Probability, Determinism and Objectivity in Classical and Quantum Physics

In classical physics all these ideas appear, at least prima facie, relatively well-defined and unproblematical. In quantum physics, their role is much more controversial.

Determinism in classical physics

It is important to distinguish the idea of *determinism* (given, say, complete initial conditions, the subsequent motion of a closed system is uniquely determined: an "ontological" statement, i.e., one concerning the *behavior* of the system) from *predictability* (it is possible under certain conditions to predict the future motion with certainty: an "epistemological" statement, i.e., one concerning our *knowledge* of that behavior). Crudely speaking, complete predictability of the behavior of a given system would seem prima facie to require it to behave deterministically, but the converse need certainly not be true. For example, if we consider a closed mechanical system described by Newton's laws, then according to our definition, it is certainly deterministic, and thus an exact knowledge of all coordinates† $q_j(0)$ and velocities $\dot{q}_j(0)$ at time zero enables prediction into the indefinite future. But some difficulties:

(1) Effect of errors in initial data: in a simple ("integrable") system, an error δq in our initial data will in general only lead to an error $\sim \delta q$ at all subsequent times (e.g., if our measurement of the position of Mars is very slightly wrong, our prediction of its position in the future will be wrong by roughly the same factor, but no worse). However, in chaotic systems, this is not necessarily so: a small error can very quickly be amplified. We could still try to define "determinism" by the statement: given some finite time T in the future, and a required error $\leq \Delta$, there exists some small number ε such that if we can achieve $\delta q(0) \leq \varepsilon$, we will get $\delta q(T) \leq \Delta$. Of course, this may not be very useful in practice!*

(2) Effect of "unforeseen" external perturbations – physically, no system is ever completely closed! (E.g., consider the impact of a comet on the Solar System. Also, cosmic rays in superfluid ^{3}He.)

(3) In special relativity (or any theory with a finite maximum propagation of causal effects), for any given observer there are *inevitably* events that at any given time are "elsewhere", and therefore inaccessible to him, but in the future will lie in his past light cone, and therefore be able to affect him; cf. lecture 15.

A further point to note is that the link between determinism and causality is not problem-free (cf. the end of lecture 7). Certainly the fact that the initial conditions "cause" the subsequent motion is not uniquely prescribed by Newtonian

†If the notation $q_j(t)$ confuses you, just substitute x for q everywhere on this page!

*On the relevance (or not) of chaos to the concept of determinism, see Cushing ch. 12, sections 4 and 5, and the cited references.

physics, which is equally compatible with either of the statements (`i`) $q(T)$ and $\dot{q}(T)$ determine $q(0)$, and (`ii`) $q(0)$ and $q(2T)$ determine $q(T)$ (or, in fact, an infinite number of similar conditions). We will need to return to this question in lecture 25.

In spite of all, however, we do have the confidence, in classical physics, that we can "in principle" attain the conditions of a sufficiently isolated system, and sufficiently accurate measurement, that we can predict outcomes at a later time with any desired accuracy. Is this confidence actually justified? In fact, already in pre-quantum days there were some suggestions (Exner 1920) that perhaps nature is fundamentally indeterministic at the microlevel, and an apparent determinism arises as a result of the "law of large numbers" (see below). This would, however, apparently imply a belief that Newtonian mechanics cannot be pushed arbitrarily far into the micro-realm before something else takes over.

Objectivity in classical physics

The word "objectivity" has, of course, an everyday usage: crudely speaking we say that someone judged the evidence for a particular proposition, e.g., that A killed B, "objectively" if we think that he/ she was not influenced by a wish that the proposition in question should or should not be true. The various technical uses of the term in philosophy are only fairly distantly related to this everyday usage. In its weakest sense, the statement that a particular phenomenon "objectively" exists may mean no more than it can be verified by any number of independent observers, and is then used to distinguish such phenomena from others such as hallucinations, fantasies, and other such things which we generally take to be in some sense properties of the mind of a particular observer (though perhaps this distinction is not quite as logically necessary as we would tend to assume). In this sense, "objectivity" in effect means no more and no less than "intersubjectivity".

Does this weak sense of the word really do justice to our intuitive prejudice that there is "something out there" in the physical world that is independent of whether or not we are observing it? E.g., does the statement that "there is a tree outside the window of this room" really mean no more than that there are appropriate sense experiences that any qualified (i.e., fully seeing, etc.) observer will have that we associate with the tree? The common sense answer is clearly that it means a good deal more than that; but it is the business of philosophers to doubt common sense, and there is a long tradition, which is often loosely called "idealism", that stems in particular from the ideas of Bishop Berkeley in the 18th century, and which holds that no such extra element exists; i.e., that objectivity, so far as the concept has a meaning, indeed reduces to nothing more than intersubjectivity. (Berkeley's own solution to the problem was somewhat remarkable.)

Note that, irrespective of this issue, the notion of objectivity is quite different from that of determinism. The fact that I'm giving a lecture in room 144, Loomis Lab, at 2:15pm on October 26, 2017, is certainly an "objective" fact in at least the minimal sense that it could be verified by any competent observer. Whether this fact has been uniquely determined by events in my past and that of the University of Illinois is a quite different, and indeed unrelated, question. This distinction may seem totally trivial, but as we shall see, the lack of it has led on occasion to considerable confusion in the discussion of so-called "hidden variable" theories in quantum mechanics.

Classical physics is prima facie compatible with the idea of objectivity; e.g., there is no particular problem in assuming that particles not only exist but have been in definite positions and velocities even when unobserved. And yet... what about, say, the concept of the electric field, which has been defined in terms of what force *would* be experienced by a charged particle, were one there? (Worse: in special relativity the magnitude of the electric field depends on the state of motion of the observer!)

Probability in general and in classical physics

The idea of probability is a typical subject of philosophical inquiry, in the sense that in everyday life we (and insurance companies!) seem to have no difficulty in using it, and yet we may have severe difficulties in analyzing and justifying what we are doing when we use it (compare bicycle-riding!).

There is a well-developed formal theory of probability that seems to pose no particular problems: that is, we can define a mathematical structure that seems to correspond closely to what we do when we argue about probability (though who is to say that it could not correspond as well to something quite unrelated? – cf. the Peano axioms). Let's review some of its basic features (I deliberately eschew mathematical rigor here).

(1) We first introduce the idea of a "population" or "sample space" (or in physics, "ensemble"). This is a large collection (effectively infinite) of things/ people/ events/ etc., on which we make random measurements of some property (variable) A, which can take (for simplicity) only a set of discrete values A_i. If we are only interested in the math, the concept of "ensemble" can be taken as a primitive undefined concept, but it is interesting to think of possible realizations in the real world. For example, the sample space might be the successive throws of a die, in which case the quantity A would correspond to "heads or tails?"; or it might be the population of the U.S., in which case the quantity A might be, for example, height measured to the nearest inch, age to the nearest year, etc. Note that it is absolutely essential to the subsequent analysis that successive trials be genuinely random; for example, to take a "random" sample of the population of the U.S., it will almost certainly not be adequate to take the names on a particular page of the telephone directory for various towns, since there is a considerable higher-than-random probability that successive names belong to individuals who are related and whose height (e.g.) may therefore be correlated. Similarly, if one's coin tosses are performed on a tabletop covered with some sticky substance, the odds are that the outcome of one toss would influence the probabilities of the next. So while the concept of a randomly sampled population is a nice one in the context of pure mathematics, actually deciding which collections of people/ things/ events in the real world would satisfy the definition may not be at all trivial. This problem is of crucial importance in many arguments about those experiments on "paranormal"

phenomena that rely on statistical considerations; also, as we shall see, they are relevant to the use of probability in quantum mechanics.

(2) Having defined one's sample space and the quantity A one is interested in, one next assigns to each of the L possible values A_i of A, a (real) number $p(A_i)$ with the following properties:

(a) $0 \leq p(A_i) \leq 1$

(b) $\sum_{i=1}^{L} p(A_i) = 1$

The first condition expresses the fact that the probability of getting any particular A_i must be positive and cannot exceed 1; the second that we must get one of the values A_i. Actually, (b) can be regarded as a special case of a more general statement: let us consider the outcome "A_i or A_j" and assign a probability to that. Since the outcomes A_i and A_j are by hypothesis mutually exclusive, we have:

(c) $p(A_i \,\text{or}\, A_j) = p(A_i) + p(A_j)$

and since the probability of getting some result, which in our notation is $p(A_1 \text{ or } A_2 \ldots \text{or } A_L)$, must be unity, the result (b) then follows.

(3) Next we introduce the idea of joint probabilities for two different variables. Suppose that for members of the same ensemble A can take values A_i, $i = 1, 2 \ldots L$, and B, and B can take values B_j, $j = 1, 2 \ldots M$. Then we can define, analogously to the above, the quantity $p(A_i, B_j)$: The obvious generalizations of postulates (a to c) are

(a) $0 \leq p(A_i, B_j) \leq 1$

(b) $\sum_{i=1}^{L} \sum_{j=1}^{M} p(A_i, B_j) = 1$

(c) $p\left((A_i, B_j) \,\text{or}\, (A_k, B_l)\right) = p(A_i, B_j) + p(A_k, B_l)$

Note that (c) should apply whether or not (say) $k = i$ (as long as $j \neq l$). The distribution of two variables A and B in the ensemble is said to be independent (uncorrelated) if for all i and j

$$p(A_i, B_j) = p(A_i)p(B_j)$$

In the case when the ensemble is the adult population of the U.S., for example, two variables that are at least approximately uncorrelated are height and age; on the other hand, certain other pairs of variables such as height and sex are likely to show correlations.

(4) A very important concept is that of conditional probability. Very commonly, we are interested in cases where the drawing of samples might be random but we nevertheless have some additional information about the specimen drawn. Again,

if we are interested in the probability of an individual drawn at random from the U.S. population being over $6'$ in height, the information (e.g.) that that individual is female is clearly relevant. Thus we introduce the quantity $p(A_i|B_j)$, meaning the probability that a randomly selected member of the ensemble has property A_i given that we know it to have property B_j. It is clear that this idea is closely related to the joint probabilities $p(A_i, B_j)$ defined above: we evidently have

$$p(A_i, B_j) = p(B_j)p(A_i|B_j)$$

(or, of course, $p(A_i)p(B_j|A_i)$, and in the formal theory this can, if necessary, be used as the definition of $p(A_i|B_j)$).

Once we have these definitions and axioms, it is a matter of pure mathematics to establish a large number of interesting theorems. A typical (and very simple) example is Bayes' theorem, which states that

$$p(A_i|B_j) = p(A_i)p(B_j|A_i)/p(B_j)$$

One theorem which is particularly relevant for physics is the so-called law of large numbers (or central limit theorem). Suppose we take a given ensemble and measure a given variable A on a large number N of its members (as always, randomly selected). For these N members, define the average value of A in the obvious way:*

$$\overline{A} = \frac{1}{N}\sum_{k=1}^{N} a_k$$

What then is the most probable value of $\overline{A}$, and what are the chances of getting a result very different from this most probable value? The answer to the first question is intuitively rather obvious:† the most probable result for $\overline{A}$ is just the quantity $\sum_{i=1}^{L} A_i p(A_i)$, i.e., the "single-system average". The answer to the second question turns out to be (not obvious!) that the "distance away" of the value of $\overline{A}$ for which the probability sinks to (say) half of its value for the most probable A_i is proportional, for large N, to $N^{-1/2}$. Let's illustrate this by the simplest possible example, the tossing of a "fair" coin:

$$p(H) = p(T) = \frac{1}{2}$$

Start with the trivial case, $N = 1$; then there are only two possible outcomes, namely H or T and each occurs with probability $1/2$. Next, consider $N = 2$, i.e., two successive tosses. Here a crucial role is played by the assumption (axiom) that the probabilities in successive throws are independent; thus, for example, $p(H \,\text{on}\, 1, T \,\text{on}\, 2) = p(H)p(T) = \frac{1}{4}$. Thus, the four possible outcomes have equal probability:

$$p(H,H) = p(H,T) = p(T,H) = p(T,T)$$

*a_k is, of course, one of the possible values A_i $(i = 1, 2 \ldots L)$.

†Although the formal proof is not entirely trivial.

But what we are interested in is the probability of the different possible values for the fraction of heads (call it x). It is clear that this can, for $N = 2$, take only the values 0, $\frac{1}{2}$, and 1. Now 0 has to correspond to (T, T) and 1 to (H, H), but $\frac{1}{2}$ can correspond to either (H, T) or (T, H), and by axiom (c) the probability of getting one or the other of them is simply the sum of the probabilities of getting each separately, i.e., $\frac{1}{2}$. Thus we find

$$p(x = 0) = \frac{1}{4}, \quad p\left(x = \frac{1}{2}\right) = \frac{1}{2}, \quad p(x = 1) = \frac{1}{4}$$

and so on. It is intuitively clear that as N increases the distribution approximates to a continuous curve having a maximum at $x = \frac{1}{2}$ and a "width" which decreases with increasing N (actually, it turns out, as $N^{-1/2}$).

None of these results are at all controversial; when actually manipulating probabilities we seem to have no special difficulties (or rather there are prima facie difficulties and paradoxes, but there does always seem to be a unique "right" answer).

But, of course, the crunch is what "probability" means. The only thing that is universally agreed is that the concept applies only under conditions of uncertainty, i.e., either (a) the outcome of a trial is in principle uniquely determined but we don't have sufficient information to predict it, or possibly (b) even with the most complete possible information the outcome is not uniquely determined, as in Exner's hypothesis. Beyond this, there exist two main "schools" or "interpretations".

(1) "Frequentist" interpretation: probabilities are defined as the limiting frequency in an indefinitely large set of trials. Formally:

$$p_i = \lim_{N \to \infty} \frac{N_i}{N}$$

The formal axioms of probability theory then turn out to be simple mathematical statements about the properties of such limits. The principal problem with the frequentist definition is that it risks being circular. We have seen that all that the "law of large numbers" assures us is that the probability of getting a result, for the "fraction of events" giving A_i in a long sequence of trials, which is very different from p_i is extremely small and can be made as small as we like by taking $N \to \infty$. But it is still a probability! Suppose one is playing bridge and is dealt a hand containing, say, no card of value more than 6. If the dealing is indeed a random process, the probability of this result is exceedingly small. So there are two obvious possible reactions:

- (a) the dealer has cheated, or
- (b) we have indeed encountered a "very improbable" event. Without detailed assumptions about the honesty of the dealer, etc., it is impossible to exclude the latter explanation!

A different kind of objection to the "frequentist" approach is that it does not prima facie do justice to many cases in everyday life where we would like to use the concept of probability, e.g., where the weather service says that there is a 40% probability of rain tomorrow, or we say we believe there is a 70% probability that the Republican candidate will win the next presidential election. In the former case, we might perhaps try to justify a frequentist approach by claiming to have identified a set of variables (temperature, humidity, wind strengh...) that are relevant to weather forecasting, and then saying that "a 40% chance of rain tomorrow" means that if we look at the total ensemble of days on which those relevant variables had the values they have today, then in 40% of cases it rained on the next day. This is not totally implausible because at least there presumably exists a number of actual days in the records in which the values have been equal to (or at least close to) today's. However, in the second case it seems much less plausible; each presidential election is at least to some extent *sui generis*. (Similar considerations apply in a physical context, e.g., with regard to the probable consequences of extremely rare cosmic-ray events).

(For a more detailed discussion of possible variants of the frequentist (objectivist) theory of probability, see Sklar pp. 94-7).

(2) "Subjectivist" approach: this essentially holds that probability statements are statements about the degree of confidence of the speaker in different hypotheses. The formal laws of probability theory are then interpreted as *consistency conditions* between different beliefs, i.e., the conditions that will prevent one making, as it were, bets that are sure to lose. To take a trivial example, suppose we are tossing dice which we do not necessarily assume to be "fair" but do assume to constitute a random ensemble, in the sense that the results of different throws are uncorrelated. I offer you the choice of two schemes: in scheme *I*, I pay you a dollar every time two sixes come up on successive throws, and you pay me X dollars (X being some unspecified number) every time sixes come up on throws separated by one (irrespective of the result of the intermediate throw); in scheme *II*, our roles are reversed, i.e., in the former case, you pay me a dollar, etc. It is clear that if X is less than 1, then it is always to your advantage to choose scheme *I*, and if X is greater than 1, scheme *II*, and that in the long run you would expect in either case to make money from me. Thus, the only "consistent" choice for me to make is $X = 1$, which expresses my confidence that the probability of two sixes on throws separated by one is exactly the same as that of two sixes on successive throws. Note carefully that the above statements are only true "in the long run"; e.g., even if I choose $X < 1$ and you therefore opt for scheme I, there is a finite chance that in a run of finite length I will still make money from you, if e.g., the proportion of pairs of sixes on successive throws is less than expected statistically. However, in accordance with the law of large numbers discussed above, the probability of this happening diminishes rapidly as the length of the sequence increases. Note also that none of the above analysis requires us to know the actual probability of a occurrence of a six.

Thus, neither the objectivist (frequentist) nor the subjectivist interpretation has particular difficulty explaining why the formal axioms of probability theory should apply to it. However, one can ask a different question: Given a typical ensemble such as the throws of a die or the population of the U.S., what is the basis for the assignment of probabilities to it? For the frequentist, there is no major prima facie problem: you simply take a large sample and count! E.g., if you are interested in the height distribution of the U.S. population, you simply select, say, 10,000 or 100,000 individuals at random, determine their height distribution, and then take that distribution as describing the population as a whole. The argument implicit here is that, by the law of large numbers, the probability of the distribution in the sample being appreciably different from that of the total population is exceedingly small.

The subjectivist also needs to be able to assign "probabilities" absolutely and not just relatively (i.e., he not only needs to respect the formal axioms of probability, but also to actually assign, numbers for the p_i in a given ensemble). In fact, he needs to do so even to bet consistently in general; e.g., consider a trivial experiment: you pay me one dollar per throw irrespective of the result and I pay you y dollars every time a six comes up or vice versa. It is clear that the only "consistent" choice is to take $y = \frac{1}{p}$ where p is the probability of obtaining a 6. So how does the subjectivist actually estimate p? In real life, any sensible person would of course conduct a series of preliminary trials and take p to be, at least approximately, the fraction of sixes coming up; in other words, he would proceed exactly like the frequentist. A central question then is why it is "rational" to base one's expectations (or bets) on the observed frequency in a finite but large sample (something that insurance companies, of course, do all the time!) And this question actually arises for the frequentist, who, after all, in real life has to "place bets" just as often as his subjectivist colleague! It is perhaps the irreducible core of the philosophical problem of probability.

However, quite irrespective of this, it is clear that there are many situations where we would like to use the idea of probability where the whole idea of doing experiments on an existing "ensemble" is clearly inapplicable. One such is "one-off" events, like the next presidential election or the Big Bang (if it exists). Another relevant example arises in physics: the attempt to assign certain "a priori" probabilities, e.g.,, to molecular velocities. For example: suppose that we know that a given molecular group in a crystal has three different possible orientations, and we know nothing else about it; is there then a "rational" way of assigning probabilities to each of these three orientations, and if so what is it? It is tempting at this point to use the so-called Principle of Indifference, which says, in effect, that if we have several possibilities and there is no physical reason to distinguish them, then they must occur with equal probability. Thus, in the case in question, if the energy is the same for all three orientations, we would simply assign equal probabilities, i.e., 1/3 each. However, it turns out that there is a serious problem with the Principle of Indifference, in that it is often not at all clear what one should count as "different equivalent possibilities". Sklar gives the example of a jar in which the volume of water is not proportional to the area wetted: we can ask whether it is more reasonable to say that it is as likely to be more than half full or less, or that it is as likely that more than half the surface is wetted as that less is. In classical physics, a rather similar situation arises when we consider the possible velocities of the molecules of a gas: is it more reasonable to assign equal "weight" to equal "cells" of velocity or of angular momentum? It turns out that the physical predictions are quite different depending on which we choose. (Actually, in a sense, quantum mechanics comes to the rescue here, as we shall see).

In classical physics, the idea of probability is most commonly applied to the behavior of the individual microscopic constituents of matter (atoms and molecules); this application is the subject matter of the branch of physics known as "statistical mechanics", which we will have to return to in more detail later (lecture 24). We will see there that the frequentist/ objectivist debate is closely mimicked in arguments between proponents of the so-called "ergodic" and "information-theoretic" approaches to statistical mechanics. For the moment it is sufficient to note two features:

(1) In classical physics, it seems at least consistent, though perhaps not necessary, to assume that the need to apply probabilistic statistical ideas is merely a result of unavoidable ignorance due to the impossibility in practice of following the motion of every individual molecule. Thus, the probabilistic description that we use in practice is perfectly compatible with an underlying picture in which the motion of each individual particle is both objective (i.e., exists in the absence of observation) and deterministic. I.e., because of our ignorance, the motion, e.g., of a Brownian particle can look completely random, but there is no good reason for us to doubt that there exists in principle a well defined cause for it.

(2) By contrast, when we go to a much larger scale than Brownian particles, the randomness tends to cancel by the law of large numbers and the behavior does appear deterministic (e.g., the laws of Boyle and Charles relating the pressure and volume of a gas). The average force (i.e., pressure) has only very tiny fluctuations around the mean value.

Determinism, objectivity, and probability in quantum mechanics

In order to examine how these key concepts look in a quantum context, it is advisable to generalize the quantum mechanical description somewhat beyond the simple experiment illustrating "wave particle duality" so far discussed. In the following, I assume until further notice the standard "textbook" version of quantum mechanics.

(1) Quantum mechanics is a formalism used to describe ensembles of identically prepared systems. The question of how in practice we actually effect this "identical preparation" is actually quite a delicate one, but let's assume for present purposes that it has been solved.† E.g., in the Young's slits experiment, we would need (at least) a source of electrons (perhaps thermionic), a monochromator and a collimator.

†For simplicity, I am here assuming that there are no "avoidable" sources of ignorance; if there are, those need to be handled just as in classical physics, by ascribing (classical) probabilities to the various possibilities.

(2) Quantum mechanics describes the ensemble so prepared by a *wave function*, or more generally a state vector (the standard notation is $|\psi(t)\rangle$), which is a set of numbers ("probability amplitudes") given with respect to a certain basis. E.g., for the motion of a point particle, we usually write the wave function in the explicit form $|\psi(x,t)\rangle$ (position basis). For given values of x and t the corresponding *number* $\psi(x,t)$ is (proportional to) the probability amplitude to find the particle at position x at time t. The crucial point to note is that these numbers can be positive or negative (actually complex). *In the standard formulation no further knowledge about the systems of the ensemble beyond that embodied in the state vector is possible.*

(3) So long as no measurement of any physical quantity is made on the systems of the ensemble (and they are not otherwise disturbed by any external influence), the state vector (wavefunction) evolves smoothly and in a way which is completely determined by the physically relevant quantities such as the particle mass, the forces acting on it, etc. (Example: propagation of a wave packet.) Because it is "wave-like", it can show the interference effects characteristic of a wave.

(4) *Measurements* are made on *individual systems* of the ensemble, and for any given measured quantity, the possible results of the measurement, and the probability of getting each of them, are determined by the state vector. In very special cases, the measurement, and the state vector, may be such that one and only one result is seen on all measured systems. In that case, the state vector is said to be an *eigenstate* of the measured quantity. More generally, there will be various possible outcomes (not necessarily a continuous set), and these correspond to the *eigenvalues* of the measured quantity. In such a case, state vector determines the *probability* of getting each of the eigenvalues. The number ("probability amplitude") assigned by the state vector to a particular outcome is just the square root of the probability of getting that outcome. As an example, in the case of the continuous variable x (position), any value of x is an eigenvalue and given a particular wave function (state vector) $|\psi(x)\rangle$ at time t, the probability of measuring the position to be (at or near) x is proportional simply to $\psi^2(x)$. By contrast, a measurement of the angular momentum (e.g., of an electron around the nucleus in a hydrogen atom) can give only the discrete results $0, \hbar, 2\hbar \ldots$ (the eigenvalues), and the probability of getting each of these is given by $\psi(xyz)$, but in a way too complicated to go into here.

(5) Suppose we have drawn our system from an ensemble that at time t is described by a state vector $|\psi\rangle$ that is not an eigenstate of some quantity A. Suppose at time t, we make a measurement of A on some of the systems of the ensemble. As noted above, we will get the various possible eigenvalues A_i with some probability p_i. What is the correct description immediately after the measurement of the subensemble of systems which have given the result A_i? The standard textbook prescription ("projection postulate" or "collapse of the wave function") is that this subensemble is described by a new state vector $|\chi_i\rangle$, which is the eigenstate of A

corresponding to the eigenvalue A_i that was measured; i.e., for this subensemble we get a discontinuous "jump":

$$|\psi\rangle \to |\chi_i\rangle$$

(on measurement if outcome is A_i)

The necessity and indeed the meaningfulness of the projection postulate is still a matter of considerable debate.

It is impossible to overemphasize that measurements are made on *individual systems*, while (at least in the standard textbook formulation) the state vector describes an *ensemble*.

Determinism in quantum mechanics

It is possible to maintain that there are at least three senses in which QM preserves a concept of *determinism*:

(a) The *evolution* of the state vector (wave function) is completely deterministic (so long as no measurement is made).

(b) For an ensemble described by a given wave function $|\psi\rangle$, the *statistical distribution* of the results of any experiment is well-defined and determined by $|\psi\rangle$.

(c) "Many-body" determinism: in some cases, even when the origin of certain effects is quite manifestly quantum-mechanical, they may to all intents and purposes obey deterministic laws. The reason for this is essentially the same as applies in classical statistical mechanics: the law of large numbers evens out the fluctuations.

However, none of these considerations removes the fact that we can have all the knowledge of the state of a system that the conceptual structure of QM ever allows us to have and yet be quite unable to predict with certainty, even in principle, the results of a measurement performed on a particular member of that ensemble. Does this mean that the theory is "fundamentally" indeterministic? Not necessarily: after all, *predictability* is an epistemological concept and *determinism* an ontological one, and it might seem, prima facie, to be possible to maintain that QM is "in principle" deterministic (in the sense that the "true" initial state does indeed determine the subsequent behavior completely), but that it is impossible, within the framework of QM, ever to *know* that true initial state and thus to *predict* the subsequent behavior. Indeed, such a point of view seems to be essentially that embraced by Ernst Cassirer in his book *Determinism and Indeterminism in Modern Physics*. For reasons to be given in the next lecture, I believe this view is fundamentally mistaken.

Probability in quantum mechanics

Of the two "classic" approaches to probability theory, probably the one that comes more naturally to the practicing physicist is the "frequentist" one. If one takes this point of view for an ensemble of N systems prepared in a specified initial state $|\psi\rangle$, the probability of getting result i is defined by

$$p_i = \lim_{N\to\infty} \frac{N_i}{N}$$

where N_i is the number of systems yielding result i. Problems (reflecting general case, p. 2 above):

(1) There is no mathematical sense in which the limit exists.

(2) It is difficult to apply to cases where no "ensemble" exists, e.g., cosmic rays or fate of false vacuum.

So, one might try the subjectivist (information-theoretic) approach: the QM state vector (hence the probability based on it) is our best description of the system/ ensemble in the light of available evidence. This makes it easier to interpret the "projection postulate": when our knowledge of the system changes, the state vector changes with it, but this is not (in this interpretation) a physical change of the system any more than in classical probability theory. Obviously, the essential difference with classical physics is that in the latter case we can always, in principle, improve our information, but in QM, by hypothesis, no improvement beyond the information embodied in the state vector description is possible.

Objectivity in quantum mechanics

Perhaps this is the slipperiest question of all. Obviously, there are some quantities that are associated with a microsystem in a given state to which it is plausible to ascribe "objective" existence: first, "inalienable" properties such as its mass and charge, and second, perhaps those quantities of which the state vector is an eigenstate. Again, one could perhaps try to ascribe "objective" existence to the probability distribution (as properties of the *ensemble*), at least in the weak sense, since certainly these are intersubjectively verifiable. However, perhaps the two fundamental questions are:

(1) Is the state vector, which we recall gives the maximum possible information about the state of the system, itself an "objective" property?

(2) We know that each individual system of the ensemble, when subjected to a "measurement", will display a particular value of the measured variable, even though the state vector was not an eigenstate of the corresponding operator. After the measurement, therefore, this value seems to possess an "objective" existence. But did it exist before the measurement?

In regard to (1), there are difficulties with either answer:

(1) If the state vector is an objective property of the ensemble (rather than a function mainly of our information), then it seems difficult to avoid the conclusion that it is a property of the individual members of that ensemble; and in that case, one wonders how it can change discontinuously when a measurement is made, as most interpretations of the quantum formalism seem to require.

(2) If the state vector is *not* an objective property of the ensemble (or is a property of the ensemble, but not of individual members of it), but only a measure of our information concerning it, then it is difficult to see how it can give rise to physical effects (e.g., the interference pattern in a Young's slits experiment). Note that there are certain "effects" that are realized even for a single member of the ensemble, e.g., that it is never found at a point where destructive interference is complete. (An interesting attempt by Popper to "have his cake and eat it" is the "propensity" interpretation.)

These two difficulties point in the direction of the two towering conceptual issues in the fundamentals of QM: (2) points toward the "hidden-variable" question, while (1) indicates the absolute necessity of facing up to the nature and description of the measurement process. We will spend the remaining lectures on QM essentially on these two issues.