementary event, hence the standard probability theory is a theory of sharp (categorical) properties. A probability measure over the collection of possible elementary events merely describes our incomplete knowledge of the actual state of the universe of discourse. The incompleteness of the initial information leads naturally to an uncertainty about the outcome of a test of a random variable in such a state; this uncertainty is described again by a probability distribution over the set of possible values the random variable can take on. Thus the distribution of a standard random variable is still a probability distribution of a sharply defined, categorical property.

A different picture arises if we inspect the probabilistic structure of quantum mechanics. Let $\mathcal{H}$ denote a complex separable Hilbert space, the stage for all performances of quantum mechanics. The quantum mechanical pure states are represented by one-dimensional orthogonal projections on $\mathcal{H}$; the set of all of them will be denoted $\Omega_{\mathcal{H}}$ (the measurable structure of $\Omega_{\mathcal{H}}$ we need for defining probability measures on it is induced by the weak topology of the Banach space of trace-class operators on $\mathcal{H}$). Let $\hat{F}$ be a self-adjoint operator on $\mathcal{H}$; we assume that it represents an observable physical quantity ("is an observable" according to quantum mechanical usage). It is convenient to understand here the spectral resolution of $\hat{F}$ as a map $\hat{E}^F$ which to any (Borel-measurable) subset of the real line $\mathbf{R}^1$ attaches an appropriate projection operator on $\mathcal{H}$ (see, e.g., Reed and Simon, 1972). The map $\hat{E}^F$ (called the spectral measure associated with $\hat{F}$) is known to produce at any point $\omega \in \Omega_{\mathcal{H}}$ a probability measure $F(\omega)$ on $\mathbf{R}^1$ according to

$$F(\omega)(X) := \mathrm{Tr}(\hat{P}_\omega \hat{E}^F(X)) \tag{1}$$

where $X \in \mathcal{B}(\mathbf{R}^1)$ is a measurable subset of the real line, $\hat{E}^F(X)$ is the projection operator attached to X by the spectral resolution of $\hat{F}$, and $\hat{P}_\omega$ is the one-dimensional projection operator representing the pure state ω. It is well known that the measure $F(\omega)$ rarely is concentrated at a point (i.e., rarely equals the Dirac measure δ_λ for some $\lambda \in \mathbf{R}^1$, called then an eigenvalue of $\hat{F}$).

The commonly accepted statistical interpretation of quantum mechanics connects the measure $F(\omega)$ directly to experiments: it should describe the statistical scatter of values of the physical quantity represented by $\hat{F}$ which results from a measurement of this quantity on an ensemble of objects prepared to be in the state ω. The same meaning is commonly attached to a distribution of a random variable provided by standard probability theory. This suggests that quantum mechanical observables should be seen as quantum counterparts of standard random variables. Assuming now that quantum pure states correspond to elementary events, we obtain a scheme resembling standard probability theory. There is, however, a crucial difference: quantum mechanical observables, contrary to random variables of standard probability theory, generate at pure states probability distributions (nontrivial in general) rather than well-defined values.

In this way we come to a natural generalization of the standard concept of random variable:

(Generalized) random variables are measure-valued functions on the space of elementary events.

The standard random variables would fit this preliminary definition if only we agree to consider the sharp values they take on as Dirac measures. Let me introduce some notation: given a set Ω of elementary events, a random variable (in the generalized sense) on Ω having outcomes in an outcome space Ξ is represented by a function $F: \Omega \to M_1^+(\Xi)$, where $M_1^+(\Xi)$ denotes the convex set of probability measures on Ξ. Clearly, Ξ has to be equipped with a measurable structure, so we should in fact speak about a measurable space $(\Xi, \mathcal{B}(\Xi))$. The standard random variables are then identified as those ranging over the set $\{\delta_\xi | \xi \in \Xi\} \subset M_1^+(\Xi)$ of Dirac measures on Ω.

It is evident that a nonstandard random variable shows a kind of inherent uncertainty, as it attaches nontrivial probability distributions on Ξ to (some) elementary events. In a generic case this feature cannot be interpreted as a result of our incomplete knowledge of the actual situation (see below, Section

6). A theory incorporating such random variables describes a universe which does not fit the rational ideal—it is not categorical and not deterministic. (It seems to be holistic, too.) This is exactly the case of quantum mechanics, but it is not the case of standard probability theory. One can easily guess that the new concept of random variable put into the elaborate structure of the latter would force its fundamental reconstruction. The emerging theory-to-be is exactly what we call fuzzy probability theory. The remaining part of this essay is the first guided sightseeing tour of some accessible regions of this unexplored and exotic land.

3. FUZZY EVENTS

Let $(\Omega, \mathcal{B}(\Omega))$ denotes a measurable space, elements of Ω representing elementary events (in physics: pure states). Consider a random variable F: $\Omega \to M_1^+(\Xi)$. As $F(\omega)$ is a measure (at least trivial) on Ξ, the random variable defines a specific function K^F on the Cartesian product $\Omega \times \mathcal{B}(\Xi)$ taking values in the unit interval [0, 1]:

$$K^F(\omega, X) := F(\omega)(X) \tag{2}$$

If we fix $\omega \in \Omega$ and let X vary over the σ-algebra $\mathcal{B}(\Xi)$ of measurable subsets of Ξ, we recover the measure $F(\omega)$. If we fix $X \in \mathcal{B}(\Xi)$ and let ω vary over Ω, we get a real-valued function on elementary events which will be called the effect of F on X.[2] Thus $K^F(\omega, \cdot) = F(\omega)$, while the effect $K^F(\cdot, X)$ will be denoted $E^F(X)$.

Incomplete information about the actual state of affairs is to be described, as in the standard theory, by a probability measure on Ω. Random variables should be meaningful also in such cases; it is natural to assume that the random variable F: $\Omega \to M_1^+(\Xi)$ generates at a probability measure μ on Ω the distribution, denoted $A_F\mu$, on Ξ according to the formula

$$A_F\mu(X) := \int_\Omega F(\omega)(X)\ \mu(d\omega) = \int_\Omega K^F(\omega, X)\ \mu(d\omega) \tag{3}$$

where $X \in \mathcal{B}(\Xi)$. The integral is well defined only if the effect $E^F(X)$ is a measurable function on Ω. This is what we have to assume to obtain the complete definition of random variable. Thus:

A function F: $\Omega \to M_1^+(\Xi)$ is a random variable on Ω if and only if for any $X \in \mathcal{B}(\Xi)$ the numerical function $F(\omega)(X)$ on Ω is measurable.

[2]The term "effect" was apparently introduced by Ludwig (1954) in his operational approach to quantum mechanics. We borrow this term to stress that fuzzy probability theory is founded on some general ideas of this approach, although we do not use here its typical formalism of ordered Banach spaces.

This completion of the definition of random variable makes also more regular the function K^F introduced by formula (2). With the condition of measurability of effects of the random variable F, K^F becomes a (Markov) kernel from $(\Omega, \mathcal{B}(\Omega))$ to $(\Xi, \mathcal{B}(\Xi))$ as defined in the standard theory of stochastic processes; see, for instance, Bauer (1981).

Notice by the way that formula (3) extends any random variable F: $\Omega \to M_1^+(\Xi)$ to the affine function A_F: $M_1^+(\Omega) \to M_1^+(\Xi)$. On the other hand, some affine functions from $M_1^+(\Omega)$ to $M_1^+(\Xi)$ define, by restriction to the set of Dirac measures on Ω, fuzzy random variables on Ω. Measure-valued affine functions on convex sets are natural representants of physical observables (see Bell, 1964).

Coming back to effects, consider a standard random variable F: $\Omega \to M_p(\Xi)$. As $F(\omega)$ is a Dirac measure for every ω, the corresponding kernel K^F: $\Omega \times \mathcal{B}(\Xi) \to [0, 1]$ assumes only two values, 0 and 1. Hence any effect $E^F(X)$ of F is a 2-valued function on Ω too. The measurability of effects implies that these functions are characteristic functions of measurable subsets of Ω. This is a manifestation of the sharp (or categorical) nature of the standard random variable. If we now take F to be a general random variable, we come to see its effects $E^F(X)$ as membership functions of fuzzy subsets of Ω, with the values of the membership functions being naturally probabilities. This fuzziness is a characteristic property of the general notion of random variable and explains the name of the generalized probability theory we propose.

A random variable F: $\Omega \to M_p(\Xi)$ attaches, *via* the kernel $K^F : \Omega \times \mathcal{B}(\Xi) \to [0, 1]$ determined by it, an appropriate effect $E^F(X)$ to any measurable subset of Ξ. It can be demonstrated that the map E^F of $\mathcal{B}(\Xi)$ into effects is an effect-valued measure on Ξ, the typical construction of the mentioned operational quantum mechanics (see an up-to-date and thorough monograph of Busch, Grabowski, and Lahti, 1995). The effect-valued measure E^F, called the semispectral resolution of the random variable F, provides an alternative way of representing random variables. Any random variable can be so decomposed into effects and constructed out of them: effects are elementary constituents of random variables.

The subsets belonging to $\mathcal{B}(\Omega)$ are traditionally called events, we can say (with a slight abuse of language) that the random variables appearing in the standard probability theory can be represented as event-valued measures on their outcome-spaces. Hence our effects are to be seen as generalizing the standard notion of event, and we can call them general (or fuzzy) events.

Let us summarize the above remarks: The basic structure of the fuzzy probability theory consists of a measurable space $(\Omega, \mathcal{B}(\Omega))$, the space of elementary events, together with the set $\mathcal{F}(\Omega)$ of all (generalized, fuzzy) random variables on it. Any random variable F: $\Omega \to M_p(\Xi)$ defines the affine map $A_F : M_p(\Omega) \to M_p(\Xi)$, and decomposes into effects (fuzzy events).

4. FUZZY PROBABILITY THEORY GENERATES QUANTUM MECHANICS

Now it is rather apparent that the standard probability theory can be obtained from the fuzzy probability theory if only we forsake the nonstandard random variables and focus our attention on the sharp ones. It is less trivial that quantal theories emerge from the fuzzy probability theory in a similar way.

The procedure of obtaining the basic structure of the classical probability theory from the fuzzy one does not influence the basic convex set $M_p(\Omega)$. It is not so in a general case. If we have (or prefer) to restrict ourself to a narrower set, say $\mathcal{F}_q(\Omega)$, of random variables on Ω, $\mathcal{F}_q(\Omega) \subset \mathcal{F}(\Omega)$, it can happen that we lose the ability of discriminating between measures on Ω.[3] Assuming that $\mathcal{F}_q(\Omega)$ indeed does not separate elements of $M_1^+(\Omega)$, we get the equivalence relation on $M_1^+(\Omega)$:

$$\mu_1 \simeq \mu_2 \qquad \text{iff} \quad \int_\Omega E^F(X)\ \mu_1(d\omega) = \int_\Omega E^F(X)\ \mu_2(d\omega)$$

for all $F \in \mathcal{F}_q(\Omega)$ and all $X \in \mathcal{B}(\Xi)$, where Ξ is the outcome space of F. In concise form it reads

$$\mu_1 \simeq \mu_2 \qquad \text{iff} \quad A_F\mu_1 = A_F\mu_2 \tag{4}$$

for all $F \in \mathcal{F}_q(\Omega)$. Thus two measures are equivalent in this meaning if they are seen as identical by all random variables in which we are interested. The factor set $M_1^+(\Omega)/\simeq$ (i.e., the family of all equivalence classes under $\simeq$) inherits the convex structure from $M_1^+(\Omega)$; nevertheless it does not have to be a simplex any more. We will show below that in this manner we can generate (from an appropriate Ω) the convex set of states typical for quantum mechanics.

It is interesting to see what happens to random variables on Ω when the simplex $M_1^+(\Omega)$ is transformed into $M_1^+(\Omega)/\simeq$. Consider the natural projection (called the reduction map)

$$R: \quad M_1^+(\Omega) \to M_1^+(\Omega)/\simeq$$

which to any measure on Ω attaches the $\simeq$-equivalence class to which it belongs. A remarkable property of the reduction map R appears when the distinguished set $\mathcal{F}_q(\Omega)$ of random variables discriminates between all points

[3]The formal notion of coarsening (Busch and Quadt, 1993) gives a correct account of this effect.

of Ω, which seems actually to be the most interesting case. Assuming this separability property of $\mathcal{F}_q(\Omega)$, we find that the reduction map R embeds the set $\delta(\Omega) := \{\delta_\omega | \omega \in \Omega\}$ of all Dirac measures on $M_1^+(\Omega)$ into the set $\partial(M_1^+(\Omega)/\simeq)$ of extreme points of the factor set in a 1-to-1 way.

Owing to this, any random variable $F \in \mathcal{F}(\Omega)$ can be literally transported to the extreme boundary of $M_p(\Omega)/\simeq$, which defines a measure-valued function on (a subset of) $\partial(M_1^+(\Omega)/\simeq)$. In spite of the formal analogy, this function cannot be in general regarded as a counterpart of the original random variable F. The reason is that, even if it could be extended over $\partial(M_1^+(\Omega)/\simeq)$, the factor set $M_1^+(\Omega)/\simeq$ is not a simplex; hence the problem of the affine extension of functions defined on its extreme boundary [the Dirichlet problem; see Alfsen (1971)] becomes highly nontrivial.

The distinguished random variables are well behaved in this respect. It is evident that the affine map A_F determined by any $F \in \mathcal{F}_q(\Omega)$ identifies measures belonging to the same equivalence class. Hence A_F with $F \in \mathcal{F}_q(\Omega)$ factorizes into the composition of the reduction map R and an affine map, say $A_{q,F}$, of the factor set $M_1^+(\Omega)/\simeq$ into $M_1^+(\Xi)$,

$$A_F = A_{q,F} \circ R: \quad M_1^+(\Omega) \overset{R}{\rightarrow} M_1^+(\Omega)/\simeq \overset{A_{q,F}}{\rightarrow} M_1^+(\Xi)$$

The map $A_{q,F}$ is the unique affine extension of $F \in \mathcal{F}_q(\Omega)$ over the factor set $M_1^+(\Omega)/\simeq$.

Taking into account that, as was mentioned above, measure-valued affine functions on a convex set (of states) provide the natural tool for describing physical observables, we see that the distinguished random variables of $\mathcal{F}_q(\Omega)$ can be regarded as "observables" of the "physical model" based on the convex set $M_1^+(\Omega)/\simeq$. Thus if the set $M_1^+(\Omega)/\simeq$ resulting from the above reduction procedure appears to be the set of states of a quantum mechanical model, then the random variables which determine this reduction procedure are to be identified as quantum observables.

All this will be illustrated by two simple examples. The first comes from Holevo (1982).

Example 1. Let $\Omega = \{\omega', \omega'', \omega''', \omega'^{v}\}$ be a four-point set. The set $M_1^+(\Omega)$ can be visualized then as a tetrahedron with vertices corresponding to the elements of Ω. Assume that we are interested only in random variables which satisfy the condition:

$$F(\omega') + F(\omega''') = F(\omega'') + F(\omega'^{v}) \tag{5}$$

where the sum of measures is understood in the usual way; for instance, $F(\omega') + F(\omega''')$ is the measure defined by $(F(\omega') + F(\omega'''))(X) = F(\omega')(X) + F(\omega''')(X)$ for any measurable subset X of the outcome space of F. The set of all such random variables, denoted $\mathcal{F}_q(\Omega)$, will be the starting point of the reduction procedure.

It is evident that $\mathcal{F}_q(\Omega)$ does not separate elements of $M_1^+(\Omega)$, which leads to the equivalence relation $\simeq$ on the tetrahedron described above. It is easy to find that in the considered case $M_1^+(\Omega)/\simeq$ bears the shape of a square. The distinguished set $\mathcal{F}_q(\Omega)$ separates points of Ω; hence the random variables satisfying condition (5) can be regarded as "observables" on the square-shaped "set of states." The obtained structure, consisting of the convex set $M_1^+(\Omega)/\simeq$ and the set of affine maps $\{A_{q,F} \mid F \in \mathcal{F}_q(\Omega)\}$, indeed resembles (to some extent of course) the basic structure of a quantum mechanical model, and is known as the Davies example (Davies, 1972). A more realistic situation of this kind is described by the next example, which essentially comes from Neumann (1985).

Example 2. Let Ω denote the set of points of the unit sphere in $\mathbf{R}^3$ equipped with the natural measurable structure. Among all possible random variables on Ω we distinguish the set $\mathcal{F}_q(\Omega) = \{F_\omega \mid \omega \in \Omega\}$, where

$$F_\omega\colon\ \Omega \to M_1^+\left(\left\{-\frac{1}{2}, +\frac{1}{2}\right\}\right), \qquad F_\omega(\omega')\left(\pm\frac{1}{2}\right) := \frac{1}{2}(1 \pm \mathbf{r}_\omega \cdot \mathbf{r}_{\omega'})$$

with $\mathbf{r}_\omega$ the unit vector of $\mathbf{R}^3$ pointing to $\omega \in \Omega$, etc. Following the reduction procedure, we define the equivalence relation $\simeq$ on $M_1^+(\Omega)$ according to formula (4). It can be easily demonstrated that the resulting factor set $M_1^+(\Omega)/\simeq$ is, as a convex set, identical with the unit ball of $\mathbf{R}^3$. As the reduction map has the 1-to-1 property on the extreme points, we can identify Ω with the surface of the ball $M_1^+(\Omega)/\simeq$.

It is easy to guess that the obtained structure has a direct connection to the quantum mechanical description of spin-1/2 objects. Indeed, it is well known that the set of all statistical operators on the two-dimensional Hilbert space $\mathbf{C}^2$ is affinely isomorphic to the unit ball of $\mathbf{R}^3$ (e.g., Beltrametti and Cassinelli, 1981). The affine map $A_{q,F_\omega}\colon M_1^+(\Omega)/\simeq\ \to M_1^+(\{-1/2, +1/2\})$ generated by $f_\omega \in \mathcal{F}_q(\Omega)$ defines then uniquely the Hermitian matrix $\hat{A}_\omega = \frac{1}{2}(r_{\omega,1}\sigma_1 + r_{\omega,2}\sigma_2 + r_{\omega,3}\sigma_3)$ which represents the projection of spin on the direction of $\mathbf{r}_\omega = (r_{\omega,1}, r_{\omega,2}, r_{\omega,3})$.

Both examples show that one can, by means of the reduction procedure, extract from fuzzy probability theory some quasiprobabilistic structures consisting of a convex set and of a family of measure-valued affine maps on this convex set. They cannot be in general regarded as models of any probability theory, because the nonsimplicial shape of the convex sets in question excludes their representation as $M_p(\Omega')$ for a measurable space Ω'.[4]

[4]Nevertheless, under some technical conditions nonsimplectic convex sets can be affinely embedded into simplexes (Singer and Stulpe, 1992; Busch *et al.*, 1993; Bugajski, 1993).

However, such nonsimplicial convex sets of states are typical for quantum mechanical models. The second example shows that a particular set of this kind can be obtained as the result of a reduction of some particular model of fuzzy probability theory. It should be stressed that this is not an exceptional case: any model provided by standard (i.e., founded on a separable complex Hilbert space $\mathcal{H}$) quantum mechanics can be obtained as the result of reduction of an appropriate model belonging to fuzzy probability theory (Beltrametti and Bugajski, 1995). Consequently quantum mechanics can be regarded as a particular subtheory of fuzzy probability theory.

This result, which in view of the preceding remarks is perhaps not very surprising, becomes comprehensible if we notice the natural conversion of quantum observables into (generalized) random variables described in Section 2 around formula (1). All random variables on $\Omega_{\mathcal{H}}$ generated in the way described there by quantum mechanical observables have to be collected into the set $\mathcal{F}_q(\Omega_{\mathcal{H}}) \subset \mathcal{F}(\Omega_{\mathcal{H}})$ and then taken as the basis of the reduction procedure of the fuzzy-probability model based on $\Omega_{\mathcal{H}}$. Then it seems natural to expect that the reduction procedure will exactly recover the original quantum mechanical model we started with. The reduction map R_M: $M_1^+(\Omega_{\mathcal{H}}) \to M_1^+(\Omega_{\mathcal{H}})/\approx$ for the equivalence relation $\approx$ defined by the set $\mathcal{F}_q(\Omega_{\mathcal{H}})$ of all standard quantum mechanical observables was first constructed by Misra (1974) and then rediscovered by Ghirardi *et al.* (1976) and Holevo (1982). Beltrametti and Bugajski (1995) generalize this result and discuss some of its physical aspects. The fuzzy probability model based on $\Omega_{\mathcal{H}}$ is called there the canonical classical extension of the corresponding model of standard quantum mechanics.

Obviously, the basic probabilistic structure of classical statistical mechanics (which is in fact the same as the basic structure of standard probability theory) can be also obtained from an appropriate fuzzy model. Thus we conclude by noticing that fuzzy probability theory provides a universal probabilistic framework for classical statistical mechanics as well as for standard quantum mechanics.

5. JOINT RANDOM VARIABLES

Once we introduce the general concept of random variable and get some acquaintance with its peculiarities and advantages, we can begin developing fuzzy probability theory following the lines of the standard theory. Even a few steps in this direction offer unexpected discoveries; they can be exemplified by a closer examination of properties of joint random variables.

The standard notion of a joint random variable is tailored to standard random variables and does not admit a direct generalization toward general random variables. However, observables of quantum mechanics, as we have seen, are nothing but fuzzy random variables hidden from view by the machinery of functional analysis, so we can look there for some hints. Indeed, the intensively studied and vividly debated concept of comeasurability of quantum observables, especially in its general form provided recently by Beltrametti and Bugajski (1995), remains meaningful in the framework of fuzzy probability theory. Thus, given two random variables F_1: $\Omega \rightarrow M_1^+(\Xi_1)$, F_2: $\Omega \rightarrow M_1^+(\Xi_2)$, we will say that a random variable $J(F_1, F_2)$: $\Omega \rightarrow M_p(\Xi_1 \times \Xi_2)$ is their joint random variable (so that F_1 and F_2 are comeasurable) if

$$F_i(\omega) = \pi_i(J(F_1, F_2)(\omega)), \qquad i = 1, 2 \tag{6}$$

for any $\omega \in \Omega$, where π_1 and π_2 are the marginal projections in $M_1^+(\Xi_1 \times \Xi_2)$.

It can be seen (Beltrametti and Bugajski, 1995) that, according to what we would expect, any two random variables of fuzzy probability theory possess a joint random variable, i.e., are comeasurable. This natural result entails a surprising conclusion: fuzzy probability theory provides joint random variables for any pair of quantum mechanical observables. This observation sheds new light on the phenomenon of noncomeasurability of quantum mechanical observables, which appears now to be merely a by-product of the traditional Hilbert space-based formalism instead an occurrence of the mysterious nature of the microworld. A noncommuting pair of quantum observables seems to have no joint observable simply because the standard formalism is not able to provide it, while a joint random variable for such a pair does in any case exist in the extended framework of fuzzy probability theory (see Example 3 below).

What makes the notion of joint random variable nontrivial and even intriguing is the fact that in general there are many joint random variables for a given pair of (generalized) random variables. The nonuniqueness of joint random variables is in fact a natural consequence of the evident nonuniqueness of joint measures.

Let us recall that a joint measure of two measures $\mu_1 \in M_1^+(\Xi_1)$ and $\mu_2 \in M_1^+(\Xi_2)$ is a measure, say $J(\mu_1, \mu_2)$, such that $J(\mu_1, \mu_2) \in M_1^+(\Xi_1 \times \Xi_2)$, and

$$\mu_i = \pi_i J(\mu_1, \mu_2), \qquad i = 1,2$$

The set of all joint measures for μ_1, μ_2 will be denoted $\mathcal{J}(\mu_1, \mu_2)$; it contains only one measure if and only if one of μ_1, μ_2 (or both) is a Dirac measure.

Take then two random variables F_1: $\Omega \rightarrow M_1^+(\Xi_1)$, F_2: $\Omega \rightarrow M_1^+(\Xi_2)$. To any point $\omega \in \Omega$ there corresponds now the set $\mathcal{J}(F_1(\omega), F_2(\omega))$ of joint measures for $F_1(\omega)$ and $F_2(\omega)$. Attaching to elementary events ω various elements of the corresponding family $\mathcal{J}(F_1(\omega), F_2(\omega))$, we obtain plenty of functions on Ω which take values in $M_p(\Xi_1 \times \Xi_2)$ and satisfy (6). A unique result occurs only if one of the random variables f_1, f_2 (or both) is standard.

Finally, we have to prove that some of the obtained measure-valued functions are really random variables, i.e., satisfy the measurability condition for effects. In this way we come to the view that the nonuniqueness of joint random variables is as natural and obvious as the nonuniqueness of joint probability measures.

This of course does not explain the nature of this effect. One can speculate about various ways of coupling together two given random variables or about a mutual disturbance of one by the other; however, an acceptable interpretation should arise naturally as a result of inspecting practical instances of occurrence of the nonuniqueness phenomenon.[5]

The following example, which comes essentially from Beltrametti and Bugajski (1995), shows the nonuniqueness of joint random variable and provides joint random variables for a pair of uncomeasurable quantum mechanical observables.

Example 3. Let us come back to the second example of Section 4. Take two particular random variables F_{ω_1}, F_{ω_2} on the sphere Ω; they are generated by the quantum observables of spin projection on the directions of $\mathbf{r}_{\omega_1}$, $\mathbf{r}_{\omega_2}$, respectively. In the framework of standard quantum mechanics the corresponding quantum observables do not have any joint observable (except the case $\mathbf{r}_{\omega_1} = -\mathbf{r}_{\omega_2}$), while the fuzzy probability model constructed over Ω [the canonical classical extension of the quantum model of spin 1/2 according to Beltrametti and Bugajski (1995)] provides many joint random variables for them. Following the above consideration we start by examining the set $\mathcal{J}(F_{\omega_1}(\omega), F_{\omega_2}(\omega))$ of all joint measures for the two measures $F_{\omega_1}(\omega)$, $F_{\omega_2}(\omega)$, ω arbitrarily fixed. Any joint measure $J(F_{\omega_1}(\omega), F_{\omega_2}(\omega)) \in \mathcal{J}(F_{\omega_1}(\omega), F_{\omega_2}(\omega))$ should have the form

$$J(F_{\omega_1}(\omega), F_{\omega_2}(\omega))(+\tfrac{1}{2}, +\tfrac{1}{2}) = \lambda(\omega)$$

$$J(F_{\omega_1}(\omega), F_{\omega_2}(\omega))(+\tfrac{1}{2}, -\tfrac{1}{2}) = \tfrac{1}{2}(1 + \mathbf{r}_{\omega_1}\cdot\mathbf{r}_{\omega}) - \lambda(\omega)$$

$$J(F_{\omega_1}(\omega), F_{\omega_2}(\omega))(-\tfrac{1}{2}, +\tfrac{1}{2}) = \tfrac{1}{2}(1 + \mathbf{r}_{\omega_2}\cdot\mathbf{r}_{\omega}) - \lambda(\omega)$$

$$J(F_{\omega_1}(\omega), F_{\omega_2}(\omega))(-\tfrac{1}{2}, -\tfrac{1}{2}) = \lambda(\omega) - \tfrac{1}{2}(\mathbf{r}_{\omega_1} + \mathbf{r}_{\omega_2})\cdot\mathbf{r}_{\omega}$$

where $\lambda(\omega)$ is an arbitrary real number satisfying the natural conditions imposed by the fact that $J(F_{\omega_1}(\omega), F_{\omega_2}(\omega))$ is a probability measure. If we let ω vary over Ω, then any particular choice of the function $\lambda(\omega)$ would

[5] The nonuniqueness of joint random variables can be actually observed in the framework of operational quantum mechanics. Indeed, it is easy to see that the nonuniqueness of the joint effect for a pair of compatible effects [noticed years ago by Ludwig (1954), p. 89 of the English translation] entails the nonuniqueness of joint observables for a pair of comeasurable operational quantum observables. See Busch *et al.* (1995) for examples.

lead to a joint random variable for $f_{\omega_1}, f_{\omega_2}$. There are many appropriate functions $\lambda(\omega)$; two of them are, for instance,

$$\lambda'(\omega) = \tfrac{1}{4}(1 + \mathbf{r}_{\omega_1}\cdot\mathbf{r}_\omega)(1 + \mathbf{r}_{\omega_2}\cdot\mathbf{r}_\omega)$$

$$\lambda''(\omega) = \min\{\tfrac{1}{2}(1 + \mathbf{r}_{\omega_1}\cdot\mathbf{r}_\omega), \tfrac{1}{2}(1 + \mathbf{r}_{\omega_2}\cdot\mathbf{r}_\omega)\}$$

The nonuniqueness of joint random variables in fuzzy probability theory implies an interesting consequence, called the Bell phenomenon (Beltrametti and Bugajski, 1996). We discuss it in the next section.

6. IS FUZZY PROBABILITY THEORY REALLY NECESSARY?

The above motivation for promoting fuzzy probability theory comes mainly from physics, especially from quantum mechanics. One could agree that the new probability theory is a natural consequence of quantum physics, that it provides a unified and universal background for physical statistical theories, even that it correctly mirrors some peculiarities of the microworld as well as some irrational aspects of psychological or sociological phenomena. Nevertheless one question has to be answered: is it really impossible to get all that from standard probability theory? In fact this is a form of the old problem of "hidden variables" that has haunted quantum mechanics almost since its beginnings.

The most essential difference between standard and fuzzy probability theories lies in the notion of random variable they adopt. The most essential difference between standard and general (fuzzy) random variables is that the latter attach nontrivial measures to some elementary events, contrary to the former, which attach to elementary events Dirac measures only.

Quantum mechanics provides many examples of situations which should be modeled by attaching to an elementary event (i.e., to a pure state) a nontrivial probability measure on an outcome space. It seems that similar situations can occur in other fields closer to everyday experience. The simplest case is if one asks somebody a question which admits only two answers, yes or no, but the person being examined cannot decide which option to choose. The most reasonable description of this "unreasonable" behavior is provided by a probability measure equally distributed over the two possible answers.[6] In such situations fuzzy random variables come into play. Standard probability theory does not admit fuzzy random variables, so it has to offer alternative ways of description.

[6] "To assign equal probabilities to two events is not in any way an assertion that they must occur equally often in any 'random experiment', as Jeffreys emphasized, it is only a formal way of saying 'I don't know'" (Jaynes, 1985).

Assume that we are rather disposed to attach a probability measure $\mu \in M_1^+(\mathbf{R}^1)$ to an elementary event $\omega_0 \in \Omega$. If nevertheless we want to stay inside the framework of standard probability theory, we could attach to ω_0 the expectation value $Exp\mu$ of μ instead of μ itself. In some cases this looks plausible: if, for instance, one is polling a group of people, one can ascribe to those undecided a third value, say 1/2, intermediate between 1 (yes) and 0 (no) of the random variable representing the question asked. This makeshift is of limited use. If, for example, we want to consider correlations of two random variables, say F_1 and F_2, which attach to ω_0 nontrivial probability measures $F_1(\omega_0)$ and $F_2(\omega_0)$, then the standard substitutes $ExpF_1(\omega_0)$ and $ExpF_2(\omega_0)$ for $F_1(\omega_0)$ and $F_2(\omega_0)$ are completely useless.

A more promising alternative arises from the observation that standard random variables are able to generate nontrivial distributions at nonextreme elements of $M_1^+(\Xi)$. If $F: \Omega \to M_1^+(\Xi)$ is a standard random variable (hence attaches only Dirac measures to points of Ω), then its affine extension A_F: $M_1^+(\Omega) \to M_1^+(\Xi)$ [see Section 3, formula (3)] attaches in general nontrivial measures to those elements of $M_p(\Xi)$ which are placed outside the extreme boundary $\partial M_p(\Omega)$. Following this line we could conjecture that the elementary event ω_0 to which we want to attach the probability measure μ is "in fact" not elementary, but rather should be represented by a probability measure over a broader "hidden" measurable space. This is exactly the quantum mechanical idea of hidden variables.

The idea of attaching a more subtle structure to elementary events can be formalized in general terms of the phase-space representations of statistical theories[7] (Singer and Stulpe, 1992; Busch *et al.*, 1993). Some shortcomings of this approach are pointed out in Bugajski (1993); a decisive argument is in fact the one raised by Bell (1964), who demonstrated on a simple example from the quantum theory of spin-1/2 systems that standard probability theory could never reproduce some properties of quantum observables. Having in mind the close connection between quantum mechanics and fuzzy probability theory, we guess that the Bell argument also solves the main question of this section in favor of fuzzy probability theory.

Detailed considerations of the Bell phenomenon can be found in Beltrametti and Bugajski (1996); here we illustrate it by an example, which is not devoid of some quantum mechanical background: a similar situation can be observed when analyzing the original Bell argument.

Example 4. Let Ω be an arbitrary space of elementary events; consider a fixed point $\omega_0 \in \Omega$ and a family of random variables $\{F_1, F_2, F_3, J(F_1,$

[7]The basic scheme of the phase-space representations was established by Prugovečki (Ali and Prugovečki, 1977; Prugovečki, 1986), but the idea can be traced back to the known Wigner distributions.

$F_2)$, $J(F_2, F_3)$, $J(F_1, F_3)\}$ on Ω. We assume that the outcome spaces of F_1, F_2, F_3 are two-point spaces $\Xi_i = \{\xi_i', \xi_i''\}$ and that

$$F_i(\omega_0) = \tfrac{1}{2}\delta_{\xi_i'} + \tfrac{1}{2}\delta_{\xi''}, \qquad i = 1, 2, 3$$

Choose now the joint random variables in such a way that

$$J(F_1, F_2)(\omega_0) = \tfrac{1}{2}\delta_{(\xi_1', \xi_2')} + \tfrac{1}{2}\delta_{(\xi_1'', \xi_2'')}$$

$$J(F_2, F_3)(\omega_0) = \tfrac{1}{2}\delta_{(\xi_2', \xi_3')} + \tfrac{1}{2}\delta_{(\xi_2'', \xi_3'')}$$

$$J(F_1, F_3)(\omega_0) = \tfrac{1}{2}\delta_{(\xi_1', \xi_3'')} + \tfrac{1}{2}\delta_{(\xi_1'', \xi_3')}$$

so the measures generated at ω_0 by $J(F_1, F_2)$ and $J(F_2, F_3)$ show a strong correlation, while the measure $J(F_1, F_3)(\omega_0)$ shows a strong anticorrelation of the corresponding marginal measures. Notice that the fuzzy probability theory can provide random variables having the prescribed properties.

It is easy to realize that there is no probability measure on the eight-point set $\Xi_1 \times \Xi_2 \times \Xi_3$ which would return the above three measures $J(F_1, F_2)(\omega_0)$, $J(F_2, F_3)(\omega_0)$, $J(F_1, F_3)(\omega_0)$ as its marginals. Indeed, if such a global joint measure did exist, then, for instance, the two measures $J(F_1, F_2)(\omega_0)$ and $J(F_2, F_3)(\omega_0)$ would force the third one to be $\frac{1}{2}\delta_{(\xi_1', \xi_3')} + \frac{1}{2}\delta_{(\xi_1'', \xi_3'')}$ instead of $\frac{1}{2}\delta_{(\xi_1', \xi_3'')} + \frac{1}{2}\delta_{(\xi_1'', \xi_3')}$. The nonexistence of a measure which would generate the above measures by means of marginalizations is exactly what is called the Bell phenomenon in Beltrametti and Bugajski (1996).

Such a situation cannot occur in standard probability theory. It is obvious that substituting the original variables F_1, F_2, F_3 by their mean values does not help much, because standard random variables would generate at ω_0 nothing but Dirac measures, which cannot produce nontrivial joint distributions.

The idea of extending the original space of elementary events is of no use either. Indeed, assume that we have, in place of ω_0, a nontrivial measure, say μ_0, on the new space $\tilde{\Omega}$ of "hidden" elementary events. The original random variables F_1, F_2, F_3 should be now extended over $\tilde{\Omega}$, which is by no means automatic (Bugajski, 1993). If we succeed in constructing correct standard representants over $\tilde{\Omega}$, say $\tilde{F}_1$, $\tilde{F}_2$, $\tilde{F}_3$, for the original random variables, then their unique joint random variable $J(\tilde{F}_1, \tilde{F}_2, \tilde{F}_3)$ would generate at μ_0 the joint measure on the eight-point space $\Xi_1 \times \Xi_2 \times \Xi_3$, which inevitably returns the unique joint distributions for pairs of the random variables in question; see the detailed discussion in Beltrametti and Bugajski (1996).

The Bell phenomenon shows that there occur situations which are not describable in terms of standard probability theory; the experimental confirmations of this phenomenon (Aspect, 1976; Aspect *et al.*, 1981, 1982) indicate that the nature itself favors fuzzy probability theory. Notice also that the example above demonstrates the occurrences of the Bell phenomenon are not specific for the particular Hilbert space-based framework of quantum mechanics, but result rather from the fuzziness of the involved random variables (Beltrametti and Bugajski, 1996). This observation encourages the search for occurrences of the Bell phenomenon outside physics, especially in the social sciences, where some deviations from the rules of rationality are likely to appear.

7. REVOLUTION OR CONTINUATION?

Fuzzy probability theory, although it repudiates some of the basic rules of standard probability theory, should be seen as a natural and smooth extension of the latter rather than its radical negation. As I have stressed above, the basic structure of the standard theory is still present inside the fuzzy framework; in particular, standard random variables can be seen as a special class of general random variables. It is natural to conjecture that all notions and results of standard probability theory would be carefully preserved by the fuzzy theory, which possibly would extend them and provide them with a new interpretation.

Moreover, fuzzy probability theory introduces a new kind of consistency into the standard framework. As mentioned above, our general random variables can be found in the standard theory of stochastic processes in the form of Markov kernels. It appears also that random strategies of the standard theory of statistical decisions (for instance, Neveu, 1965) are formally identical to our random variables. The appearance of general random variables at various points of the standard probabilistic landscape signals some inherent tension, which perhaps is caused by a tendency to go beyond the limitations of the standard concept of a random variable. This seems to suggest that fuzzy probability theory could help in formulating a new consistent framework connecting various theories deriving from standard probability theory.

ACKNOWLEDGMENTS

I am greatly indebted to Prof. Enrico Beltrametti, the coauthor of the joint papers (Beltrametti and Bugajski, 1995, 1996) which were basic for the above essay, for many discussions which stimulated the ideas presented above. Many thanks also go to Prof. Karl-Eberhard Hellwig for his interest in this work as well as for his kind hospitality in the Institute of Theoretical Physics of the Technical University of Berlin, where the final version of this paper was prepared. This work was supported by the Alexander von Humboldt-

Foundation (Bonn) and by the State Committee for Scientific Research (KBN, Warsaw) under grant 2/2420/92/03.

REFERENCES

Alfsen, E. M. (1971). *Compact Convex Sets and Boundary Integrals*, Springer-Verlag, Berlin.

Ali, S. T., and Prugovečki, E. (1977). *Journal of Mathematical Physics*, **18**, 219.

Aspect, A. (1976). *Physical Review D*, **14**, 1944.

Aspect, A., Grangier, P., and Roger, G. (1981). *Physical Review Letters*, **47**, 460.

Aspect, A., Grangier, P., and Roger, G. (1982). *Physical Review Letters*, **49**, 91.

Bauer, H. (1981). *Probability Theory and Elements of Measure Theory*, 2nd English ed., Academic Press, London.

Bell, J. S. (1964). *Physics*, **1**, 195.

Beltrametti, E. G., and Bugajski, S. (1995). *Journal of Physics A*, **28**, 3329.

Beltrametti, E. G., and Bugajski, S. (1996). *Journal of Physics A*, in press.

Beltrametti, E. G., and Cassinelli, G. (1981). *The Logic of Quantum Mechanics*, Addison-Wesley, Reading, Massachusetts.

Bugajski, S. (1993). *International Journal of Theoretical Physics*, **32**, 969.

Busch, P., and Quadt, R. (1993). *International Journal of Theoretical Physics*, **32**, 2261.

Busch, P., Hellwig, K.-E., and Stulpe, W. (1993). *International Journal of Theoretical Physics*, **32**, 399.

Busch, P., Grabowski, M., and Lahti, P. J. (1995). *Operational Quantum Physics*, Springer-Verlag, Berlin.

Davies, E. B. (1972). *Journal of Mathematical Physics*, **13**, 39.

Ghirardi, G. C., Rimini, A., and Weber, T. (1976). *Nuovo Cimento*, **36B**, 97.

Holevo, A. S. (1982). *Probabilistic and Stochastic Aspects of Quantum Theory*, North-Holland, Amsterdam.

Jaynes, E. T. (1985). Bayesian methods: General background, in *Proceedings of the Fourth Annual Workshop on Bayesian/Maximum Entropy Methods in Geophysical Inverse Problems (Calgary 1985)*, J. H. Justice, ed., Cambridge University Press, Cambridge, Massachusetts.

Ludwig, G. (1954). *Die Grundlagen der Quantenmechanik*, Springer-Verlag, Berlin [English translation, Springer-Verlag, Berlin (1983)].

Misra, B. (1974). On a new definition of quantal states, in *Physical Reality and Mathematical Description*, C. P. Enz and J. Mehra, eds., Reidel, Dordrecht, Holland, pp. 455–476.

Newmann, H. (1985). The size of sets of physically possible states and effects, in *Recent Development in Quantum Logic*, P. Mittelstaedt and E.-W. Stachow, eds., Bibliographisches Institut, Mannheim, pp. 337–348.

Neveu, J. (1965). *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco.

Prugovečki, E. (1986). *Stochastic Quantum Mechanics and Quantum Space Time*, 2nd ed., D. Reidel, Dordrecht, Holland.

Reed, M., and Simon, B. (1972). *Methods of Modern Mathematical Physics 1. Functional Analysis*, Academic Press, New York.

Singer, M., and Stulpe, W. (1992). *Journal of Mathematical Physics*, **33**, 131.