



Prof. Dr. Carl Friedrich von Weizsäcker (1992, Griesser Alm). © Lili Bansa who granted permission to use this photo

# Chapter 8

## Probability and Abstract Quantum Theory

### 8.1 Probability and Experience

In memoriam Imre Lakatos

This and the following section are from my essay “Probability and Quantum Mechanics“ (1973).<sup>1</sup> Imre Lakatos had read this work in the last years of his life, when he was a member of the Scientific Council of the Max Planck-Institute in Starnberg. He viewed it as an example of a ‘rational reconstruction’ and had it published in the *British Journal for the Philosophy of Science*, 24, 321–337. I dedicate these sections to his memory.

The theory of probability had its origin in an empirical question: Chevalier de Mere’s gambling problem. Equally, the present-day physicist finds no difficulty in empirically testing probabilities which have been theoretically predicted, by measuring the relative frequencies of the occurrence of certain events. On the other hand, the epistemological discussion on the meaning of the application of the so-called mathematical concept of probability is by no means settled. The battle is still raging between ‘objectivist,’ ‘subjectivist,’ and even other interpretations of probability. Probability is one of the outstanding examples of the “epistemological paradox” that we can successfully use our basic concepts without actually understanding them. In many apparent paradoxes associated with fundamental philosophical problems, the first step toward their solution consists in accepting the seemingly paradoxical situation as a phenomenon, and in this sense as a fact. Thus we must understand that it is the very nature of basic concepts to be practically useful without, or at least before, being analytically clarified. This clarification must use other concepts in an unanalysed manner. It may mean a step forward in such an analysis to see whether a hierarchy exists in the practical use of basic concepts, and which concepts then practically depend on the availability of which other concepts, and also to see where concepts interlink in a non-hierarchical manner. I will try to show that one of the traditional difficulties in the

---

<sup>1</sup> This chapter was first published in: *British Journal for the Philosophy of Science*, 24 (1973): 321–337; it was republished in: *The Structure of Physics* (Heidelberg: Springer, 2006), which is a translation of: *Aufbau der Physik* (Munich: Hanser, 1985): 100–111.

empirical interpretation of probability stems from the idea that experience can be treated as a given concept and probability as a concept to be applied to experience. This is what I call a mistaken epistemological hierarchy. I will try to point out that, on the contrary, experience and probability interlink in a manner that will preclude understanding experience without already using some concept of probability. I will offer a particular way to introduce probability in several steps.

We will interpret the concept of probability in a strictly empirical sense. We consider probability to be a measurable quantity whose value can be empirically tested, much like, for example, the value of an energy or temperature. What we need for defining a probability is an experimental situation in which different “events”  $E_1, E_2, \dots$  are the possible results of one experiment. We further need the possibility of meaningfully saying that an equivalent experimental situation (in short “the same situation”) prevails in different cases (in different “realizations,” “at different times,” “for different individual objects” etc.) and that, given this situation, an equivalent experiment (in short “the same experiment”) is carried out in each case. Let there be  $N$  performances of the same experiment, and assume that event  $E_k$  occurred  $n_k$  times. In this series of cases we call the fraction

$$f_k = \frac{n_k}{N}$$

the relative frequency with which event  $E_k$  occurred in the series. Now consider a future series of performances of the same experiment. Let us assume that our (theoretical and experimental) knowledge enables us to calculate a probability  $p_k$  of the event  $E_k$  in this experiment. Then we will take the meaning of this number  $p_k$  to be that it is a prediction of the relative frequency  $f_k$  for the future series of performances.<sup>2</sup> Finally,  $p_k$  will be empirically tested by comparing it with the values of  $f_k$  found in this and subsequent series of the experiment under consideration.

This is the simplistic view of the ordinary experimentalist. I think it is essentially correct and it will only need to be defended against the objections of the epistemologists. Of course we hope to understand it better by defending it.

Let us use a simple example in formulating the main objection. Our experiment will consist in the single cast of a die. There are six possible events. Let us choose the event that a “5” appears as the event of interest. Its probability  $p_5$  will be  $1/6$  if the die is “good.” Now let us cast the die  $N$  times. Even if  $N$  is divisible by 6, the fraction  $f_5$  will only rarely be exactly  $1/6$ , and, what is more important, the theory of probability does not expect  $f_5$  to be  $1/6$ . The theory predicts a distribution of the measured values of  $f_5$  in different series of casts around the theoretical probability  $p_5$ . The probability is only the *expectation value* of the relative frequency. But the concept “expectation value” is generally defined by making use of the concept “probability.” Hence it seems impossible to define probability by referring it to measurable relative frequencies, since that definition itself, if rigorously formulated,

---

<sup>2</sup> This formulation was proposed by Drieschner (1967).

would necessarily contain the concept of probability. It would, so it seems, be a circular definition.

We will not evade the problem by defining the probability as the limiting value of the relative frequency for long series, since there is no strict meaning to a limiting value in an *empirical* series which is essentially finite. These difficulties have induced some authors to abandon the ‘objectivist’ interpretation altogether in favour of a ‘subjectivist’ one which, e.g., reads the equation  $p_5 = 1/6$  as meaning: “I am ready to lay odds of 1–5 that a 5 will come up next time.” The theory of probability is then a theory of the consistency of a betting system. But this is not the problem of the physicist. He wishes to discover empirically who will become a rich man by his betting system. I am not going to enter into the discussion of these proposals,<sup>3</sup> but instead immediately offer my own.

The origin of the difficulty does not lie in the particular concept of probability but more generally in the idea of an empirical test of any theoretical prediction. Consider the measurement of a position coordinate  $x$  of a planet at a certain time; let its value predicted by the theory be  $\xi$ . A single measurement will give a value  $\xi_1$ , different from  $\xi$ . The single measurement may not suffice to convince us whether this result is to be considered a confirmation or refutation of the prediction. Thus we will repeat the measurement  $N$  times and apply the theory of errors. Let  $\bar{\xi}$  be the average of the measured values. Then, comparing the distance  $|\xi - \bar{\xi}|$  with the average scatter of the measured values, we can formally calculate a ‘probability’ with which the predicted value will differ from the ‘real’ value  $\xi_r$  (“ $\xi$  real”) by a quantity  $d = |\xi - \xi_r|$ . This ‘probability’ is itself a prediction of the relative frequency with which the measured distance  $|\xi - \bar{\xi}|$  will assume the value  $d$ , if we repeat the series of measurements many times. This structure of the empirical test of a theoretical prediction is slightly complicated, but well known. We can compress it into an abbreviated statement: “An empirical confirmation or refutation of any theoretical prediction is never possible with certainty but only with a greater or lesser degree of probability.” This is a fundamental feature of all experience. Here I am satisfied to describe it and to accept it; its philosophical relevance is to be discussed in another context.<sup>4</sup> Whoever works in an empirical science has already tacitly accepted it by his practice. In this sense the concept of scientific experience in practical use presupposes the applicability of some concept of probability, even if this concept is not explicitly articulated. Hence the very attempt to give a complete definition of probability by recurring on a given concept of experience is likely to result in a circular definition. Of course it would be equally impossible to define the concept of an empirical test by using a preconceived concept of probability. These two concepts, experience and probability, are not in a relationship of hierarchical subordination.

In practice every application of the theory of errors implies that we consider relative frequencies of events to be predictable quantities. In this sense probability

<sup>3</sup> Cf. Weizsäcker (1992): I 4.

<sup>4</sup> Cf. Weizsäcker (1992): I 3 and I 4.

is a measured quantity. This implies that our “abbreviated statement” also applies to probability itself: The empirical test of a theoretical probability is only possible with some degree of probability. The appearance of the probabilistic concept of an expectation value in the ‘definition’ of probability is therefore not a paradox but a necessary consequence of the nature of the concept of probability; or it is a ‘paradox’ inherent in the concept of experience itself. Still, probability is not on the same methodological level as all other empirical concepts. The precise measurement of any other quantity refers us to the measurement of relative frequencies, that is, to probabilities; the precise measurement of probability refers us to probabilities again. Due to this higher level of abstraction the predictions of the theory are better defined. The scatter of the measured values of any quantity about its average value depends on the nature of the measurement device; the scatter of relative frequencies about their expectation values is itself defined by the theory.

## 8.2 The Classical Concept of Probability

We have yet to achieve a definition of probability that can avoid the objection of being circular. I will now sketch a systematic theory of probability as an empirical concept, i.e., a concept of a quantity which can be empirically measured. This is not a rigorously developed classical theory of probability, but a sketch for an analysis of its concept of probability that emphasizes aspects of the theory where epistemological difficulties usually arise. I hope that this analysis will suffice for the construction of a consistent classical theory of probability, where for the mathematical details we might use any good textbook. The word ‘classical’ means here only “not yet quantum theoretical.”

This is done in three stages. We first formulate a “preliminary concept” of probability. It does not aim for precision; it is meant to describe in comprehensible terms how probability concepts are actually used in practice. Secondly we formulate a system of axioms of the *mathematical theory* of probability. In this section we can adopt Kolmogorov’s system. Thirdly we give empirical meaning, *physical semantics*, so to speak, to the concepts of the mathematical theory by identifying some of its concepts with some concepts associated with the preliminary concept of probability. This three-step procedure can also be described as a process of giving mathematical precision to the preliminary concept. The most important part of the third stage is a study of the consistency of the whole process. The interpreted theory of the third stage offers a mathematical model of those structures which were imprecisely described in the preliminary concept. I propose to call a theory *semantically consistent* if it permits one to use the preliminary concepts without which it would not have been given meaning in such a manner

that this use is correctly described by the mathematical model offered in the theory itself.<sup>5</sup> The *preliminary concept* is described by three postulates:

- A. A probability is a predicate of a formally possible future event or, more precisely, a modality of the proposition which asserts that this event will happen.
- B. If an event, or the corresponding proposition, has a probability very near to 1 or 0, it can be treated as practically necessary or practically impossible. A proposition (event) with a probability not very near to 0 is called possible.
- C. If we assign a probability  $p$  ( $0 < p < 1$ ) to a proposition or to the corresponding event, we thereby express the following expectation: out of a large number  $N$  of cases in which this probability is correctly assigned to this proposition there will be approximately  $n = pN$  cases in which the proposition will turn out to be true.

The language in which we formulated these postulates needs further explanation. We first see that restrictive concepts like ‘practically,’ ‘approximately,’ “expressing an expectation” are used. Their task is to indicate that our preliminary concept is not precise but should be made more precise. We will see that in this process these restrictive concepts will not be eliminated but be made more precise themselves. The word “correctly” in *C* indicates that we consider ascribing a probability to an event not an act of free subjective choice, but a scientific assertion subject to test.

The language of the postulates refers to the logic of temporal propositions. For propositions about the future this logic proposes not to use the traditional truth values ‘true’ and ‘false,’ but the “future modalities”: ‘possible,’ ‘necessary,’ ‘impossible.’ The postulate proposes to use probabilities as a more precise form of future modalities. With respect to the ordinary use of the word ‘probability’ this can be considered a terminological convention; further on, we wish to restrict the use of this word to statements about the future. But behind this convention lies the view that this is the primary meaning of probability and that other uses of the word can be reduced to it. For example, we apply it to the past in saying “it is probable that it was raining yesterday” or “the day before yesterday it was probable that it would rain the following day.” But in the second example probability is referred to what was then the future; characteristically we say here “it *was* probable.” In the first example we first of all admit lack of knowledge concerning the past; to make the statement operative we must apply it to the future in the sense “It is probable that, if I investigate, I will find out that it was raining yesterday.”

For the *mathematical theory* we can adopt Kolmogorov’s text literally, changing only some notation:

Let  $M$  be a set of elements  $\xi, \eta, \zeta \dots$  which we call ELEMENTARY EVENTS,  
and  $F$  a set of subsets of  $M$ ; the elements of  $F$  will be called EVENTS.

---

<sup>5</sup> Cf. v. Weizsäcker (2006), 9.2 and v. Weizsäcker (1992), 5.2.7.

- I.  $F$  is a [Boolean<sup>6</sup>] lattice of sets.
- II.  $F$  contains the set  $M$ .
- III. To every set  $A$  of  $F$  we assign a non-negative number  $p(A)$ . This number  $p(A)$  is called the probability of the event  $A$ .
- IV.  $p(M) = 1$
- V. If  $A_1$  and  $A_2$  are disjoint, then  $p(A_1 + A_2) = p(A_1) + p(A_2)$ .

We leave out axiom VI, which formulates a condition of continuity, since we will not discuss its problems here. We need, however, the definition of the expectation value:

Let there be a partition of the original set  $M$

$$M = A_1 + A_2 + \cdots + A_R,$$

and let  $x$  be a real function of the elementary event  $\xi$  which is equal to a constant  $A_q$  on every set  $A_q$ . Then we call  $x$  a STOCHASTIC QUANTITY and consider the sum

$$E(x) = \sum a_q p(A_q)$$

the mathematical expectation of the quantity  $x$ .

We now turn to the *physical semantics*. In order to simplify the expression and concentrate on the essentials, we assume the set  $M$  of elementary events to be finite. We call the number of elementary events  $K$ ; in the case of the die  $K = 6$ . We further consider a finite *ensemble* of  $N$  equivalent cases, e.g., of casts in the case of die. To every elementary event  $E_k$  (we write  $E_k$  instead of Kolmogorov's  $\xi$ ;  $1 < k < K$ ) we assign a number  $(nk)$  which indicates how many times this event  $E_k$  (say the "5" on the die) has actually happened in the particular series of  $N$  experiments which forms our given ensemble. Correspondingly we assign a  $n(A)$  to every event  $A$ . It is easy to see that the quantities

$$f(A) = \frac{n(A)}{N} \tag{8.1}$$

satisfy Kolmogorov's axioms I to V if we insert them for  $p(A)$ . This model of the axioms is, however, not the one intended by the theory of probabilities, but we reach our goal by adding a fourth postulate to the preliminary concept:

D. The probability of an event (of a proposition) is the expectation value of the relative frequency of its happening (its coming true).

The expectation value used in  $D$  is not defined on the original lattice of events  $F$ . It can be defined on a lattice  $G$  of 'meta-events.' We call a meta-event an ensemble of  $N$  events belonging to  $F$  which happen under equivalent conditions. We here use the language that the 'same' events can happen several times ("it has been raining and it will be raining again").  $G$  is not a subset of  $M$  or  $F$ , but it is a set of elements

---

<sup>6</sup> Added by the *ed.MD*.

of  $F$  with repetitions. Now we can assign a probability function  $p(A)$  to  $F$  (it may express our expectation of the events  $A$  according to the preliminary concepts). Then the rules of the mathematical theory of probability permit us to *calculate* a probability function for the elements of  $G$ ; it is only necessary to assume that the  $N$  events which together form a meta-event can be treated as independent. Assuming the validity of Kolmogorov's axioms for  $F$  we can then *prove* their validity for  $G$  and the validity of the formula

$$p(A) = E\left(\frac{n(A)}{N}\right). \quad (8.2)$$

We can now forget our preliminary ideas of the meaning of the  $p(A)$  in  $F$ . Instead we can apply the three postulates  $A, B, C$  to the lattice  $G$  of meta-events. After having thus given an interpretation (in the preliminary sense) to the  $p$  in  $G$ , we use (8.2) to deduce an interpretation of the  $p$  in  $F$ . It is exactly what postulate  $D$  says:  $p(A)$  is the expectation value of the relative frequency of  $A$ . If we now remember again how we would have interpreted the  $p$  in  $F$  without this construction, we would only have used  $A, B, C$ . This preliminary concept is now justified as a weaker formulation of  $D$ . The concepts 'practically,' 'approximately,' 'expectation' can now be more precisely interpreted by estimating probable errors. The mathematical "law of large numbers" proves that the expectation values of these errors tend to zero as  $N$  increases.

What have we gained epistemologically? We have not gotten rid of the imprecise preliminary concepts, we have merely transferred the lack of precision from events to meta-events, i.e., to large ensembles of events. The physical semantics for probabilities rest on the preliminary semantics for meta-probabilities. This is a more precise expression of our earlier statement, that a probability can only be empirically tested with some degree of probability. The solution of the paradox lies in its acceptance as a phenomenon; no theory of empirical probabilities can be meaningfully expected to yield more than this justification, which at least makes its consistency more evident.

If we wish, we can iterate our process and call this ladder of meta-probabilities a "recursive definition" of probability. While a typical recursive definition offers a fixed starting point ( $n = 1$ ) and a rule of recursion from  $n + 1$  to  $n$ , the recursion here can go as high as we like. At some step of the ladder we must halt and rely on the preliminary concepts. Due to the "law of large numbers" it will suffice for this highest step to postulate  $A$  and  $B$ . This will yield  $A, B,$  and  $C$  for the next lower step, and  $D$  for the ones below that.



### 8.3 Empirical Determination of Probabilities

We distinguish the *probability of an event* from the *probability of a rule*, but assume, however, (in contrast to Carnap's 1962<sup>7</sup>) that the two quantities are of exactly the same nature but at different levels of the application. The probability of an event  $x$  is the prediction (the expectation value) of the relative frequency  $f(x)$  with which an event of this type  $x$  will occur, upon frequent repetitions of the experiment in which  $x$  can happen. The content of a rule (an "empirical law of nature") is the specification of probabilities of events. Rules always specify *conditional* probabilities: "If  $y$  then  $x$  will occur with the probability  $p(x)$ ." But in the same way, by the conditional probability one also means the probability of an event. After all, the relative probability can only be measured if "the same" experiment is always performed, i.e., if equivalent conditions are realized. One can say that empirically testable probabilities are by nature conditional probabilities. The probability of a rule is now meant as the probability that the rule is 'true.' An empirical rule is true if it proves itself in experience. Its probability is then the prediction of the relative frequency with which just that rule  $R$  proves itself, upon frequent repetitions of the same empirical situation for testing that rule. This initially formal definition we have only to elucidate in detail, and we naturally arrive at an interpretation of Bayes' rule. What we are looking for can simply be called an iterated probability  $P(p(x))$ . In *Zeit und Wissen* II.4.4b<sup>8</sup> we will see that one does better to speak of a "higher-order probability"  $P(f(x))$ . For the present discussion such finesse does not matter.

One can (and in general will) approach the empirical determination of a probability from a starting point that expresses the *prior knowledge*. Methodologically we remember that an objective, empirically testable probability is at the same time related to the prior knowledge of a subject. As an example we choose two dice, each successively thrown once. Observers  $A$  and  $B$  are to state the probabilities for obtaining a 12,  $A$  before the double throw,  $B$ , however, after the first die had been cast.  $A$  gives  $p(12) = 1/36$ ,  $B$  on the average in one sixth of the cases  $p(12) = 1/6$ , on the average in five sixths of the cases  $p(12) = 0$ . Empirically both are correct as they refer to different statistical ensembles due to different prior knowledge.

The starting point for the desired rule expresses what one already knows before the test series. For simplicity let us first assume that one can describe the setup conceptually, but has never experimented with precisely this realization of the concepts. For example, one might be about to throw 'heads' or 'tails' with a coin or a "1...6" with a die, or to draw a ball from an urn containing  $w$  white and  $b$  black balls. Here is the legitimate application of Laplace's concept of *equal possibilities*, i.e., a *symmetry argument*. One knows which 'cases' are possible,

<sup>7</sup> Cf. Weizsäcker (1992): II.4.3.

<sup>8</sup> Cf. Weizsäcker (1992): II.4.4b.

i.e., one knows the catalog of all possible events. One does not know what would distinguish one of the elementary events (the atoms of the lattice of events) from any other. In this sense they are all equally possible. *Therefore* one assesses them to be all equally probable, i.e., one predicts an equal relative frequency for their occurrence. The empirically motivated assumption of symmetry is at this phase of the experiment essentially an expression of ignorance. This is the legitimate meaning of Laplace's approach, as subsequent investigation of the experiment will show.

At any rate, certain relative frequencies will be found in this experiment. Roughly speaking, we can distinguish three cases:

- (a) relative frequencies arise that are consistent with the starting point
- (b) relative frequencies arise which correspond to a different starting point
- (c) relative frequencies arise which do not correspond to any uniform statistical distribution

The phrase "are consistent with" means "in agreement with the expected distribution," within the limits of error the observer has set himself, using the calculus of probability. For the observer there is essentially no certainty, only a probability he can assign to each of the available scenarios; how he does this we will discuss shortly in more detail in connection with Bayes' problem. That the frequencies correspond at all to a unique starting point is by no means self-evident. This is indicated by the listing of case c. In this case one suspects that the catalog of events needs to be expanded to bring conditions into view which do not vary statistically but systematically. In view of these possibilities, cases *a* and *b* are hardly self-evident, and one could ask by what right they can be expected to occur at all. At the present stage of our epistemological examination we can only recognize Hume's problem in this difficulty and reply that according to our present understanding the occurrence of regular statistical distributions is a condition for the possibility of experience. At a later stage we will recognize in Laplace's symmetries basic symmetries of the world, namely in the equal probabilities of the sides of a uniform coin or a uniform die representations of the group of spatial rotations, realized in objects with negligible interaction with their environment, and a representation of the group of permutations of objects in the equal probabilities for picking any of the balls in the urn. There we must justify, through a discussion of the interaction, why the inclusion of new objects cannot remove the symmetry of the world under this group, such that every deviation of individual objects from the symmetry only stems from their individual interaction with other objects.

The classical model of Bayes' problem involves several urns (say 11) with different mixing ratios of black and white balls (say in the zeroth urn 0 white and 10 black balls, in the  $k$ th urn  $k$  white and  $10 - k$  black ones). In drawing from each urn we assume with Laplace equal probabilities; each of the urns is thus characterized by the probability  $p_w$  of drawing a white ball and  $p_b$  of drawing a black ball. According to our assumptions we have for the  $k$ th urn

$$p_w(k) = \frac{k}{10}, \quad (8.3)$$

and always

$$p_w + p_b = 1. \quad (8.4)$$

Now one picks one of the urns, without knowing which, and proceeds to draw and immediately return a single ball,  $n$  times in succession. If the outcome was  $n_w$  white and  $n_b$  black balls ( $n_w + n_b = n$ ), how probable is it that this had been the  $k$ th urn? In other words, one then determines a probability  $P_k$ . One can interpret  $P_k$  as a prediction of a relative frequency in a twofold way. On the one hand,  $P_k$  is, according to Laplace's assumption applied to the selection of one urn, the predicted relative frequency with which a particular urn turns out to be the  $k$ -th one *if* just  $n_w$  white and  $n_b$  black balls have been drawn from the urn. On the other hand,  $P_k$  enables one to predict new probabilities  $p_w'$  and  $p_b'$  for subsequent drawing from the urn, according to the formulas

$$p_w' = \sum_k P_k \cdot p_w(k), \quad (8.5)$$

$$p_w + p_b = 1. \quad (8.6)$$

Before the start of the experiment, according to Laplace's assumption, one would set each  $P_k$  to  $1/11$  for the selection of one urn, and compute from it the 'a priori probabilities'  $p_w^{(0)}$  and  $p_b^{(0)}$ , which in our case would both be  $1/2$ . The test series of drawing  $n$  balls is then the empirical determination of newer, i.e., 'a posteriori probabilities'. Bayes' procedure thus assigns to each of the 11 possible rules (8.3) a probability of the rule,  $P_k$ , and determines the probabilities  $p_w'$  proposed for the practical usage from the probabilities *according* to the rule ( $p_w(k)$ ) and the probability of *the rule*  $P_k$  according to (8.5).

Bayes' procedure thus corrects an initially assumed equipartition by means of insight into possible cases leading to different rules, for which again an equipartition is assumed. Naturally this can also be modified. One can introduce unequal a priori probabilities for picking one urn. This again can be reduced to an equipartition by assuming different numbers of urns of each type. The practical value of the procedure depends on the fact that for large values of  $n$  the influence of the assumed a priori probabilities gradually disappears. With an ontological assumption that all phenomena are built from equally possible elementary events one can thus even further justify the empirical determination of probabilities. Without such an assumption one can still describe the empirical determination "as if" such an assumption were justified; we need the assumption to *count* cases, and thereby are able to *define* in this way absolute, as well as relative, frequencies.

## 8.4 Reconstruction of Abstract Quantum Theory, Methodological Aspects

The title of this section initially suggests three questions:

1. What is meant by reconstruction?
2. What is meant by abstract quantum theory?
3. What ways are there for a reconstruction of abstract quantum theory?

### 8.4.1 *The Concept of Reconstruction*

I mean by reconstruction its retrospective derivation from the most plausible postulates. I articulate, as I have done before, the difference between two kinds of such postulates. They may either express conditions which make experience possible, thus conditions of human knowledge; then we call them *epistemic*. Or they formulate very simple principles which we hypothetically, inspired by concrete experience, want to assume as universally valid for the particular area of reality; we call these postulates *realistic*.

I emphasized in the first chapter of my ‘Structure of Physics’ that my method of a *Kreisgang*<sup>9</sup> does not permit a completely sharp distinction between these two kinds of postulates. I merge in the *Kreisgang* two traditions of thought which, in the history of philosophy, were in hostile opposition most of the time. All our knowledge of nature is subject to the conditions of human knowledge; that is the epistemological question. Humans are children of nature and their knowledge itself is a process in nature; that is the evolutionary question. Even our evolutionary knowledge is, as human knowledge, subject to the conditions of such knowledge, as studied in epistemology. The back of the mirror,<sup>10</sup> we only see in the mirror as well. But the mirror in which we see the back of the mirror is also just the mirror with this rear surface; epistemology, like the cognition it investigates, is also an event in nature. In this way every epistemological postulate is at the same time a statement about a process of nature, and every realistic postulate is formulated subject to the conditions of our knowledge.

The historical phenomenon that there are closed theories, however, permits us a relative distinction of epistemic and realistic postulates, as regards a particular

---

<sup>9</sup> Weizsäcker chose the word ‘Kreisgang’ to characterize his overall philosophical method. The term is difficult to translate (and is not a common German notion, either). It is used as a technical term throughout Weizsäcker 2006. In its literal meaning it refers to a ‘circular movement’ of knowledge and cognition. The largest circle possible is captured by Weizsäcker’s often used phrase: “Nature is older than humankind, humankind is older than natural science”, which should indicate that there is no hierarchy of cognition, but that it is necessary to circle back to former insights again and again, every time on a new level of knowledge. *Eds. of v. Weizsäcker (2006) and of this volume.*

<sup>10</sup> Lorenz (1978).

theory. “Only theory decides what can be measured” (Einstein to Heisenberg<sup>11</sup>). We will begin the reconstruction of quantum theory with *one* postulate which, for quantum theory, is epistemic: the existence of separable, empirically decidable *alternatives*. An alternative characterized in this way expresses the quantum theoretical concept of an observable, reduced to its logical foundation. The fact that quantum theory is so successful, and that one is able to succeed with the concept of the alternative in the totality of physical experience known to us, is an empirical fact which a priori does not appear to be certain. In this sense the postulate of alternatives is realistic, but it is also epistemic in another sense. First, as just noted, it is epistemic in the context of quantum theory: it formulates a condition without which the concepts of quantum theory are inapplicable. But second also as a matter of principle: we scarcely can imagine how scientific knowledge might be possible at all without separable, empirically decidable alternatives. The high degree of generality of quantum theory thereby confers upon its basic postulate a position reminiscent of Kant’s perception a priori: *that* experience is possible at all we cannot know a priori; we can only know what circumstances must obtain *in order for* experience to be possible.

However, the *second* central postulate for the actual quantum theory, which we call the postulate of *expansion* or *indeterminism*, must also be considered *realistic* in the context of quantum theory. We cannot imagine a theory of probability predictions about decidable alternatives in which this postulate is not applicable. This question, however, we can only discuss after the reconstruction is accomplished.

## 8.4.2 Abstract Quantum Theory

Terminologically we distinguish abstract and concrete quantum theory. One can characterize abstract quantum theory by means of four *theses*. We use the concept of ‘thesis’ to distinguish it from the reconstructive concept of ‘postulate.’ These could be at the foundation of a formally axiomatic deduction of the theory. But they cannot claim to be ‘evident’ as we require it of the postulates. Rather, their explanation is the goal of our reconstruction.

- A. *Hilbert space*. The states of every object are described by one-dimensional subspaces of a Hilbert space.
- B. *Probability metric*. The absolute square of the inner product of two normalized Hilbert vectors  $x$  and  $y$  is the conditional probability  $p(x, y)$  of finding the state belonging to  $y$ , if the state belonging to  $x$  is present.

---

<sup>11</sup> cf. Heisenberg (1971).

- C. *Composition rule.* Two coexisting objects  $A$  and  $B$  can be considered to be a composite object  $C = AB$ . The Hilbert space of  $C$  is the tensor product of the Hilbert spaces of  $A$  and  $B$ .
- D. *Dynamics.* Time is described by the real coordinate  $t$ . The states of an object are functions of  $t$ , described by a unitary mapping  $U(t)$  of the Hilbert space onto itself.

We call this theory abstract because it is universally valid for any object. One example of an abstract theory is classical point mechanics. There is an equation that characterizes the universally valid law of motion for arbitrary numbers  $n$  of mass points with arbitrary masses  $m_i$ , and arbitrary force laws  $f_{ik}(x_1 \dots x_n)$ . Von Neumann's quantum theory is even more abstract, as it does not presume the concept of a point mass and the existence of a three-dimensional configuration space. These concepts enter into quantum theory itself only via the special choices of the dynamics and the selection of certain observables associated with the dynamics. They belong to the *concrete* theory of specific objects.

## 8.5 Reconstruction via Probabilities and the Lattice of Propositions

This reconstruction path was chosen by Drieschner (1967) and described later (Drieschner 1979) in improved form. It follows most closely Jauch (1968) and the usual axiomatic theories; it goes beyond these in the way of the justification and the choice of postulates thereby implied. The reconstruction is sketched here to facilitate the connection to existing axiomatic quantum theories. This offers the opportunity to explain the abstract basic concepts within a familiar context.<sup>12</sup>

### 8.5.1 Alternatives and Probabilities

Physics formulates probability predictions about the outcome of future decisions of empirically decidable alternatives. The concept of probability is described in sections 1–3. Here, however, we will replace axiom I of Kolmogorov by another one; the catalogue of events is not the lattice of the subsets of a set.

We describe all possible observations as decision of  $n$ -fold alternatives. There  $n$  means either a natural number  $>2$  or the denumerable infinite. An  $n$ -fold alternative represents a set of  $n$  formally possible events which satisfy the following conditions:

---

<sup>12</sup> We follow the layout and sometimes the wording of the presentation in Weizsäcker (1980): II.5.4. *Addendum by the ed.MD*: That presentation is reprinted in this volume as “Quantum Theory”, part d, because it differs in many details from the present one.

1. The alternative is *decidable*, i.e., a situation can be created in which one of the possible events becomes an actual event and subsequently a fact. We then say that this event has happened.
2. If an event  $e_k$  ( $k = 1 \dots n$ ) has happened then none of the other events  $e_j$  ( $j \neq k$ ) has occurred. The results of an alternative are *mutually exclusive*.
3. If the alternative has been decided and all events except one, thus all  $e_j$  ( $j \neq k$ ), have not occurred, then the event  $e_k$  has happened. The alternative is defined as being *complete*.

*Note about the nomenclature.* Probabilities can be considered to be predicates of possible *events* or of *propositions*. About the philosophical interpretation of the difference between the two expressions see *Zeit und Wissen*<sup>13</sup> II 4. Here in this sketch we use both expressions indiscriminately; sometimes the one, sometimes the other is more convenient. This leads to the following expressions:

An alternative is a set of either *events* or of *propositions*. Both we call its *elements*. An event consists of the determination of a formally possible (conditional) *property* of an object at one time. Instead of this we also say that the object is at this time in a certain *state*. The word ‘state’ is in this context not restricted to “pure cases.” The *proposition* which asserts the existence of a property or state is formulated in the *present* tense. This means that one can often decide “the same” alternative. An alternative can also be referred to as a *question*; the propositions are then its possible *answers*.

### 8.5.2 Objects

The elements of an alternative consist of the determination of formally possible properties of an object at one time.

We introduce the ‘ontological’ concept of an object in addition to the ‘logical’ concept of an alternative. The alternatives for an object are, speaking quantum theoretically, its observables. We follow here the mode of thinking customary in all of physics, in particular in quantum theory, which interprets all its catalogues of propositions as propositions about, respectively, an ‘object’ or a ‘system.’ These two words are practically synonymous in contemporary physics. ‘Object’ is perhaps the more general concept as it encompasses composite as well as the possibly existing elementary objects, whereas the word ‘system’ is more indicative of compositeness (systema, standing together). In this chapter we will therefore in general choose the term ‘object.’

In the reconstruction we need the concept of the object to define the lattice of propositions which in each case is determined as the lattice of propositions about a fixed object (or the properties of a fixed object).

---

<sup>13</sup> v. Weizsäcker (1992).

The concept of the object, however, contains a fundamental problem which we will discuss in [Sect. 8.5.5](#).

### 8.5.3 *Ultimate Propositions About an Object*

For every object there ought to be ultimate propositions as well as alternatives whose elements are, logically speaking, ultimate propositions. As an ultimate (contingent) proposition about an object we define a proposition which is not implied by any other proposition about the same object.<sup>14</sup> In the quantum theoretical language this means that there are pure cases. Lattice-theoretically these ultimate propositions are ‘atoms,’ i.e., the lowest elements of the lattice; Drieschner (1979) therefore calls them atomic propositions. Drieschner argues for the postulate of the existence of atomic propositions from the requirement that it ought to be possible in principle to completely describe every object in terms of its properties.

### 8.5.4 *Finitism*

Drieschner (1967) introduced the postulate of finitism, which one might perhaps formulate thus: “The number of elements of an arbitrary alternative for a given object does not exceed a fixed positive number  $K$  which is characteristic of that object.” In contrast, we have also admitted denumerably infinite alternatives in 5.1. Furthermore, Drieschner (1979) no longer requires finitism. The technical benefit of the finitism postulate is that it avoids mathematical complications of an infinite-dimensional Hilbert space in the axiomatic reconstruction of quantum theory. Philosophically, behind this is the observation that no alternative with more than a finite number of elements can actually be decided by an experiment.

For convenience we will use here only finite alternatives. For physics, the infinite dimensions of Hilbert space become indispensable if we wish to unitarily represent in it the non-compact transformation groups of special relativity. In other words, we need it for relativistic quantum theory. In that regard, the present chapter is restricted to non-relativistic quantum theory. In [Chap. 4](#) of my (2006) I define the simplest objects, particles, as representations of relativistic transformation groups, following Wigner; thereby for every object  $K = \infty$ . The ‘objects’ of finitism, however, retain an assignable meaning as representations of the compact part of the group in finite-dimensional subspaces. We will then call them “sub-objects.”

---

<sup>14</sup> With the trivial exception of the “always false proposition”  $\wedge$ , which, by definition, implies any proposition—*ex falso quodlibet*.



### 8.5.5 *Composition of Alternatives and Objects*

Several alternatives can be combined to a *composite alternative*. This is done by “Cartesian multiplication.” Given  $N$  alternatives ( $N$  finite or perhaps denumerably infinite):  $\{e_k^\alpha\}$  ( $k = 1 \dots n^\alpha$ ;  $\alpha = 1 \dots N$ ), a combined event means that an event from each alternative occurs (not necessarily simultaneously). This is an element of the combined alternative which has  $n = \prod_\alpha n^\alpha$  elements.

Now  $N$  objects also define a total object of which they are parts. The Cartesian product of any alternatives of the parts is an alternative of the total object. In particular, the product of ultimate alternatives of the parts is a ultimate alternative of the total object.

The concept of an object, as we now see, contains some sort of self-contradiction which one cannot eliminate without eliminating all of physics known to us, which is built upon the concept of the object. Objects are known to us only through their interaction with other objects, ultimately with our own body. Completely isolated objects, free of any interaction, are no objects at all to us. The Hilbert space of an object describes just the possible states of only this one object. The introduction of dynamics, as we will perform it afterward, i.e., of a Hamiltonian operator, describes the influence of a fixed environment on the object and, insofar as one considers the object to be composite, the interaction of its parts with one another. To describe its influence on the environment one must combine it with other objects, thereby forming an aggregate object. In the Hilbert space of the aggregate object, however, the pure product states, in which the individual objects are in a definite state, are a set of measure zero. But it is just these definite states in terms of which quantum theory describes the individual objects. It appears that quantum theory could be formulated only approximately, which, if the theory is correct, would practically never be exactly valid. In short, the feasibility of theoretical physics rests upon its character as an approximation.

The philosophical problems herein I have discussed in detail in previous essays.<sup>15</sup> Let us accept here the concept of an object in its common usage.

### 8.5.6 *The Probability Function*

Between any two states  $a$  and  $b$  of the same object there is defined a probability function  $p(a, b)$  which gives the probability of finding  $b$  if  $a$  is necessary. The formulation and content of this postulate depended on the assumption that everything which can be said about an object in an empirically decidable way must be equivalent to the prediction of certain probabilities. The empirical verification of a proposition lies in the future at the time to which this proposition refers. About the future, however, only probabilities can be stated, which of course may

---

<sup>15</sup> Weizsäcker (1980): II.3.5, IV.6.4; Weizsäcker (1988): II.1.9.

approach the values 1 and 0, certainty and impossibility. The formulation of the condition in  $p(a, b)$  by means of “if  $a$  is necessary” includes the case “if  $a$  is present,” as  $a$  is then, due to the reproducibility, necessary in the future, as well as the case that one knows the necessity of  $a$  for other reasons.

The really strong assumption in this postulate remains inconspicuous in the above formulation, namely that this probability function assigns to each pair of states  $a$  and  $b$  the value  $p(a, b)$  independent of the state of the environment. This means at the same time that the states of an object admit an “internal description,” consisting only of its relative probabilities without reference to ‘external’ objects. How one can identify the respective states through observation, however, is then only determined in terms of the interaction of the object with its environment.

This strong assumption of independence is the form in which the *identity* of an object with itself expresses itself in this reconstruction, which ought to hold independently of its changing environment. Here is a specification of the concept of an object which we need for the reconstruction but which here we do not justify any further.

### 8.5.7 Objectivity

If a certain object actually exists, then an ultimate proposition about it is always necessary. This, too, is a strong statement. For its justification we refer to Drieschner (1979: 115–117). There it is described as being equivalent to the statement: “Every object has at any time as property a probability distribution of all its properties.” The premise “If a certain object actually exists” is necessary, because in states of composite objects which are not product states of the partial objects no ultimate statement about such a partial object is necessary. We then say that this partial object does not actually exist in such a state.

We call this postulated fact the *objectivity* of the properties of actually existing objects. For an actually existing object there is always an ultimate proposition, independently of whether we know it, i.e., it must inevitably be found if one looks for it. In other words, when one says that an object actually exists, one means that in principle one can know with certainty something about the object. Knowledge is not “merely a subjective state of the mind.” To know means, putting it tautologically, knowing that the known is as we know it. Here as well we refrain from following up on the philosophical implications of our assertion.

### 8.5.8 Indeterminism

To any two mutually exclusive ultimate propositions  $a_1$  and  $a_2$  about an object, there is an ultimate proposition  $b$  about the same object which does not exclude

either of the two. Two propositions  $x$  and  $y$  exclude one another if  $p(x, y) = p(y, x) = 0$ .

This is the central postulate of quantum theory. Following Drieschner it is called here the postulate of indeterminism. Within the context of the reconstruction it turns out to be equivalent to, e.g., the principle of superposition formulated by Jauch (1968: 106). It is the ‘realistic’ fundamental postulate; for it is at least not immediately obvious that experience without this postulate is not possible.

We can denote this postulate equivalently by the more abstract term *postulate of expansion*. The connection between the two names is as follows. Every alternative of ultimate propositions is expanded through this postulate by ultimate propositions about the same object which are not elements of the lattice of propositions which form the original alternative. The expansion is here formulated as a requirement on the probability function, i.e., on predictions: there are always predictions which have neither the value of certainty nor impossibility. This is juxtaposed with the postulate of objectivity according to which there are always necessary predictions. Both always exist. The requirement is at the same time formulated universally: it holds for *any* pair of mutually exclusive ultimate propositions. It implies that there can be no probability assignment of the catalogue of propositions about any object whatsoever for which every proposition is either true ( $p = 1$ ) or false ( $p = 0$ ). It thus implies the openness of the future as a matter of principle.

### 8.5.9 Sketch of a Reconstruction of Quantum Theory

For the implementation of the reconstruction we refer to Drieschner (1979). Here we merely mention the most important steps.

The catalogue of propositions is constructed about an object. Negation, disjunction, and implication are defined in terms of obvious requirements on the probability functions, such that the catalogue proves to be a lattice, and, in fact, for the case of finitism, a modular lattice. It can be shown that, with the imposed requirements, it is even a projective geometry. This can be represented as the lattice of the linear subspaces of a vector space. There remains the question of the field of numbers in which the vector space is erected. As a real metric is defined in it by means of the probability function, the field of numbers must contain the real numbers. Following Stückelberg (1960) Drieschner concludes from the uncertainty relations that it specifically must be the field of complex numbers. The dynamics is to be described in it, i.e., the time-dependence of the state, in terms of transformations under which the probability function remains invariant. These must be unitary transformations. In this fashion, abstract quantum theory is reconstructed.

For the time being, we forgo any attempt to examine how close the individual postulates have come to the ideal of an epistemic justification.

### 8.5.10 Historical Remark

The first formulation of the ideas utilized here, in my version, is given in the work *Komplementarität und Logik* (Weizsäcker 1955). To Drieschner's indeterminism axiom, there corresponds, for example, the "theorem of complementarity" (Sect. 6): "To every elementary proposition there are complementary elementary propositions." But only the work of Drieschner transformed this "complementary-logical" way of thinking, together with the axiomatic quantum theory of Jauch (1968), into a reconstruction of quantum theory. The goal of the present historical note is to point out the reconstruction of quantum theory previously begun by F. Bopp. Bopp's work of 1954 I quoted in 1955 (Sect. 5). It provided me with essential suggestions for the elaboration of my arguments at that time; see also his more recent work (Bopp 1983). Bopp begins, as we do in Sect. 4.1, with a simple alternative ("Sein oder Nichtsein als Grundlage der Quantenmechanik"). He postulates, as in Drieschner's uncertainty postulate, the existence of additional states defined in terms of relative probabilities and the continuity of this state space to make a continuous kinematics of the states possible. He, however, takes the spacetime continuum for granted and considers the alternative to depend on position ("ur fermion").

## References

- Bopp, F., 1954: *Zeitschrift für Naturforschung*, 9a: 579.
- Bopp, F., 1983: "Quantenphysikalischer Ursprung der Eichidee", in: *Annalen der Physik*, 7. Folge, 40: 317–333.
- Carnap, R., 1962: *Logical Foundations of Probability* (Chicago: Chicago UP).
- Drieschner, M., 1967: *Quantum Mechanics as a General Theory of Objective Prediction*. Univ. (Hamburg, Dissertation).
- Drieschner, M., 1979: *Voraussage—Wahrscheinlichkeit—Objekt. Über die begrifflichen Grundlagen der Quantenmechanik* (Berlin etc.: Springer).
- Heisenberg, W., 1971: *Physics and beyond; encounters and conversations* (New York: Harper & Row), Translation of: *Der Teil und das Ganze*.
- Jauch, J. M., 1968: *Foundations of Quantum Mechanics* (Reading, MA: Addison-Wesley).
- Lorenz, K., 1978: *Behind the Mirror* (San Diego, CA: Harcourt Brace), Translation of: *Die Rückseite des Spiegels*, literally, "The Back of the Mirror".
- Stueckelberg, E. C. G., 1960: "Quantum Theory in Real Hilbert Space", in: *Helv. Phys. Acta*, 33: 727–752.
- Weizsäcker, C. F. v., 1955: "Komplementarität und Logik", in: *Die Naturwissenschaften*, 42(19): 521–529; (20): 545–555.
- Weizsäcker, C. F. v., 1980: *The Unity of Nature* (New York: Farrar Straus Giroux), Translation of: *Die Einheit der Natur*.
- Weizsäcker, C. F. v., 1988: *The Ambivalence of Progress* (New York: Paragon House), Translation of: *Der Garten des Menschlichen*.
- Weizsäcker, C. F. v., 2006: *The Structure of Physics* (Heidelberg etc.: Springer), Translation of *Aufbau der Physik*, edited by Thomas Görnitz and Holger Lyre.
- Weizsäcker, C. F. v., 1992.: *Zeit und Wissen* (Munich etc.: Hanser).