Reprint from Philosophy of Science, The Delaware Seminar, Vol. 2, 1962-63, edited by B. Baumrin, pp. 319-337, New York: Wiley and Sons, 1963.

The Role of Probability in Quantum Mechanics *

Patrick Suppes, *Stanford University*

The view I attempt to support in this paper by various sorts of arguments is that if we examine the structure of quantum mechanics for paradoxical or surprising implications, it is not the interpretation of the uncertainty principle as showing the untenability of determinism that is most important. I take it that many would now agree that strict determinism was a fragile flower nurtured for a few decades in the special climate of classical mechanics, particularly as formulated by LaPlace. The once broad claims for determinism have in many ways been replaced by universal claims for the theory of probability and the methodology of statistics, although I do not wish to suggest for a moment that determinism as a thesis and probability theory as a methodological cornerstone of science are on the same conceptual footing. Rather, my point is to try to show that the really radical intellectual thing about quantum mechanics is not its challenge to determinism as a philosophical thesis but its challenge to probability theory and modern mathematical statistics as the universal methodology of all empirical science.

Let us begin by considering some of the standard interpretations of the Heisenberg uncertainty relation. Landau and Lifshitz (ref. 2, p. 4) have this to say:

> whereas in classical mechanics a particle has definite coordinates and velocity at any given instant, in quantum mechan-

* It is a pleasure to dedicate this paper to Professor Charles Loewner on the occasion of his seventieth birthday.

ics the situation is entirely different. If, as a result of measurement, the electron is found to have definite coordinates, then it has no definite velocity whatever. Conversely, if the electron has a definite velocity, it cannot have a definite position in space. For the simultaneous existence of the coordinates and velocity would mean the existence of a definite path, which the electron has not. Thus, in quantum mechanics, the coordinates and velocity of an electron are quantities which cannot be simultaneously measured exactly, i.e., they cannot simultaneously have definite values. We may say that the coordinates and velocity of the electron are quantities which do not exist simultaneously.

Similar views are to be found in Reichenbach (ref. 6, p. 13). He says that the uncertainty

> relation . . . can be interpreted in the form: When the position of the particle is well determined, the momentum is not sharply determined and vice versa.

As can be seen from these two quotations the orthodox viewpoint is to emphasize that the uncertainty relation means that when position is determined exactly the momentum is not determined at all, or when momentum is sharply determined, position is badly determined, if at all. It is particularly interesting to note that Landau and Lifshitz, and Reichenbach as well, do not give a quantitative interpretation of the uncertainty relation.

A somewhat more sophisticated characterization is given in Schiff's well-known textbook (ref. 7, p. 7):

> The relation . . . means that the component of the momentum of a particle cannot be precisely specified without our loss of all knowledge of the corresponding component of its position at that time, that a particle cannot be precisely localized in a particular direction without our loss of all knowledge of its momentum component in that direction, and that in intermediate cases the product of the uncertainties of the simultaneously measurable values of corresponding position and momentum components is at least of the order of magnitude of $\hbar$.

Schiff does mention the "intermediate cases" when neither position nor momentum is sharply determined, but none of the authors cited give anything like a thorough probability interpretation of the inequality. I have, for instance, seen no book on quantum mechanics in which an "intermediate case" is described in quantitative or experimental detail. What single experiment can be cited on behalf of the presumed intermediate cases, that is, in what experiment can we determine approximately, apart from other sources of experimental error, that

$$\sigma_x > 0$$

$$\sigma_p > 0$$

$$\sigma_x \sigma_p \geq \hbar/2$$

and more particularly, what numerical values of the standard deviations σ_x and σ_p are found? The third of the three inequalities is a standard formulation of the uncertainty principle. From its form one would naturally infer that it is possible simultaneously and jointly to determine the standard deviations σ_x and σ_p, but, as we shall see, this is not in general possible.

If we consider this inequality from the standpoint of probability theory, there seem to be no persuasive arguments whatsoever for accepting the orthodox interpretation. To say this is not to say that the orthodox interpretation is necessarily wrong but to emphasize that there is a very large logical gap between the claims that are ordinarily made for the meaning of the uncertainty principle and the literal interpretation of the principle itself as a statement about the product of two standard deviations. At this point it may be instructive to consider other interpretations that are possible within the framework of standard probability theory.

Perhaps the first point that needs clarification is the distinction between measurements that are exact when taken at a given instant and the probability distribution of quanti-

tative properties of objects, when those objects are "identically prepared" for experimental manipulation. Consider the following simple example. Let H be the random variable defined on the set of all human beings such that for any person x, $H(x)$ is the height of x, and let W be the random variable measuring the weights of human beings. Now we do not need to collect any new experimental data in order to know that both of these random variables will have considerable variation in the population of human beings. This variation in itself has nothing to do with the measuring process. For purposes of any investigation in which we would be interested in the height and weight of human beings, we can assume that both these variables may be measured with absolute accuracy. It will still be the case that the product of the standard deviations of the distributions of the two random variables will exceed some fairly large positive number, in spite of the fact that for any individual x, $H(x)$ and $W(x)$ are perfectly definite. On the basis of the positive variances of both height and weight, we could paraphrase Landau and Lifshitz and say that if the height of a person has a definite value, then his weight has no definite value whatsoever. The patent absurdity of this is obvious, which is just to say that Landau and Lifshitz have explained the meaning of the uncertainty relation rather badly.

The immediate rejoinder in considering this example in relation to the Heisenberg uncertainty inequality is that of course there is a drastic difference. For we know that the members of the set of human beings are distinct individuals with quite different properties of height and weight. On the other hand, when we prepare an electron or some other "elementary" particle for experimental observation, we think of all electrons brought into the experimental situation as being in some sense identical or equivalent, at least being identical in a sense that we do not expect the various human beings to be identical. Yet there does not seem any reason to accept this principle of identity or equivalence for elementary particles other than as a kind of

prejudice derived from two hundred years of classical particle mechanics. It seems apparent that if this principle of identity is not adopted, then it is perfectly consistent to hold that in any given instant, position and momentum of an elementary particle may be measured exactly, but if the measurements are repeated on other elementary particles of the same kind under what the experimenter thinks are identical experimental conditions, then a genuine probability distribution of both kinds of measurements will be obtained, and in fact the product of the standard deviations of these two distributions will satisfy the Heisenberg uncertainty relation.

This is the kind of situation that obtains in a wide range of psychological experiments. It may be useful to describe in detail one kind of such experiment that has a number of affinities from a methodological standpoint with the issues we are discussing in connection with the uncertainty principle. I have in mind a simple paired-associate experiment. This simple example from mathematical behavior theory conjoins four properties that collectively challenge overly simple analyses of the uncertainty relation: (*1*) homogeneity or 'identity' of items is assumed; (*2*) a detailed theory of the learning phenomena is given; (*3*) the measurements of the trial of last error are absolutely exact; (*4*) a positive variance for the distribution of last errors is predicted. The orthodox interpretation of quantum mechanics needs to make clearer why a similar conjunction of properties could not be postulated in quantum mechanics. On this point I must be absolutely clear. I am not saying that an interpretation of quantum mechanics in terms of exact joint measurements of position and momentum is correct. We need to be told in more detail why it is wrong. (It is the burden of the latter part of this paper to show it must in fact be rejected for deeper reasons than those provided by the uncertainty relation.)

In our simple paired-associate experiment, the task for the subject is to learn to associate each one of a list of nonsense syllables with an appropriate response. In a typical

setup the list might consist of 20 nonsense syllables of the form CVC (consonant, vowel, consonant). The responses are given by pressing one of two keys. On a random basis ten of the syllables are assigned to key 1 and ten to key 2. The subject is shown each nonsense syllable in turn, is asked to make a response, and is then shown the correct response by one of several devices, for example, by the illumination of a small light above the correct key. After the subject has proceeded through the list once, he is taken through the list a second time, but the order or presentation of the 20 items is randomized. A criterion of learning is set; for example, four times through the list without a mistake. The subject is asked to continue to respond until he satisfies this criterion. The criterion is selected so as to give substantial evidence that the subject has indeed learned the correct association between each stimulus item and its appropriate response. There are a number of statistics to be derived from a simple and very adequate model for these phenomena. For the present purpose it will be sufficient to concentrate on the distribution of the trial of last error.

The simple model to be applied to the phenomena is the following. The subject begins the experiment by not knowing the arbitrary association established by the experimenter between individual stimuli and the response keys. He is thus in the unconditioned state U. On each trial there is a probability c that he will pass from the unconditioned state to the conditioned state C. It is postulated that this probability c is constant over trials and independent of responses on preceding trials. Once the subject passes into the conditioned state it is also postulated he remains there for the balance of the experiment. A simple transition matrix for the model, which is a first-order Markov chain in the two states U and C, is the following:

	C	U
C	1	0
U	c	$1-c$

To complete the model for the analysis of experimental data it is also necessary to state what the probabilities of response are in the two states U and C. When the subject is in the unconditioned state, it is postulated that there is a guessing probability p of making a correct response, and this guessing probability is independent of the trial number and the preceding pattern of responses. When the subject is in the conditioned state, the probability of making a correct response is postulated to be 1. With these assumptions it may be shown that the distribution of the trials on which state C is entered is geometric, where c is the parameter of the geometric distribution. We recall that the mean of the geometric distribution with parameter c is $1/c$ and its variance is the quantity $(1 - c)/c^2$. Unfortunately, it is not possible to observe when the subject enters state C, for he may or may not have guessed the correct responses on immediately preceding trials. What is behaviorally observable is the trial number of the last error on each item. This distribution is, as would be expected, approximately geometric. The exact distribution is as follows. Let n be any trial. Then the probability $f(n)$ that the last error with respect to responses to a given stimuli occurs on trial n is:

$$f(n) = \begin{cases} bp & \text{for } n = 0 \\ b(1 - p)(1 - c)^{n-1} & \text{for } n > 0 \end{cases}$$

where $b = c/[1 - p(1 - c)]$. The mean, μ, and variance, σ^2, of this distribution are:

$$\mu = b(1 - p)/c^2$$

$$\sigma^2 = \mu[2/c - 1 - \mu]$$

(Note that if $p = 0$, we just obtain the geometric distribution.) For our purposes the important thing about this example is that the variance of the trial of last error is certainly not zero, in any except degenerate cases. It is important to realize that this positive variance is not due to any hypothesis of individual differences. The hypothesis of indi-

vidual differences among subjects or among stimulus items would be brought into the theory by postulating variation in the parameter c and possibly the guessing parameter p. The assumption of a constant c and a constant p either for all stimulus items or a group of subjects is equivalent to assuming homogeneity of items or of subjects. In spite of this assumption of homogeneity, the predicted behavior has a positive variance. The experimenter attempts to select items for a given subject as carefully as he can in terms of empirical criteria of homogeneity. No matter how carefully such items are selected, a wide variety of studies show that a distribution of last errors with a positive variance is always obtained in such paired-associate experiments.

In connection with the earlier quotations about the interpretation of the uncertainty relation, it should be remarked that in the case of the empirical observation of the distribution of trials of last error in a given experiment there is no problem whatsoever concerning errors of measurement. The measurements are absolutely exact. They are discrete observations, for which it is necessary only to record which of two keys was pressed by the subject on a given trial. In other words, the positive variance and thus the uncertainty in the trial of last error, are in no sense due to the measurement procedure.

Let us try transferring these results on exact measurements with a positive variance to a quantum-mechanical context. As a simple example we may consider a one-dimensional harmonic oscillator in the ground state. It may be shown that the theoretical distribution $f(x)$ of position for this physical system is given by the following expression:

$$f(x) = \frac{\alpha}{\pi^{1/2}} e^{-\alpha^2 x^2}$$

This expression is that for a normal density with mean zero and variance $\sigma^2 = \alpha^2/2$, where α is a physical constant defined in terms of h and mass m. More precisely,

$$\sigma^2 = \frac{\hbar}{2}\sqrt{Km}$$

where K is the constant arising in the expression for the potential energy: $V(x) = \frac{1}{2}Kx^2$. This distribution with its positive variance for position is derived from the theory in quite straightforward terms. The important point at the moment is that this derivation in no way mentions any processes of measurement. An obvious empirical interpretation of the result is that if measurements of position are made for a harmonic oscillator in the ground state, then the obtained measurements will fit a normal distribution with mean zero and variance as indicated. There is nothing as such in the theory that says the procedures of measurement themselves give rise to the distribution.

A related point of interpretation of the uncertainty relation is the following. The standard verbal formulations or interpretations imply, or tend to imply, that the experimenter may at will decide to measure a particular physical variable with arbitrary precision. Of course, if he does so he must sacrifice corresponding precision in conjugate physical variables. But the example of the linear oscillator in the ground state suggests that this interpretation is wrong. For a given energy state of a system that is a linear oscillator it is not possible to make a precise measurement of position or momentum. In the case of the oscillator in the ground state, all that can be obtained, no matter how exact the experimental procedure, is a normal distribution with mean zero and positive variance as indicated for the position, and similar results obtain for the momentum.

The tenor of these remarks suggests that I am proposing it may be consistent with the fundamental theory of quantum mechanics to measure position and momentum exactly on any given occasion. As the measurements are repeated on identically prepared particles, distributions with positive variance will, of course, be obtained, and these will satisfy the uncertainty relation. This state of affairs would

fit in very well with ordinary probability concepts. Unfortunately, it is not a possible interpretation of the uncertainty relation. Such an interpretation requires that there be a *joint* probability distribution of position and momentum at any given instant. However, if the usual formalism is applied to the computation of the joint distribution of two conjugate physical variables like position and momentum, then the distribution obtained by standard arguments is in general not a *proper* probability distribution at all. This result raises serious difficulties for any attempt to interpret quantum mechanics within a standard probability framework. It is extremely hard to think of any other empirical examples of physical variables having this property of not possessing a genuine joint probability distribution.

In order to bring out the full significance of having a joint distribution of position and momentum, it will be useful first to consider a much simpler case that will require no detailed computations of any sort. Suppose that we have two fair coins and we are flipping them in the following two ways. In the first case, we flip the first coin and then whatever comes up, either heads or tails, we automatically turn the second coin so that the same side is facing up. In the second case, we flip the two coins at the same time and assume that the physical handling of the coins is such that the outcomes are statistically independent.

First of all, then, let us write down the joint distribution for these two different cases. We shall use a subscript 1 for the first case and a subscript 2 for the second case, and we shall use a subscript a for the first coin and a subscript b for the second coin. In the first case the four possible outcomes have the following probabilities.

$$P_1(H_a, H_b) = \tfrac{1}{2} \qquad P_1(T_a, H_b) = 0$$
$$P_1(H_a, T_b) = 0 \qquad P_1(T_a, T_b) = \tfrac{1}{2}$$

Because of the physical conditions of this case, it is clear that the probability must be zero of having H_a and T_b, or also T_a and H_b.

In the second case, the joint distribution of heads and tails of the two coins assumes quite a different form. On the assumption of statistical independence, we have at once that the probability of any two outcomes, for example, H_a and T_b, is simply one-fourth:

$$P_2(H_a, H_b) = \tfrac{1}{4} \qquad P_2(T_a, H_b) = \tfrac{1}{4}$$
$$P_2(H_a, T_b) = \tfrac{1}{4} \qquad P_2(T_a, T_b) = \tfrac{1}{4}$$

In spite of the clear difference between these two joint distributions, that is, P_1 and P_2, there are a number of probability statements that are the same for both of them. In particular, the two marginal distributions for coins a and b are precisely the same. By marginal distributions we mean just the probability statements about the single coins. Thus, we have

$$P_1(H_a) = P_1(H_b) = P_2(H_a) = P_2(H_b) = \tfrac{1}{2}$$
$$P_1(T_a) = P_1(T_b) = P_2(T_a) = P_2(T_b) = \tfrac{1}{2}$$

Another way of putting it is that if we are able only to observe one coin at a time, that is, either coin a or coin b, we are not able to infer from such observations which of the two joint distributions is the true one. As this example clearly illustrates, the marginal distributions tell us a great deal less about the behavior of a "system" of two coins than does the joint distribution.

To obtain some numerical results similar in spirit, but of course much simpler than those we obtain in quantum mechanics in dealing with such physical variables as position and momentum, we may define for each coin a random variable. This random variable is a function defined on the two possible outcomes, heads or tails, and for each coin the random variable has the value 1 when the outcome is heads and the value zero when the outcome is tails. We shall use X_a for the random variable corresponding to coin a, and X_b for the random variable corresponding to coin b. As would be expected from what we have said about the

identity of the marginal distributions for the two different joint distributions P_1 and P_2, the expected values or means $E(X_a)$ and $E(X_b)$ are the same for P_1 and P_2. Namely,

$$E_1(X_a) = E_1(X_b) = E_2(X_a) = E_2(X_b) = \tfrac{1}{2}$$

Similar results obtain for any other quantities we want to compute for X_a and X_b separately, that is, the quantities are the same for the two distributions P_1 and P_2. For example, the variances of X_a and X_b have the following form:

$$\mathrm{Var}_i(X_a) = \mathrm{Var}_i(X_b) = \tfrac{1}{2}(1 - \tfrac{1}{2}) = \tfrac{1}{4}, \text{ for } i = 1, 2$$

But once we move to any considerations involving both variables at the same time, the results are quite different. For example, consider the covariance of X_a and X_b, defined as follows:

$$\mathrm{Cov}(X_a, X_b) = \sum_{x_a} \sum_{x_b} (x_a - \bar{x}_a)(x_b - \bar{x}_b) P(X_a = x_a, X_b = x_b)$$

where $\bar{x}_a = E(X_a)$ and $\bar{x}_b = E(X_b)$. The covariance has the following distinct values for P_1 and P_2.

$$\begin{aligned}\mathrm{Cov}_1(X_a, X_b) = {} & (1 - \tfrac{1}{2})(1 - \tfrac{1}{2}) \cdot \tfrac{1}{2} \\ & + (1 - \tfrac{1}{2})(0 - \tfrac{1}{2}) \cdot 0 \\ & + (0 - \tfrac{1}{2})(1 - \tfrac{1}{2}) \cdot 0 \\ & + (0 - \tfrac{1}{2})(0 - \tfrac{1}{2}) \cdot \tfrac{1}{2} = \tfrac{1}{4}\end{aligned}$$

And

$$\begin{aligned}\mathrm{Cov}_2(X_a, X_b) = {} & (1 - \tfrac{1}{2})(1 - \tfrac{1}{2}) \cdot \tfrac{1}{4} \\ & + (1 - \tfrac{1}{2})(0 - \tfrac{1}{2}) \cdot \tfrac{1}{4} \\ & + (0 - \tfrac{1}{2})(1 - \tfrac{1}{2}) \cdot \tfrac{1}{4} \\ & + (0 - \tfrac{1}{2})(0 - \tfrac{1}{2}) \cdot \tfrac{1}{4} = 0\end{aligned}$$

Similarly, the correlation coefficient ρ of X_a and X_b is defined in general as follows

$$\rho(X_a, X_b) = \frac{\mathrm{Cov}(X_a, X_b)}{\sigma(X_a)\sigma(X_b)}$$

and from the calculations already made, it is clear that for the joint distribution P_1, $\rho_1(X_a, X_b) = 1$ and for the joint distribution P_2, $\rho_2(X_a, X_b) = 0$. The covariance and correlation are measures of the dependency relationship holding between two events or random variables. When the events are independent, the covariance and correlation coefficient are both zero, as is the case for X_a and X_b under distribution P_2.

These computations for the simple random variables we have defined correspond to computations for expected positions and the like in quantum mechanics. It will be useful now to look at some corresponding results for a simple physical case in quantum mechanics. I shall not here justify the results stated, but I have indicated in a previous paper [8] how they may be derived in a straightforward fashion from the usual formulations of classical quantum mechanics. For purposes of discussion, I selected one of the simplest cases possible, namely, the one-dimensional harmonic oscillator already mentioned. By applying the usual methods of calculation, we may show that in the ground state the harmonic oscillator does indeed have a genuine joint distribution of position and momentum, given by the expression

$$f(p, x) = \frac{1}{\hbar\pi} \exp\left(-\alpha^2 x^2 - \frac{p^2}{\hbar^2\alpha^2}\right) \tag{1}$$

where α is the familiar physical coefficient defined earlier. Integrating out x, we obtain for the marginal distribution of momentum

$$f(p) = \frac{1}{\alpha\hbar\sqrt{\pi}} \exp\left(-\frac{p^2}{\hbar^2\alpha^2}\right)$$

which is a normal density with mean zero and variance $\sigma_p^2 = \hbar^2\alpha^2/2$. Integrating out p, we obtain for the marginal distribution of position

$$f(x) = \frac{\alpha}{\sqrt{\pi}} \exp\left(-\alpha^2 x^2\right)$$

which is the normal density with mean zero and variance $\sigma_x^2 = 1/2\alpha^2$, previously mentioned. Then we know at once that the Heisenberg uncertainty relation is satisfied by the product of the standard deviations σ_p and σ_x of the two marginal distributions:

$$\sigma_p\sigma_x = \frac{\hbar\alpha}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}\alpha} = \frac{\hbar}{2}$$

It is perhaps worth noting that in the present case we get an actual equality rather than the weak inequality of the general case. Also of particular interest is the observation that the joint distribution (1) shows that the distributions of momentum and position are statistically independent because $f(p)f(x) = f(p,x)$. And it of course follows from this independence result that the covariance and correlation coefficient of position and momentum are zero.

The situation is quite different when we turn to the first excited state of a harmonic oscillator. For simplicity, let us replace the momentum p by the propagation vector $k = p\hbar$. We then have for the joint density of k and x, the following expression:

$$f(k, x) = \frac{4}{\pi}\left[\exp\left(-\alpha^2x^2 - \frac{k^2}{\alpha^2}\right)\right]\left[\alpha^2x^2 + \frac{k^2}{\alpha^2} - \frac{1}{2}\right]$$

Unfortunately, the function f in this case is not a genuine probability density. In particular, the function is negative for those values of k and x such that

$$\alpha^2x^2 + \frac{k^2}{\alpha^2} < \frac{1}{2}$$

Similar results are obtained for other states of the harmonic oscillator; namely, the joint distributions obtained by the usual arguments are not genuine distributions at all. This situation is characteristic in quantum mechanics. It is only the unusual or special case for which a genuine joint distribution is obtained. From a statistical standpoint, the

covariance or correlation of position and momentum would be at least as natural and important to consider as the product of their standard deviations, but when their joint distribution does not exist we cannot in a meaningful way talk of their covariance or correlation.

The strangeness of these results from a methodological standpoint is difficult to overemphasize. As we all know, the applications of probability theory range over all domains of science, and have become increasingly important in the past several decades. The mathematical techniques of probability theory, as well as the conceptual foundations, have received an enormous amount of attention from mathematicians, statisticians, and philosophers. In the several domains of application with which I am familiar or which I have at least cursorily inspected I have not been able to find a single example having the conceptual status of these results about the non-existence of a joint distribution in quantum mechanics. Moreover, the idea that two physical variables that we are able to observe even separately must have a joint probability distribution is a very deep and natural one from the standpoint of probability theory. It may, of course, not be the case that we can make direct observations on this joint distribution, but what is unusual and dismaying is to find a case in which the postulate of the mere *existence* of a joint distribution is inconsistent with the fundamental theory of the phenomena.

Those of you among the audience who are not already familiar with these results will, roughly speaking, probably have one of two reactions. Those whose backgrounds are in physics will tend to accept the non-existence of the joint distribution of conjugate physical variables with scarcely a murmur, and say that here is simply another piece of evidence of the inadequacy of classical ideas. I shall have more to say in a moment about this rather casual attitude toward probability theory. On the other hand, those who are more thoroughly acquainted with probability theory than quantum mechanics will undoubtedly tend to

think initially that it is perhaps a simple and not very deep matter to change the quantum-mechanical computations in such a way as to obtain genuine distributions. In the present lecture I have not attempted to examine this rather technical point, or, in fact, even to substantiate my own computations in any mathematical detail. I do, however, offer assurances that the framework within which I have performed these computations is perfectly standard and derives from classical and fundamental work by Wigner [9] and Moyal.[5] It is true that other approaches to probability questions within classical quantum mechanics can be found in the literature. But generally speaking, these lead to even more strange results. A typical example is to be found in Diracs' paper,[1] in which he formulates a theory that leads to probability distributions that are complex valued. It need hardly be said that no direct empirical significance from a probability standpoint can be attached to such mathematical functions. It would, in fact, seem to be a complete misnomer to call them probability distributions in any sense.

The recent proposals of Margenau and Hill [3,4] to avoid these difficulties by a bold redefinition of the joint probability distribution are clearly inconsistent with the standard formulations of quantum mechanics. Reasons for skepticism about their proposals are too technical to enter into here. Suffice it to say that it is seldom possible to redefine a derived quantity like that of a joint distribution in the context of a complex theory and hope to end up with a complete theory that is both mathematically exact and consistent.

It is doubtful that any simple way can be found of avoiding the perplexing and paradoxical problems that arise in quantum mechanics, when probability notions are developed with anything even faintly approaching the thoroughness with which they are used in other disciplines. For those familiar with the applications of probability and mathematical statistics in mathematical psychology or mathematical economics, it is surprising indeed to read the treatments

of probability even in the most respected texts of quantum mechanics. I do not mean by this remark to belittle in any sense the magnificent and profound accomplishments embodied in quantum mechanics. These accomplishments, as well as the subtlety of the kind of results achieved, we surely all recognize. What is surprising is that the level of treatment both in terms of mathematical clarity and in terms of mathematical depth and detail is surprisingly low. Probability concepts even have a strange and awkward appearance in quantum mechanics, as if they had been brought within the framework of the theory only as an afterthought and with apology for their inclusion. I cannot, for instance, recall reading a single book or article on quantum mechanics by a physicist which uses the fundamental notion of a random variable in an explicit manner, although this notion is central to every modern treatise on probability theory. That many physicists still find it difficult to feel at home with probability notions is well attested to by the large number of people who prefer the more classical physical language of wave mechanics to talk about probability distributions and expected values. The intellectual tension that exists between the widely applicable methodology of modern probability and statistical theory on the one hand, and quantum mechanics as the boldest and most important scientific theory of the twentieth century, with its peculiar and strange use of probability—this tension will surely receive a different final resolution than is now in view.

If my diagnosis of the antinomy between probability theory and quantum mechanics, exemplified in the nonexistence of the joint probability distribution of conjugate physical variables, is correct, then it is very likely also correct to say that we are confronted with one of those fundamental conflicts that have in the past been so important to the history of mathematics and science. Perhaps the oldest clear example of such a conflict is to be found in the discovery of the irrational by the Pythagoreans. The Pythagorean philosophers and mathematicians found that the

ratio of the diagonal to the side of a square could not be expressed by a "number," that is, by what mathematicians would now call a rational number—a number which is the ratio of two integers. We can, if we like, regard the Pythagorean derivation of the irrationality or incommensurability of the diagnoal of the unit square as the derivation of a contradiction within the mathematical framework accepted at that time. In the same way, we could by adding plausible postulates to the usual methodology of probability theory construe the non-existence of a joint distribution of conjugate physical variables as the derivation of a contradiction from the joint assumptions of probability theory and quantum mechanics. The decision as to whether we are to regard the conflict as one issuing in a formal contradiction is not of critical importance. What is of fundamental importance is the recognition that two of the most powerfully entrenched ways of looking at the world lead to a fundamental conceptual conflict.

What happened in the Pythagorean case of the discovery of incommensurable line segments is well known. The fundamental concept of number was changed and extended to include irrational numbers, but the chronicle of this change and the acceptance of irrational numbers is a long and torturous one in the history of mathematics. The Greeks themselves did not really directly accept it, but in fact, rejected numerical algebra as an adequate instrument for the study of geometrical problems. The theory of proportion of Theatetus and Eudoxus, beautiful as it is from a mathematical standpoint, is clearly a geometrical escape. It was not really until many centuries later, due to the work of Weierstrass, Dedekind, and others in the nineteenth century, that the theory of real numbers, in particular of irrational numbers, was put on a sound mathematical basis as a theory in its own right.

It may be claimed that I have exaggerated the conflict between probability theory and quantum mechanics. Those who wish to support such a claim will emphasize that the

problem of the non-existence of the joint distribution of conjugate variables arises in a natural way out of the quantum-mechanical theory of measurement. I have not in this paper said a great deal about the difficult and numerous problems of measurement in quantum mechanics. For the purposes of the present point, I think it is sufficient to mention, however, that it is not the theory of measurement *per se* that can be made to assume the burden of the conflict. For the kind of peculiar probability results derivable from quantum-mechanical theory exist nowhere else in the theory of measurement. Secondly, these non-existence results do not depend on any specific assumptions about the processes of measurement, but follow in a natural and simple way from the fundamental theory.

As I have already said, it is certainly possible to avoid any implication of a formal contradiction from the non-existence of a joint distribution of conjugate variables. What I think is much more difficult to avoid is the sense of tension and instability in the confrontation of quantum mechanics and probability theory. The existence of this conflict will surely lead to further fundamental conceptual changes in our basic scientific view of the world.

References

1. P. A. M. Dirac, *Reviews of Modern Physics*, **17**, 195 (1945).
2. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, Pergamon Press, London, 1958.
3. H. Margenau, *Philosophy of Science*, **30**, 138 (1963).
4. H. Margenau and R. N. Hill, *Progress of Theoretical Physics*, **26**, 722 (1961).
5. J. E. Moyal, *Proceedings of the Cambridge Philosophical Society*, **45**, 99 (1949).
6. H. Reichenbach, *Philosophic Foundations of Quantum Mechanics*, University of California Press, Berkeley, California, 1944.
7. L. I. Schiff, *Quantum Mechanics*, McGraw-Hill, New York, 1949.
8. P. Suppes, *Philosophy of Science*, **28**, 378 (1961).
9. E. Wigner, *Physical Review*, **40**, 749 (1932).