

CHAPTER V

Probability and Induction

CHAPTER V

PROBABILITY AND INDUCTION

CHAPTER V

PROBABILITY AND INDUCTION

§ 32. The two forms of the concept of probability

The concept of probability has been represented in the preceding inquiries by the concept of weight. However, we did not make much use of this equivalence; we dealt with the concept of weight in an independent manner, not regarding the delimitations involved in its presumed equivalence to the concept of probability. We showed that there is such a concept of weight, that knowledge needs it in the sense of a predictional value, and that it is applied in everyday language as well as in scientific propositions—but we did not enter into an analysis of the concept, relying on a layman's understanding of what we meant by the term. We made use of the fact that the handling of a concept may precede an analysis of its structure. We constructed the triplet of predicates—meaning, truth-value, and weight—and found that it is the latter concept to which the others reduce. Truth has turned out to be nothing but a high weight and should not be considered as something other than an idealization approximately valid for certain practical purposes; meaning has been reduced to truth and weight by the verifiability theory—thus we found that the logical place for the concept of weight is at the very foundation of knowledge. It will now be our last task to enter into the analysis of this concept and to prove its equivalence to the probability concept; we may also hope to clarify its functions by their derivation from a concept as definitely determined as the concept of probability.

Turning to this task, we meet the fact that there are two different applications of the concept of probability, only one of which seems to be identical with the concept of weight as introduced by us. At the beginning of our inquiry into the nature of probability, we find ourselves confronted with the necessity of studying this distinction; we have to ask whether we are justified in speaking of only one concept of probability comprising both applications.

There is, first, the sharply determined concept of probability occurring in mathematics, mathematical physics, and all kinds of statistics. This mathematical concept of probability has become the object of a mathematical discipline, the calculus of probability; its qualities have been exactly formulated in mathematical language, and its application has found a detailed analysis in the well-known methods of mathematical statistics. Though this discipline is rather young, it has been developed to a high degree of perfection. This line of development starts with the inquiries of Pascal and Fermat into the theory of games of chance, runs through the fundamental works of Laplace and Gauss, and finds its continuation in our day in the comprehensive work of a great number of mathematicians. Any attempt at a theory of this mathematical concept of probability must start from its mathematical form. Mathematicians, therefore, have endeavored to clarify the foundations of the concept; among modern investigators of this subject, we may mention the names of v. Mises, Tornier, Dörge, Copeland, and Kolmogoroff.

There is, however, a second concept of probability which does not present itself in mathematical form. It is the concept which appears in conversation as "probably," "likely," "presumably"; its application is, however, not confined to colloquial language but is extended to scientific language also, where suppositions and conjectures cannot be

avoided. We pronounce scientific statements and scientific theories not with the claim of certainty but in the sense of probable, or highly probable, suppositions. The term "probable" occurring here is not submitted to statistical methods. This logical concept of probability, though indispensable for the construction of knowledge, has not found the exact determination which has been constructed for the mathematical concept. It is true that logicians of all times have considered this concept, from Aristotle to our day; thus the scientific treatment of this concept is much older than that of the mathematical concept which began with the investigations of Pascal and Fermat. But the theory of the logical concept of probability has not been able to attain the same degree of perfection as the theory of the mathematical concept of probability.

It was the great merit of the creators of logic that they contemplated, from the very beginning, a logic of probability which was to be as exact as the logic of truth. Leibnitz already had demanded "une nouvelle espèce de logique, qui traiterait des degrés de probabilité"; but this demand for a probability logic, like his project of a calculus of the logic of truth, was actualized only in the nineteenth century. After some attempts of De Morgan, it was Boole who developed the first complete calculus of a probability logic, which, in spite of some mistakes later corrected by Peirce, must be regarded as the greatest advance in the history of the logical concept of probability since Aristotle. It was a prophetic sign that the exposition of this probability logic was given in the same work which stands at the basis of the modern development of the logic of truth and falsity: in Boole's *Laws of Thought*. In the subsequent development, the problems of the logic of truth have assumed a much wider extent; probability logic was carried on by isolated authors only, among whom we may mention

Venn and Peirce, and among contemporary writers, Keynes, Lukasiewicz, and Zawirski.

If we regard these two lines of development, the supposition obtrudes that underlying them are two concepts which may show certain similarities and connections, but which are in their logical nature entirely disparate. This *disparity conception* of the two probability concepts has indeed been maintained by a great many authors, in the form either of a conscious or of a tacit assumption. On the other hand, the idea has been maintained that the apparent difference of the two concepts is only superficial, that a closer investigation reveals them as identical, and that only on the basis of an *identity conception* can a deeper understanding of the two probability concepts be obtained. The struggle between these two conceptions occupies to a great extent the philosophic discussion of the probability problem. The issue of this struggle is, indeed, of the greatest importance: since the theory of the mathematical concept of probability has been developed to a satisfactory solution, the identity conception leads to a solution of the philosophic probability problem as a whole, whereas the disparity conception leaves the problem of the logical concept of probability in a rather vague and unsatisfactory state. The latter consequence originates from the fact that a satisfactory theory of this concept, as different from the mathematical one, has not yet been presented.

The disparity conception has its genesis in the fact that the mathematical concept of probability is interpreted in terms of frequency, whereas the logical concept of probability seems to be of a quite different type.

Indeed, the great success of the mathematical theory of probability is due to the fact that it has been developed as a theory of relative frequencies. It is true that the original

definition of the degree of probability construed for an application in games of chance was not of the frequency type; Laplace gave the famous formulation of the ratio of the favorable cases to the possible cases, valid under the controversial presupposition of "equally possible" cases. This definition, apparently natural for cases of the type of the die, was abandoned however in all applications of the theory to cases of practical value: statisticians of all kinds did not ask for Laplace's "equally possible" cases but interpreted the numerical value of the probability by the ratio of two frequencies—the frequency of the events of the narrower class considered and the frequency of the events of the wider class to which the probability is referred. The mortality tables of life insurance companies are not based on assumptions of "equally possible" cases; the probabilities occurring there are calculated as fractions the numerator of which is given by the class of the cases of decease, and the denominator of which is determined by the class of the population to which the statistics are referred. The relative frequency thus obtained turned out to be an interpretation of the degree of probability much more useful than that of Laplace. The far-reaching extensions of the mathematical theory, indicated by such concepts as average, dispersion, average error, probability function, and Gaussian law, are due to a definitive abandonment of the Laplace definition and the transition to the frequency theory.

The logical concept of probability, on the contrary, seems to be independent of the frequency interpretation, which for many cases of logical probability appears not at all applicable. We ask for the probability of determinate events, say, of good weather tomorrow, or of Julius Caesar's having been in Britain; there is no statistical concept expressed in the question. It is the problem of the

probability of the single case which constitutes the origin of the disparity theory; authors such as Keynes,¹ therefore, base their concept of logical probability essentially on this problem.

Such authors even go so far as to deny a numerical value to logical probability. Keynes has developed the idea that logical probability is merely concerned with establishing an order, a series determined by the concepts of "more probable" and "less probable," in which metrical concepts such as "twice as probable" do not occur. These ideas have been continued by Popper.² For these authors, logical probability is a merely topological concept. Other authors do not want to admit such a restriction. Their concept of logical probability is metrical, but not of the frequency type. Logical probability, they say, is concerned with the "rational degree of expectation," a concept which already applies to a single event. It is here that the "equally possible" cases of Laplace find their field of application as furnishing the point of issue for the determination of the degree of expectation which a reasonable being should learn to put in place of feelings as unreasonable as hope and fear.

It will be our first task to enter into a discussion of these questions. We must decide in favor of either the disparity conception or the identity conception of the two forms of the probability concept.

§ 33. Disparity conception or identity conception?

The disparity conception is sometimes substantiated by saying that the mathematical concept of probability states a property of events, whereas the logical concept of probability states a property of propositions.

¹ J. M. Keynes, *A Treatise on Probability* (London, 1921).

² K. Popper, *Logik der Forschung* (Berlin, 1935).

If this were to be the whole content of the disparity conception, we would not attack it; for it is indeed possible to make such a distinction. If we interpret probability as a frequency of events, a probability statement would concern events; if we consider, on the contrary, probability as a generalization of truth, we have to conceive probability as concerning propositions. This is made necessary by the nature of the truth concept; only propositions, not things, can be called true, and our predicate of weight which we want to identify with probability has been introduced also as a predicate of propositions. But, if we apply these reflections to the probability concept, we find that they have only a formal signification and do not touch the central problem of the disparity conception. For, if we interpret the logical concept of probability also by a frequency, both concepts become isomorphic; the mathematical concept is then interpreted by a frequency of events, and the logical concept by a frequency of propositions about events.³ What the identity conception wants to maintain is just the applicability of the frequency interpretation to the logical concept of probability; thus we see that the thesis of the identity conception is, strictly speaking, an isomorphism of both concepts, or a structural identity. Even from the standpoint of the identity conception we may, therefore, consider the logical concept of probability as a concept of a higher linguistic level: such a distinction involves no difficulties for the theory of probability, as we are in any case obliged to introduce an infinite scale of probabilities of different logical levels (cf. § 41).

There is a second sense in which we have to speak of an identity here. If the frequency interpretation is accepted

³ This isomorphism follows strictly from the axiomatic construction of the calculus of probability which shows that all laws of probability can be deduced from the frequency interpretation (cf. § 37).

for the logical concept, this concept may be applied also to the statements of mathematical statistics: that is to say, even purely statistical statements admit both the mathematical and the logical conception of probability. A statement about the probability of death from tuberculosis may therefore be interpreted as concerning statistics of cases of tuberculosis, or as concerning statistics of propositions about cases of tuberculosis. On the other hand, the examples given for a logical meaning of the probability concept admit both interpretations as well.

For these reasons we shall use in the following inquiries the term "identity conception" without always mentioning that there is, strictly speaking, a difference of logical levels involved. We use the word "identity" here in the sense of an identity of structure, and our thesis amounts to maintaining *the applicability of the frequency interpretation to all concepts of probability*.

It is this thesis which the disparity conception attacks. We shall have to discuss this question now; if we cannot admit the disparity conception, this is because this conception involves consequences incompatible with the principles of empiricism.

There is, first, the principle of verifiability which cannot be carried through within the disparity conception. If a probability of a single event is admitted, in the sense of a prediction value—i.e., of signifying something concerning future events—there is no possibility of verifying the degree of probability by the observation of the future event in question. For instance we throw a die and expect with the probability $5/6$ to obtain a number greater than 1: how can this be verified if we watch one throw only? If the event expected does not occur, this is no refutation of the presumption because the probability $5/6$ does not exclude the case of the number 1 occurring. If the event expected

occurs, this is not a proof of the correctness of the presumption because the same might happen if the probability were $1/6$ only. We might at least say that the occurrence of the event is more compatible with the presumption than is the nonoccurrence. But how distinguish then between different degrees of probability both greater than one-half? If we had said that the probability of the event is not $5/6$ but $3/4$, how is the verification of this presumption to differ from that of the other?

The difficulty is not removed if we try to restrict probability statements to merely topological statements, eliminating the degree of probability. A statement of the form, "This event is more probable than the other," cannot be verified either, if it concerns a single case. Take two mutually exclusive events which are expected with the respective probabilities $1/6$ and $1/4$; the second one may happen. Is this a proof that this event was more probable than the other? This cannot be maintained because there is no principle that the more probable event must happen. The topological interpretation of logical probability is accordingly exposed to the same objections as the metrical one.

This analysis makes it obvious that a verification cannot be given if the probability statement is to concern a single case only. The single-case interpretation of the probability statement is not compatible with the verifiability theory of meaning because neither the degree nor the order asserted with the probability statement may be controlled if only one event is considered. One of the elementary principles of empiricism, therefore, is violated with this interpretation.

There is a second difficulty with the disparity conception, occurring if the degree of probability is to be quantitatively determined. We said that, if the frequency interpretation is denied, the concept "equally probable" de-

mands a substantiation by the concept of "equally possible cases," such as in Laplace's formulation. This leads, however, into apriorism. How do we know the "equal possibility"? Laplace's followers are obliged to admit here a kind of "synthetic a priori" judgment; the principle of "insufficient reason" or of "no reason to the contrary" does nothing but maintain this in a disguised form. This becomes obvious if we pass to a frequency statement, which in many cases, such as for dice, is attached to the "equal possibility" statement. How do we know that "equal possibility" implies equal frequency? We are forced to assume a correspondence of reason and reality, such as Kant had postulated.

We shall not enter here into a discussion of this second point, although it has played a great role in older philosophical discussions of the probability problem. We may only mention that the problem of the equally probable cases, such as occur in games of chance, finds a rather simple solution within the mathematical theory; no such presupposition as the principle of "no reason to the contrary" is needed there, and the whole question may be reduced to presuppositions such as occur within the frequency theory of probability.⁴ It is obvious that the question would not have assumed so much importance if the frequency theory of probability had been thoroughly accepted. The main point of difference in the discussion between the disparity conception and that of identity is to be sought in the problem of the interpretation of the single case. If it can be shown that the single-case interpretation is avoidable, and that the examples which seem to demand such an interpretation may be submitted to the frequency interpreta-

⁴ Cf. the report on this problem in the author's *Wahrscheinlichkeitslehre* (Leiden, 1935), § 65. For all other mathematical details omitted in the following inquiries we may also refer to this book.

tion, the superiority of the identity conception is demonstrated. To carry through this conception thus is identical with showing that the frequency interpretation of probability may always be applied. We shall inquire now whether this is possible.

For the frequency interpretation, a verification of the degree of the probability is possible as soon as the event can be repeated; the frequency observed in a series of events is considered as a control of the degree of probability. This interpretation presupposes, therefore, that the event is not described as an individual happening but as a member of a class; the "repetition" of the event means its inclusion within a class of similar events. In the case of the die, this class is easily constructed; it consists of the different throws of the die. But how construct this class in other examples, such as the case of a historical event of which we speak with a certain probability, or the case of the validity of a scientific theory which we assume not with certainty but only with more or less probability?

It is the view of the adherents of the identity conception that such a class may always be constructed and *must* be constructed if the probability statement is to have meaning. The origin of the single-case interpretation is to be found in the fact that for many cases the construction of the class is not so obviously determined as in the case of the die, or in the fact that ordinary language suppresses a reference to a class, and speaks incorrectly of a single event where a class of events should be considered. If we keep this postulate clearly in mind, we find that the way toward the construction of the corresponding class is indicated in the origin and use of probability statements. Why do we ascribe, say, a high probability to the statement that Napoleon had an attack of illness during the battle of Leipzig, and a smaller probability to the statement that Caspar

Hauser was the son of a prince? It is because chronicles of different types report these statements: one type is reliable because its statements, in frequent attempts at control, were confirmed; the other is not reliable because attempts at control frequently led to the refutation of the statement. The transition to the type of the chronicle indicates the class of the frequency interpretation; the probability occurring in the statements about Napoleon's disease, or Caspar Hauser's descent, is to be interpreted as concerning a certain class of historical reports and finds its statistical interpretation in the frequency of confirmations encountered within this class. Or take a statement such as pronounced by a physician, when he considers death in a certain case of tuberculosis highly probable: it is the frequency of death in the class of similar cases which is meant by the degree of probability occurring in the statement.

Although it cannot be denied that the corresponding class is easily determined in such cases, another objection may be raised against our interpretation of the probability statement. It is true, our opponents may argue, that the frequency within such a class is the *origin* of our probability statement; but does the statement *concern* this frequency? The physician will surely base his prediction of the death of his patient on statistics about tuberculosis; but does he mean such statistics when he talks of the determinate patient before him? The patient may be our intimate friend, it is his personal chance of death or life which we want to know; if the answer of the physician concerns a class of similar cases, this may be interesting for a statistician but not for us who want to know the fate of our friend. Perhaps he is just among the small percentage of cases of a happy issue admitted by the statistics; why should we believe in a high probability for his death be-

cause statistics about other people furnish such a high percentage?

It is the problem of the *applicability of the frequency interpretation to the single case* which is raised with this objection. This problem plays a great role in the defense of the disparity conception; it is said that the frequency theory may at best furnish a substantiation of the degree of probability but that it cannot be accepted as its interpretation as soon as the probability of a single case is demanded. The objection seems very convincing; I do not think, however, that it holds.

A clarification of the problem can only be given by an analysis of the situation in which we employ probability statements. Why do we ask for the probability of future events, or of past events about which we have no certain knowledge? We might be content with the simple statement that we do not know their truth-value—this attitude would have the advantage of not being exposed to logical criticism. If we do not agree with such a proposal it is because we cannot renounce a decision regarding the event at the moment we are faced with the necessity of acting. Actions demand a decision about unknown events; with our attempt to make this decision as favorable as possible the application of probability statements becomes unavoidable. This reflection determines the way in which the interpretation of probability statements is to be sought: *the meaning of probability statements is to be determined in such a way that our behavior in utilizing them for action can be justified.*

It is in this sense that the frequency interpretation of probability statements can be carried through even if it is the happening or not happening of a single event which is of concern to us. The preference of the more probable event is justified on the frequency interpretation by the

argument in terms of behavior most favorable on the whole: if we decide to assume the happening of the most probable event, we shall have in the long run the greatest number of successes. Thus although the individual event remains unknown, we do best to believe in the occurrence of the most probable event as determined by the frequency interpretation; in spite of possible failures, this principle will lead us to the best ratio of successes which is attainable.

Some examples may illustrate this point. If we are asked whether or not the side 1 of a die will appear in a throw, it is wiser to decide for "not-1" because, if the experiment is continued, in the long run we will have the greater number of successes. If we want to make an excursion tomorrow, and the weather forecast predicts bad weather, it is better not to go—not because the possibility of good weather is excluded but because, by applying the principle underlying this choice for all our excursions, we shall reduce the cases of bad weather to a minimum. If the physician tells us that our friend will probably die, we decide that it is better to believe him—not because it is impossible that our friend will survive his disease but because such a decision, repeatedly applied in similar cases, will spare us many disappointments.

It might be objected against the frequency interpretation that the principle of the greatest number of successes does not apply in cases in which only one member of the class concerned is ever realized. Throws of a die, or excursions, or cases of disease, are events which often recur; but how about other cases in which there is no repetition? This objection, however, conceives the class to be constructed too narrowly. We may incorporate events of very different types in one class, in the sense of the frequency interpretation, even if the degree of the probability changes from event to event. The calculus of probability has developed a type of probability series with changing

probabilities;⁵ for this type the frequency interpretation may also be carried through, the frequency being determined by the average of the probabilities which occur. Thus every action of our lives falls within a series of actions. If we consider the numerous actions of daily life which presuppose the probability concept—we press the electric button at the door because there is a probability that the bell will ring, we post a letter because there is a probability that it will arrive at the address indicated, we go to the tram station because there is a certain probability that the tram will come and take us, etc.—these actions combine to form a rather long series in which the frequency interpretation is applicable. The actions of greater importance may be included in another series, including those events which, in a narrower sense, are not repeated. The totality of our actions forms a rather extensive series which, if not submitted to the principle of assuming the most probable event, would lead to a remarkable diminution of successes.

We said that we do best to assume the most probable event; this needs a slight correction for cases in which different degrees of importance are attached to the cases open to our choice. If we are offered a wager in which the stakes are ten to one for the appearance of "number 1" and "some number other than 1" on the face of the die, of course it is more favorable to bet on "number 1." It is, however, again the frequency interpretation which justifies our bet; because of the terms of the wager, we will win more money in the long run by so betting. This case, therefore, is included in our principle of behavior most favorable on the whole. Instead of an amount of money, it may be the importance of an event which assumes a function analogous to that of the winnings in the game. If we expect the arrival of a friend with the prob-

⁵ Cf. *ibid.*, § 54.

ability of one-third, we had better go to the station to meet him. In this example, the inconvenience of our friend's arriving without our being at the station is so much greater than the inconvenience of our going there in vain that we prefer having the latter inconvenience in two-thirds of all cases to having the first inconvenience in one-third of all cases. Here it is again the frequency interpretation which justifies our behavior; if the probability of the arrival of our friend is one-hundredth only, we do not go to the station because our inconvenience in going ninety-nine times in vain to the station is greater than his inconvenience in arriving one time without our presence.

These considerations furnish a solution of the problem of the applicability of the frequency interpretation to the single case. Though the meaning of the probability statement is bound to a class of events, the statement is applicable for actions concerned with only a single event. The principle carried through in our foregoing investigations stating that there is as much meaning in propositions as is utilizable for actions, becomes directive once more and leads to a determination as to the meaning of probability statements. We need not introduce a "single-case meaning" of the probability statement; a "class meaning" is sufficient because it suffices to justify the application of probability statements to actions concerned with single events. The disparity conception of the two concepts of probability may be eliminated; the principle of the connection of meaning and action decides in favor of the identity conception.

§ 34. The concept of weight

With these considerations, the superiority of the identity conception is demonstrated in principle. But, to carry through the conception consistently, we are obliged to

enter into a further study of the logical position of statements about the single case.

If it is only the frequency of the class which is involved in the probability statement, the individual statement about the single case remains entirely indeterminate so long as it is not yet verified. We expect, say, the appearance of numbers other than 1 on the face of the die with the probability $5/6$; what does this mean for the individual throw before us? It does not mean: "It is true that a number other than 1 will appear"; and it does not mean: "It is false that a number other than 1 will appear." We must also add that it does not mean: "It is probable to the degree $5/6$ that a number other than 1 will appear"; for the term "probable" concerns the class only, not the individual event. We see that the individual statement is uttered as neither true, nor false, nor probable; in what sense, then is it uttered?

It is, we shall say, a *posit*.⁶ We posit the event to which the highest probability belongs as that event which will happen. We do not thereby say that we are convinced of its happening, that the proposition about its happening is true; we only decide to *deal with it* as a true proposition. The word "posit" may express this taking for true, without implying that there is any proof of the truth; the reason why we decide to take the proposition as true is that this decision leads, in repeated applications, to the greatest ratio of successes.

Our posit, however, may have good or bad qualities. If the probability belonging to it is great, it is good; in the contrary case it is bad. The occurrence of considerations of this type is best observed when we consider the gam-

⁶ The verb "to posit" has been occasionally used already; I shall venture to use it also as a noun by analogy with the corresponding use of the word "deposit."

bler. The gambler *lays a wager* on the event—this is his posit; he does not thereby ascribe a determinate truth-value to it—he says, however, that positing the event represents for him a determinate value. This value may even be expressed in terms of money—the amount of his stake indicates the value the posit possesses for him. If we analyze the way in which this value is appraised, we find that it contains two components: the first is the amount of money which the man would win if his wager were successful; the second is the probability of success. The arithmetical product of both components may be regarded, in correspondence with concepts in use within the calculus of probability, as the measure of the value the wager has for the gambler.⁷ We see that, within this determination of the value, the probability plays the role of a weight; the amount of the possible winnings is weighed in terms of the probability of success, and only the weighed amount determines the value. We may say: *A weight is what a degree of probability becomes if it is applied to a single case.*

This is the logical origin of the term “weight” which we used throughout the preceding inquiries. We understand now why the weight may be interpreted as the predictional value of the sentence; it is the predictional component of the whole value of the sentence which is measured by the weight. With this interpretation, the transition from the frequency theory to the single case is performed. The statement about a single case is not uttered by us with any pretense of its being a true statement; it is uttered in the form of a *posit*, or as we may also say—if we prefer an established word—in the form of a *wager*.⁸ The frequency

⁷ The occurrence of the arithmetical product here is due to the frequency interpretation. If the wager is frequently repeated, the product mentioned determines the total amount of money falling to the gambler's share.

⁸ The German word *Setzung* used in the author's *Wahrscheinlichkeitslehre* has both these significations.

within the corresponding class determines, for the single case, the weight of the posit or wager.

The case of the game may be considered as the paradigm of our position in the face of unknown events. Whenever a prediction is demanded, we face the future like a gambler; we cannot say anything about the truth or falsehood of the event in question—a posit concerning it, however, possesses a determinate weight for us, which may be expressed in a number. A man has an outstanding debt, but he does not know whether his debtor will ever meet his liability. If he wants money today, he may sell his claim for an amount determined by the probability of the debtor's paying; this probability, therefore, is a measure of the present value of the claim in relation to its absolute value and may be called the weight of the claim. We stand in a similar way before every future event, whether it is a job we are expecting to get, the result of a physical experiment, the sun's rising tomorrow, or the next world-war. All our posits concerning these events figure within our list of expectations with a predictional value, a weight, determined by their probability.

Any statement concerning the future is uttered in the sense of a wager. We wager on the sun's rising tomorrow, on there being food to nourish us tomorrow, on the validity of physical laws tomorrow; we are, all of us, gamblers—the man of science, and the business man, and the man who throws dice. Like the latter, we know the weights belonging to our wagers; and, if there is any difference in favor of the scientific gambler, it is only that he does not content himself with weights as low as accepted by the gambler with dice. That is the only difference; we cannot avoid laying wagers because this is the only way to take future events into account.

It is the desire for action which necessitates this gam-

bling. The passive man might sit and wait for what will happen. The active man who wants to determine his own future, to insure his food, and his dwelling, and the life of his family, and the success of his work, is obliged to be a gambler because logic offers him no better way to deal with the future. He may look for the best wagers attainable, i.e., the wagers with the greatest weights,⁹ and science will help him to find them. But logic cannot provide him with any guaranty of success.

There remain some objections against our theory of weights which we must now analyze.

The first objection concerns the definition of the weight belonging to the statement of a single event. If probability belongs to a class, its numerical value is determined because for a class of events a frequency of occurrence may be determined. A single event, however, belongs to many classes; which of the classes are we to choose as determining the weight? Suppose a man forty years old has tuberculosis; we want to know the probability of his death. Shall we consider for that purpose the frequency of death within the class of men forty years old, or within the class of tubercular people? And there are, of course, many other classes to which the man belongs.

The answer is, I think, obvious. We take the narrowest class for which we have reliable statistics. In our example, we should take the class of tubercular men of forty years

⁹ This remark needs some qualification. The wager with the greatest weight is not always our best wager, if the values, or gains, co-ordinated to events of different probabilities are different in a ratio which exceeds the inverse ratio of the probabilities, the best wager is that on the less probable event (cf. our remark at the end of § 33). Reflections of this type may determine our actions. If we call the wager with the highest weight our best wager, we mean to say "our best wager as far as predictions are concerned." We do not want to take into account in such utterances the value or relevance of the facts concerned. By the use of the word "posit" this ambiguity is avoided, as the term "best posit" is always to signify this narrower meaning.

of age. The narrower the class, the better the determination of the weight. This is to be justified by the frequency interpretation because the number of successful predictions will be the greatest if we choose the narrowest class attainable.¹⁰ A cautious physician will even place the man in question within a narrower class by making an X-ray; he will then use as the weight of the case, the probability of death belonging to a condition of the kind observed on the film. Only when the transition to a new class does not alter the probability may it be neglected; thus the class of persons whose name begins with the same letter as the name of the patient may be put aside.

It is the theory of the classical conception of causality that by including the single case into narrower and narrower classes the probability converges to 1 or to 0, i.e., the occurrence or nonoccurrence of the event is more and more closely determined. This idea has been rejected by quantum mechanics, which maintains that there is a limit to the probability attainable which cannot be exceeded, and that this limit is less than certainty. For practical life, this question has little importance, since we must stop in any case at a class relatively far from the limit. The weight we use, therefore, will not alone be determined by the event but also by the state of our knowledge. This result of our theory seems very natural, as our wagers cannot but depend on the state of our knowledge.¹¹

¹⁰ Imagine a class A within which an event of the type B is to be expected with the probability $1/2$; if we wager, then, always on B , we get 50 per cent successes. Now imagine the class A split into two classes, A_1 and A_2 ; in A_1 , B has a probability of $1/4$, in A_2 , B has a probability of $3/4$. We shall now lay different wagers according as the event of the type B belongs to A_1 , or A_2 ; in the first case, we wager always on non- B , in the second, on B . We shall then have 75 per cent successes (cf. the author's *Wahrscheinlichkeitslehre*, § 75).

¹¹ It has been objected against our theory that the probability not only depends on the class but also on the order in which the elements of the class are arranged. The latter is true, but it does not weaken our theory. First, it is an important feature of many statistical phenomena that the frequency structure

Another objection has its origin in the fact that in many cases we are not able to determine a numerical value of the weight. What is the probability that Caesar was in Britain, or that there will be a war next year? It is true that we cannot, for practical reasons, determine this probability; but I do not think that we are to infer from this fact that there is no probability determinable on principle. It is only a matter of the state of scientific knowledge whether there are statistical bases for the prediction of unknown events. We may well imagine methods of counting the success ratio belonging to the reports of historical chronicles of a certain type; and statistical information about wars in relation to sociological conditions is within the domain of scientific possibilities.

It has been argued that in such cases we know only a comparison of probabilities, a "more probable" and "less probable." We might say, perhaps, that this year a war is less probable than last year. This is not false; it is certainly easier to know determinations of a topological order than of a metrical character. The former, however, do not exclude the latter; there is no reason to assume that a metrical determination is impossible. On the contrary, the statistical method shows ways for finding such metrical determinations; it is only a technical matter whether or not we can carry it through.

There are a great many germs of a metrical determination of weights contained in the habits of business and daily life. The habit of betting on almost every thing unknown

is independent, to a great extent, of changes in the order. Second, if the order is relevant for the determination of the weight, it is to be included in the prescription; such is the case for contagious diseases (where the probability of an illness occurring depends on the illness or lack of illness of the persons in the environment), or for diseases having a tendency to repeat (where the probability changes if the illness has once occurred), etc. The mathematical theory of probabilities has developed methods for such cases. They do not imply any practical difficulty as to the definition of the weight.

but interesting to us shows that the man of practical life knows more about weights than many philosophers will admit. He has developed a method of instinctive appraisal which may be compared to the appraisal of a good contractor concerning the funds needed in opening a new factory, or to the appraisal by an artillery officer of spatial distances. In both cases, the exact determination by quantitative methods is not excluded; the instinctive appraisal may be, however, a good substitute for it. The man who bets on the outcome of a boxing match, or a horse race, or a scientific investigation, or an explorer's voyage, makes use of such instinctive appraisals of the weight; the height of his stakes indicates the weight appraised. The system of weights underlying all our actions does not possess the elaborate form of the mortality tables of insurance companies; however, it shows metrical features as well as topological ones, and there is good reason to assume that it may be developed to greater exactness by statistical methods.

§ 35. Probability logic

The logical conception considers probability as a generalization of truth; its rules must be developed, therefore, in the form of a logical system. It is this probability logic which we shall now construct.

Let us assume a class of given symbols a, b, c, \dots ; they may be propositions, or something similar to them—this may be left open for the present. To every symbol there is co-ordinated a number, the value of which varies between 0 and 1; we call it the probability belonging to the symbol and denote it by

$$P(a)$$

E.g., we may have

$$P(a) = \frac{1}{6}$$

In addition, we have logical symbols at our disposal, such as the signs \neg for "not," \vee for "or," a period (.) for "and," \supset for "implies," and \equiv for "is equivalent to." Performing with these signs operations based on the postulate that $P(a)$ is to assume functions similar to those of truth and falsehood in ordinary logic, we obtain a kind of logic which we shall call *probability logic*. As there is no further determination of the term "probability" as it here occurs, probability logic is a formal system, to which we may later give interpretations.

How we are to develop this formal system is not, logically speaking, sufficiently determined. We might invent any system of rules whatever and call it probability logic. This is the reason why the problem of probability logic, and the related problem of a logic of modality, have recently occasioned lively discussion; we have been presented with a great number of ingenious systems, especially in the case of the logic of modality, the advantages of each being emphasized by their various authors. I do not think, however, that the question is to be decided by logical elegance, or by other logical advantages of the proposed systems. The logic we seek is to correspond to the practice of science; and as science has developed the qualities of the probability concept in a very determinate way, there is, practically speaking, no choice left for us. This means that the laws of probability logic must be conformable to the laws of the mathematical calculus of probability; by this relation the structure of probability logic is fully determined. A similar remark applies to the logic of modality; the concepts of "possibility," "necessity," and the like, considered here are used in practice as a topological frame of the probability concept; therefore their structure is to be formulated in systems deducible from the general system of probability logic. The construction of this system by means of a de-

duction from the rules of the mathematical calculus of probability is, therefore, the fundamental problem of the whole domain. This construction has been carried out; however, we cannot present it in detail but must confine ourselves to a report of the results.¹²

The rules occurring in probability logic resemble the rules of ordinary or alternative logic (the term "two-valued logic" is in use also). However, there are two decisive differences.

The first is that the "truth-value" of the symbols a, b, c, \dots , is not bound to the two values "truth" and "falsehood," which may be denoted by 1 and 0, but varies continuously within the whole interval from 0 to 1.

The second is a difference concerning the rules. In the alternative logic, the truth-value of a combination $a \vee b$, or $a \cdot b$, etc., is determined if the truth-values of a and b are given individually. If we know that a is true and b is true, then we know that $a \cdot b$ is true; or, if we know that a is true and b is false, we know that $a \vee b$ is true, whereas $a \cdot b$ in this case would be false. Such a rule does not hold for probability logic. We cannot enter here into a detailed substantiation of this statement; we can only summarize the results obtained.¹³ It turns out that the "truth-value" of a combination of a and b is determined only if, in addition to the "truth-values" of a and b separately, the "truth-value" of one of the other combinations is given. That is

¹² For a detailed exposition cf. the author's article, "Wahrscheinlichkeitslogik," *Berichte der Berliner Akademie der Wissenschaften* (math.-phys. Kl., 1932); and the author's book *Wahrscheinlichkeitstlehre*. As to other publications of the author cf. chap. i, n. 14. For a summary of all contributions to the problem cf. Z. Zawirski, "Über das Verhältnis der mehrwertigen Logik zur Wahrscheinlichkeitstlogik," *Studia philosophica*, I (Warsaw, 1935), 407.

¹³ Cf. the author's *Wahrscheinlichkeitstlehre*, § 73. Instead of making the "truth-value" of a combination dependent on that of another combination, we may introduce as a third independent parameter the "probability of b relative to a " which we write $P(a, b)$. This is the way followed in *Wahrscheinlichkeitstlehre*. Both ways amount to the same.

to say: if $P(a)$ and $P(b)$ are given, the value of $P(a \vee b)$, or of $P(a \cdot b)$, and so on, is not determined; there may be cases in which $P(a)$ and $P(b)$ are, respectively, equal, whereas $P(a \vee b)$ and $P(a \cdot b)$ are different. If, however, the "truth-value" of *one* of the combinations is known, those of the others may be calculated. We may, e.g., introduce $P(a \cdot b)$ as a third independent parameter and then determine the "truth-values" of the other combinations as a function of $P(a)$, $P(b)$, and $P(a \cdot b)$. We have, for instance, the formula

$$P(a \vee b) = P(a) + P(b) - P(a \cdot b) \quad (1)$$

The necessity of a third parameter for the determination of the "truth-value" of the combinations distinguishes probability logic from alternative logic; it cannot be eliminated but originates from a corresponding indeterminacy in the mathematical calculus. If a and b mean the sides 1 and 2 of the same die, we have

$$P(a \cdot b) = 0$$

because the sides cannot occur together; the probability of the disjunction then becomes $2/6$, which follows from

$$P(a) = P(b) = \frac{1}{6}$$

and our formula (1). If on the contrary a and b mean the sides numbered 1 on *two* dice which are thrown together, we have on account of the independence of the throws¹⁴

$$P(a \cdot b) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

and our formula (1) furnishes $11/36$ for the probability of the disjunction, in correspondence with well-known rules of the calculus of probability.

¹⁴ We may note that our general formulas are not restricted to the case of independent events but apply to any events whatever.

A similar formula is developed for implication. It is shown to be

$$P(a \supset b) = 1 - P(a) + P(a \cdot b) \quad (2)$$

This case differs from the case of disjunction in so far as two indications, the probability of a and that of the product $a \cdot b$, suffice to determine the probability of the implication; the latter probability turns out to be independent of the probability of b . We cannot, however, replace the indication of $P(a \cdot b)$ by that of $P(b)$; this would leave the probability of the implication indeterminate.

For equivalence the equation is

$$P(a \equiv b) = 1 - P(a) - P(b) + 2P(a \cdot b) \quad (3)$$

In this case, the three probabilities $P(a)$, $P(b)$, and $P(a \cdot b)$ are again needed for the determination of the probability of the term on the left-hand side of the equivalence.

Only for the negation \bar{a} does a formula similar to that of alternative logic obtain:

$$P(\bar{a}) = 1 - P(a) \quad (4)$$

The probability of a suffices to determine that of \bar{a} .

These formulas indicate a logical structure more general than that of the two-valued logic; they contain this, however, as a special case. This is easily seen: if we restrict the numerical value of $P(a)$ and $P(b)$ to the numbers 1 and 0, the formulas (1)–(4) furnish automatically the well-known relations of two-valued logic, such as are expressed in the truth-tables of logistic; we have only to add the two-valued truth-table for the logical product $a \cdot b$, which, in the alternative logic, is not independently given but is a function of $P(a)$ and $P(b)$.¹⁵

¹⁵ It may be shown that for the special case of truth-values restricted to 0 and 1, the truth-value of the logical product is no longer arbitrary but determined by other rules of probability logic (cf. *Wahrscheinlichkeitslehre*, § 73).

These brief remarks may suffice to indicate the nature of probability logic; this logic turns out to be a generalization of the two-valued logic, since it is applicable in case the arguments form a continuous scale of truth-values. Let us turn now to the question as to the interpretation of the formal system.

If we understand by a, b, c, \dots , propositions, our probability logic becomes identical with the system of weights which we explained and made use of in our previous inquiries. We shall speak in this interpretation of the *logic of weights*.

However, we may give another interpretation to the symbols a, b, c, \dots . We may understand by the symbol a not one proposition but a series of propositions defined in a special manner. Let us consider a propositional function such as " x_i is a die showing 'side 1'"; the different throws of the die, numbered by the index i , then furnish a series of propositions which are sometimes true, sometimes false, but which are all derived from the same propositional function. We shall speak here of a *propositional series* (a_i). The parentheses are to indicate that we mean the whole series formed by the individual propositions a_i . Or take the propositional function: " x_i is a case of tuberculosis with lethal issue"; it will be sometimes true, sometimes false, if x_i runs through all the domain of tubercular people. If we substitute the symbols $(a_i), (b_i), \dots$, in our formulas, we may interpret $P(a_i), P(b_i), \dots$, as the limits of the frequencies with which a proposition is true in the propositional series. As to the logical operations, we add the definitions

$$\left. \begin{aligned} [(a_i) \vee (b_i)] &\equiv (a_i \vee b_i) \\ [(a_i) \cdot (b_i)] &\equiv (a_i \cdot b_i) \\ [(a_i) \supset (b_i)] &\equiv (a_i \supset b_i) \\ [(a_i) \equiv (b_i)] &\equiv (a_i \equiv b_i) \end{aligned} \right\} (5)$$

which postulate that a logical operation between two propositional series is equivalent to the aggregate of these logical operations between the elements of the propositional series. Our system of formulas then furnishes the laws of probability according to the frequency interpretation. We shall speak, in this case, of the *logic of propositional series*. We see that by these two interpretations the logical conception of probability splits into two subspecies. Probability logic is, formally speaking, a structure of linguistic elements; but we obtain two interpretations of this structure by different interpretations of these elements. If we conceive propositions as elements of this structure, and their weights as their "truth-values," we obtain the *logic of weights*. If we conceive propositional series as elements of the logical structure and the limits of their frequencies as their "truth-values," we obtain the *logic of propositional series*.

We explained above that the identity conception maintains the structural identity of the logical and the mathematical concept of probability; we can proceed now to another form of this thesis. Our logic of weights is the probability logic of propositions; it formulates the rules of what the adherents of the disparity conception would call the logical concept of probability. On the other hand, our logic of propositional series formulates the logical equivalent of the mathematical conception of probability, i.e., a logical system based on the frequency interpretation. What the identity conception maintains is the *identity of both these logical systems*; i.e., first, their structural identity, and, second, the thesis that the concept of weight has no other meaning than can be expressed in frequency statements. The concept of weight is, so to say, a fictional property of propositions which we use as an abbreviation for frequency statements. This amounts to saying that

every weight may be conceived, in principle, as determined by a frequency; and that, inversely, every frequency occurring in statistics may be conceived as a weight. If the adherents of the disparity conception will not admit this, it is because in certain cases they see only the weight form of probability and, in others, only the frequency form. There are, however, both forms in every case. In cases such as historical events these philosophers regard only the weight function of probability and do not consider the possibility of constructing a series in which the weight is determined by a frequency. In cases such as the game of dice, or social statistics, these philosophers see only the frequency interpretation of probability and do not observe that the probability thus obtained may be conceived as a weight for every single event of the statistical series. One throw of the die is an individual event in the same sense as Julius Caesar's stay in Britain; both may be incorporated in the logic of weights—but that does not preclude the weight's being determined by a frequency. The statistics necessary for this determination are easily obtained for the die but are very difficult to obtain in the case of Caesar's stay in Britain. We must content ourselves in this case with crude appraisals; but this does not prove an essential disparity of the two cases.

§ 36. The two ways of transforming probability logic into two-valued logic

We must now raise the question as to the transformation of probability logic into alternative logic. By the word "transformation" we do not mean a transition of the type indicated before. The transition by restriction of the domain of variables is a specialization; whether it applies depends on the nature of the variables given. We seek now for a transition which may be carried through for any kind

of variables, and which transforms any system of probability logic into two-valued logic.

There are two ways of effecting such a transformation. The first is the method of *division*. In its simplest form, the division is a *dichotomy*. We then cut the scale of probability into two parts by a demarcation value p , for instance, the value $p = \frac{1}{2}$, and make the following definitions:

If $P(a) > p$, a is called true
 If $P(a) \leq p$, a is called false

This procedure furnishes a rather crude classification of probability statements, but it is always applicable and suffices for certain practical purposes.

A more appropriate method of division introduces a three-valued logic. We proceed then by a *trichotomy*; we choose two demarcation values, p_1 and p_2 , and define:

If $P(a) \geq p_2$, a is called true
 If $P(a) \leq p_1$, a is called false
 If $p_1 < P(a) < p_2$, a is called indeterminate

If we choose for p_2 a value near 1 and for p_1 a value near 0, the trichotomy method has the advantage that only high probabilities are regarded as truth and only low probabilities as falsehood. As to the intermediate domain of the indeterminate, the procedure corresponds to actual practice: there are many statements which we cannot utilize because their truth-value is unknown. If we drop these indeterminate statements, we may regard the rest as statements of a two-valued logic; in this sense the method of trichotomy also leads to a two-valued logic.

As to the validity of the rules of the two-valued logic for the propositions defined as "true" or "false" by dichotomy or trichotomy, the following remark is to be added.

The operation of negation applies for dichotomy because it leads from one domain into the other on account of the relation expressed in (4), § 35. The same is valid for trichotomy if the limits p_1 and p_2 are situated symmetrically; on account of (4), § 35, the negation of a true statement is then false, and conversely. In the case of the other operations, however, the application of the rules of two-valued logic is permissible only in the sense of an approximation. If, for instance, according to our definitions, a is true, and b is true, we may not always regard the logical product $a \cdot b$ as also true, for there are certain exceptions. This is the case when $P(a)$ and $P(b)$ are near the limit p_1 or p_2 ; it may happen then that $P(a \cdot b)$ is below the limit. Thus if a and b are independent, the value of $P(a \cdot b)$ is given by the arithmetical product of $P(a)$ and $P(b)$; as these numbers are fractions below 1, their product may lie below the limit, whereas each of them lies above the limit. A similar case is possible for disjunction. In general, if a is false, and b is false, their disjunction $a \vee b$ is false also; it may happen however in our derived logic that in such a case the disjunction is true. This possibility is involved in our formula (1), § 35; if $P(a)$ and $P(b)$ lie below the limit, $P(a \vee b)$ may lie above the limit.

The two-valued logic derived from probability logic by dichotomy is seen to be an approximative logic only. The same is valid for the two-valued or three-valued logic derived by trichotomy. The latter becomes a strict logic only if $p_1 = 0$ and $p_2 = 1$, i.e., if the whole domain between 1 and 0 is called indeterminate. Then exceptions such as those mentioned cannot occur; only in case both a and b are indeterminate is there a certain ambiguity.¹⁶ Such a logic, however, does not apply to physics, as the cases $P(a) = 1$ or $P(a) = 0$ in practice do not occur; there would

¹⁶ Cf. the author's *Wahrscheinlichkeitslehre*, §§ 72 and 74.

be no true or false statements at all in physics if this logic were used. A transformation by division is accordingly bound to remain an approximation.

We turn now to the second method of transformation. It is made possible by the frequency interpretation of probability. We started from a relational system L between elements a, b, c, \dots ,

$$L[a, b, c, \dots]$$

As the "truth-value" of the elements a, b, c, \dots , varies continuously from 0 to 1, L has the character of a logic with continuous scale and signifies probability logic. We said that we may replace the elements a, b, c, \dots , by another set of elements $(a_i), (b_i), (c_i), \dots$, called propositional series; we have then the system

$$L[(a_i), (b_i), (c_i), \dots]$$

The truth-value of the elements $(a_i), (b_i), (c_i), \dots$, also varies on a continuous scale. Now the propositional series $(a_i), (b_i), \dots$, are built up of elements which are propositions of two truth-values only, and the "truth-value" of the propositional series (a_i) may be interpreted as the frequency with which the propositions a_i are true. By this interpretation, the relational system L is transformed into another relational system L_o

$$L_o [a_i, b_i, c_i, \dots]$$

We may compare this transition to the introduction of new variables in mathematics. L_o is nothing but the ordinary two-valued logic.

That is to say: Any statement about propositional series, within the frame of probability logic, may be transformed into a statement within the frame of two-valued

logic about the frequency with which propositions in a propositional series are true.

It is upon this transformation that the significance of the frequency interpretation is founded. The frequency interpretation allows us to eliminate the probability logic and to reduce probability statements to statements in the two-valued logic.

This transformation seems to be, in opposition to that by dichotomy or trichotomy, not of an approximative but of a strict character; however, it is so only if two conditions are fulfilled:

1. If the new elements a_i, b_i, \dots , are propositions of a strictly two-valued character; and
2. If the statement about the frequency with which propositions are true within a propositional series is of a strictly two-valued character.

These conditions are fulfilled for the purely mathematical calculus of probability; that is the reason why this calculus can be built up entirely within the frame of the two-valued logic. As for the application of this calculus to reality, i.e., to physical statements, these two conditions, however, are not fulfilled; for all statements of empirical science the transition indicated remains nothing but an approximation.

As to the second condition, the difficulty arises from the infinity of the series. A mathematically infinite series is given by a prescription which provides the means of calculating its qualities as far as they are demanded; in particular its relative frequency can be calculated. This is why the second condition offers no difficulties for mathematics. A physically infinite series, however, is known to us only in a determinate initial section; its further continuation is not known to us and remains dependent on the problematical means of induction. A statement about the

frequency of a physical series, therefore, cannot be uttered with certainty: this statement is in itself only probable. These reflections lead, as we see, into a theory of probability statements of higher levels; as these considerations involve some additional analyses, we may postpone the discussion of this theory to later sections (§§ 41 and 43). It may be sufficient for the present to state that the second condition cannot be fulfilled for statements of the empirical sciences.

At this point the first condition must be subjected to closer consideration. This condition is not fulfilled in empirical science because there are no propositions which are absolutely verifiable. Such was the result of our previous inquiries; we showed that it is only a schematization when we talk of a strictly true or false proposition. Before the throw of the die, we have only a probability statement about the result of the throw; after the throw we say that we know the result exactly. But, strictly speaking, this is only the transition from a low to a high probability; it is not absolutely certain that there is a die before me on the table showing the side 1. The same is valid for any other proposition whatever; we need not enter again into a discussion of this idea. If we consider the second condition as fulfilled—and for certain purposes this may be practical—this assumption is valid, therefore, only in the sense of a schematization.

We may indicate now what is performed in this schematization. Strictly speaking, the elementary propositions a_i possess for us a weight only; if we replace this weight by truth or falsehood, we perform a transformation by dichotomy or trichotomy. Thus the transformation from L to L_o , by the frequency interpretation, presupposes another transformation by division concerning the new set of elements.

The frequency interpretation, in introducing the two-valued logic, cannot thereby free us from the approximative character of this logic, even if we take no account of the second condition. This does not involve, however, the view that such a transition is superfluous; on the contrary, it is a procedure with which the degree of the approximation is highly enhanced. That is the reason why this transformation plays a dominant role among the methods of science.

We might try to construct our system of knowledge by giving every proposition an appraised weight; we should then find, however, that in this way we obtain a rather bad system of weights. The actual procedure of science replaces such a direct method by an indirect one, which must be regarded as one of the most perspicacious inventions of science. We begin with a trichotomous transformation, accept the propositions of high and low weight only, and drop the intermediate domain. Applying, then, the frequency interpretation of probability, we construct by counting-processes the weight of the propositions before omitted. This is not the only aim of our calculations; we may even control the weight of the propositions accepted in the beginning and possibly shift them from the supposed place within the scale of weights to a new place. Thus a proposition originally assumed to be true may afterward turn out to be indeterminate or false. This is not a contradiction within statistical method because the alteration of the truth-value of some of the elementary propositions does not, on the whole, greatly influence the frequency. We must constantly insist⁸ that what was assumed by appraisal as the weight is confirmed later on by a reduction to the frequency of other statements which are judged by appraisals as well. The original appraisals are thus submitted to a process of dissolution, directed by the frequency interpre-

tation. This process of dissolution leads to a new set of appraisals; the improvement associated with this procedure consists in the fact that every individual appraisal becomes less important, that its possible falsehood influences the whole system less. Thus by concerted action of trichotomy and frequency interpretation we construct a system of weights much more exact than we could obtain by a direct appraisal of the weights.

Within this procedure, the essential function of the frequency interpretation becomes manifest. Although our logic of propositions is not two-valued but of a continuous scale, we need not start knowledge with probability logic. We start with an approximative two-valued logic and develop the continuous scale by means of the frequency interpretation. The same method applies inversely: if a probability statement is given, we verify it by means of the frequency interpretation, in reducing it to statements of an approximative two-valued logic. This approximative logic is better than the original probability logic because it omits the doubtful middle domain of weights. It is the frequency interpretation of probability which makes this reduction possible, for in dissolving weights into frequencies it permits us to confine the direct appraisal of weights to such as are of a high or a low degree. The frequency interpretation frees us from the manipulation of a logical system which is too unhandy for direct use.

We must not forget, however, that the two-valued logic always remains approximative. The system of knowledge is written in the language of probability logic; the two-valued logic is a substitute language suitable only within the frame of an approximation. Any epistemology which overlooks this fact runs the risk of losing itself on the bare heights of an idealization.

§ 37. The aprioristic and the formalistic conception of logic

We must now turn to the question of the origin of the laws of probability logic. This question cannot be separated from the question concerning the origin of logic in general; we must enter, therefore, into an inquiry concerning the nature of logic.

In the history of philosophy there are two interpretations of logic which have played dominant roles, and which have endured to form the main subject matter of discussions on logic in our own day.

For the first interpretation, which we may call the *aprioristic interpretation*, logic is a science with its own authority, whether it is founded in the a priori nature of reason, or in the psychological nature of thought, or in intellectual intuition or evidence—philosophers have provided us with many such phrases, the task of which is to express that we simply have to submit to logic as to a kind of superior command.

Such was the conception of Plato, with visionary insight into ideas superadded; such was the doctrine of most scholastics for whom logic revealed the laws and nature of God; such was the conception of the modern rationalists, Descartes, Leibnitz, and Kant, men who must be considered as the founders of modern apriorism in logic and mathematics. The founders of the modern logic of probability, moreover, were not far removed from such a conception. They discovered that the laws of this logic are as evident as the laws of the older logic; they therefore conceived probability logic as the logic of “rational belief” in events the truth-value of which is not known, and thus as a continuation of a priori logic. Boole conceived his probability logic as an expression of the “laws of thought,” choosing this term as the title of his major work; Venn called prob-

ability logic “a branch of the general science of evidence,” and Keynes, the representative of this conception of probability logic in our day, renews the theory of “rational belief.” The dominion of apriorism, therefore, extends even into the ranks of the logicians.

The second interpretation does not acknowledge logic as a material science and may be called the *formalistic interpretation* of logic. The adherents of this interpretation do not believe in an a priori character of logic. They refuse even to talk of the “laws” of logic, this term suggesting that there is something in the nature of an authority in logic which we have to obey. For them logic is a system of rules which by no means determine the content of science, and which do nothing but furnish a transformation of one proposition into another without any addition to its intension. This conception of logic underlay the struggle of the nominalists in the Middle Ages; it was recognized by those empiricists, such as Hume, who saw the need of an explanation of the claim of necessity by logic; and it was to constitute the basis of the modern development of logistic associated with the names of Hilbert, Russell, Wittgenstein, and Carnap.¹⁷ Wittgenstein gave the important definition of the concept of tautology: A tautology is a formula the truth of which is independent of the truth-values of the elementary propositions contained in it. Logic in this way was defined as the domain of tautological formulas; the view as to the material emptiness of logic found its strict formulation in Wittgenstein’s definition.

Carnap added a point of view which was essential for the explanation of the claim of necessity by logic. Logic, he said, in continuation of the ideas of Wittgenstein, deals

¹⁷ It is to be noted here that we use the term “formalistic” in a sense somewhat wider than the sense in use within the discussion of modern logistic, where the formalists are represented by the narrower group centering around Hilbert. The differences between these groups are, however, not essential for our survey.

with language only, not with the objects of language. Language is built up of symbols, the use of which is determined by certain rules. Logical necessity, therefore, is nothing but a relation between symbols due to the rules of language. There is no logical necessity "inherent in things," such as the prophets of all kinds of "ontology" emphasize. The character of necessity is entirely on the side of the symbols; such necessities, however, say nothing about the world because the rules of language are constructed in such a way that they do not restrict the domain of experience.

Logic is accordingly called by Carnap the syntax of language. There are no logical laws of the world, but only syntactical rules of language. What we called a logical fact (§ 1), is to be called in this better terminology a syntactical fact. Instead of speaking of the logical fact that a sentence b cannot be deduced from a sentence a , it is better to speak of a syntactical fact: the structure of the formulas a and b is of such a kind that the syntactical relation "deducibility" does not hold between them.

The formalistic conception of logic frees us from all the problems of apriorism, from all questions of a correspondence between mind and reality. It is for this reason the natural logical theory of every empiricism. It does not demand from us any belief in nonempirical laws. What we know about nature is taken from experience; logic does not add anything to the results of experience because logic is empty, is nothing but a system of syntactical rules of language.

Let us ask now whether we may insert probability logic into the formalistic conception of logic. It is obvious that this is, for every variety of empiricism, a basic question. We found that the concept of probability is indispensable for knowledge, that probability logic determines the methods of scientific investigation. If we could not give a formalistic

interpretation of probability logic, all efforts of the anti-metaphysicians would have been in vain; in spite of their having overcome the difficulties of the two-valued logic, they would now fail before the concept which forms the very essence of scientific prediction—before the concept of probability. A logistic empiricism would be untenable if we should not succeed in finding a formalistic solution of the probability problem.

There is such a solution. To present it we shall proceed by two steps.

The first step is marked by the frequency interpretation. We showed that probability logic can be transformed into the two-valued logic by the frequency interpretation. Our statement of this transformation needs a supplementary remark. Though it is easily seen that such a transformation is obtained by the frequency interpretation, we do not know immediately whether or not this reduction requires axioms of another kind for which we may have no justification. This question can only be answered by an axiomatical procedure which reduces the mathematical calculus of probability to a system of simple presuppositions sufficient for the deduction of the whole mathematical system; the nature of these axioms has then to be considered.

This procedure has been carried through; it leads to a result of the highest relevance for our problem. It turns out that all theorems of probability reduce to one presupposition only: this is just the frequency interpretation. If probability is interpreted as the limit of the relative frequency in an infinite (or finite) series, all laws of probability reduce to arithmetical laws and, with this, become tautological. The demonstration of this theorem involves some complications, as the theory of mathematical probability refers to a great many types of probability series, the normal series, such as occur in games of chance, being

only a special type within this manifold. Even a short indication of this demonstration would unduly lengthen our exposition, so we must content ourselves with a statement of the result.¹⁸

The consequences of this result for the insertion of probability logic into the formalistic interpretation of logic are obvious: the problem of the justification of the laws of probability logic disappears. These laws are justified, as arithmetical laws, within the formalistic interpretation of mathematics. To see the effect of this result, let us remember the difficulties of the older writers on probability logic. They saw that the laws of probability, although admitted by everybody, cannot be logically deduced from the concept of probability if this concept is to mean something like reasonable expectation, or the chance of the occurrence of a single event; the laws, then, were to be synthetical and a priori. The conception of the "laws of rational belief" which expressed this idea originated from the fact that the deducibility of these laws from the frequency interpretation was not seen. We need no "science of evidence" to prove the laws of probability if we understand by probability the limit of a frequency. On the other hand, this is one of the reasons we must insist on the identity conception of the two probability concepts: if they were disparate, if there were a nonstatistical concept of probability, the justification of its laws by the frequency interpretation could not be given, and the formalistic interpretation of probability logic could not be carried through.¹⁹ We should

¹⁸ This reduction of the calculus of probability to one axiom concerning the existence of a limit of the frequency has been carried through in the author's paper, "Axiomatik der Wahrscheinlichkeitsrechnung," *Mathematische Zeitschrift*, XXXIV (1932), 568. A more detailed exposition has been given in the author's *Wahrscheinlichkeitslehre*.

¹⁹ This fact has not been sufficiently noticed by some modern positivists who have tried to defend the disparity conception against me (cf. my answer to Popper and Carnap in *Erkenntnis*, V [1935], 267).

be driven back into the aprioristic position and should be obliged to believe in laws we cannot justify. It is only the frequency interpretation which frees us from metaphysical assumptions and links the problem of probability with the continuous dissolution of the a priori which marks the development of modern logistic empiricism.

The reduction of the laws of probability to tautologies by the frequency interpretation is only the first step in this direction however. There remains a second step to be taken.

§ 38. The problem of induction

So far we have only spoken of the useful qualities of the frequency interpretation. It also has dangerous qualities.

The frequency interpretation has two functions within the theory of probability. First, a frequency is used as a *substantiation* for the probability statement; it furnishes the reason why we believe in the statement. Second, a frequency is used for the *verification* of the probability statement; that is to say, it is to furnish the meaning of the statement. These two functions are not identical. The observed frequency from which we start is only the basis of the probability inference; we intend to state another frequency which concerns *future observations*. The probability inference proceeds from a known frequency to one unknown; it is from this function that its importance is derived. The probability statement sustains a prediction, and this is why we want it.

It is the problem of induction which appears with this formulation. The theory of probability involves the problem of induction, and a solution of the problem of probability cannot be given without an answer to the question of induction. The connection of both problems is well known; philosophers such as Peirce have expressed the idea that a

solution of the problem of induction is to be found in the theory of probability. The inverse relation, however, holds as well. Let us say, cautiously, that the solution of both problems is to be given within the same theory.

In uniting the problem of probability with that of induction, we decide unequivocally in favor of that determination of the degree of probability which mathematicians call the *determination a posteriori*. We refuse to acknowledge any so-called *determination a priori* such as some mathematicians introduce in the theory of the games of chance; on this point we refer to our remarks in § 33, where we mentioned that the so-called determination a priori may be reduced to a determination a posteriori. It is, therefore, the latter procedure which we must now analyze.

By "determination a posteriori" we understand a procedure in which the relative frequency observed statistically is assumed to hold approximately for any future prolongation of the series. Let us express this idea in an exact formulation. We assume a series of events A and \bar{A} (non- A); let n be the number of events, m the number of events of the type A among them. We have then the relative frequency

$$h^n = \frac{m}{n}$$

The assumption of the determination a posteriori may now be expressed:

For any further prolongation of the series as far as s events ($s > n$), the relative frequency will remain within a small interval around h^n ; i.e., we assume the relation

$$h^n - \epsilon \leq h^s \leq h^n + \epsilon$$

where ϵ is a small number.

This assumption formulates the *principle of induction*. We may add that our formulation states the principle in a

form more general than that customary in traditional philosophy. The usual formulation is as follows: induction is the assumption that an event which occurred n times will occur at all following times. It is obvious that this formulation is a special case of our formulation, corresponding to the case $h^n = 1$. We cannot restrict our investigation to this special case because the general case occurs in a great many problems.

The reason for this is to be found in the fact that the theory of probability needs the definition of probability as the limit of the frequency. Our formulation is a necessary condition for the existence of a limit of the frequency near h^n ; what is yet to be added is that there is an h^s of the kind postulated for every ϵ however small. If we include this idea in our assumption, our postulate of induction becomes the hypothesis that there is a limit to the relative frequency which does not differ greatly from the observed value.

If we enter now into a closer analysis of this assumption, one thing needs no further demonstration: the formula given is not a tautology. There is indeed no logical necessity that h^s remains within the interval $h^n \pm \epsilon$; we may easily imagine that this does not take place.

The nontautological character of induction has been known a long time; Bacon had already emphasized that it is just this character to which the importance of induction is due. If inductive inference can teach us something new, in opposition to deductive inference, this is because it is not a tautology. This useful quality has, however, become the center of the epistemological difficulties of induction. It was David Hume who first attacked the principle from this side; he pointed out that the apparent constraint of the inductive inference, although submitted to by everybody, could not be justified. We believe in induction; we even cannot get rid of the belief when we know the impossibility

of a logical demonstration of the validity of inductive inference; but as logicians we must admit that this belief is a deception—such is the result of Hume's criticism. We may summarize his objections in two statements:

1. We have no logical demonstration for the validity of inductive inference.

2. There is no demonstration *a posteriori* for the inductive inference; any such demonstration would presuppose the very principle which it is to demonstrate.

These two pillars of Hume's criticism of the principle of induction have stood unshaken for two centuries, and I think they will stand as long as there is a scientific philosophy.

In spite of the deep impression Hume's discovery made on his contemporaries, its relevance was not sufficiently noticed in the subsequent intellectual development. I do not refer here to the speculative metaphysicians which the nineteenth century presented to us so copiously, especially in Germany; we need not be surprised that they did not pay any attention to objections which so soberly demonstrated the limitations of human reason. But empiricists, and even mathematical logicians, were no better in this respect. It is astonishing to see how clear-minded logicians, like John Stuart Mill, or Whewell, or Boole, or Venn, in writing about the problem of induction, disregarded the bearing of Hume's objections; they did not realize that any logic of science remains a failure so long as we have no theory of induction which is not exposed to Hume's criticism. It was without doubt their logical apriorism which prevented them from admitting the unsatisfactory character of their own theories of induction. But it remains incomprehensible that their empiricist principles did not lead them to attribute a higher weight to Hume's criticism.

It has been with the rise of the formalistic interpretation of logic in the last few decades that the full weight of Hume's objections has been once more realized. The demands for logical rigor have increased, and the blank in the chain of scientific inferences, indicated by Hume, could no longer be overlooked. The attempt made by modern positivists to establish knowledge as a system of absolute certainty found an insurmountable barrier in the problem of induction. In this situation an expedient has been proposed which cannot be regarded otherwise than as an act of despair.

The remedy was sought in the principle of retrogression. We remember the role this principle played in the truth theory of the meaning of indirect sentences (§ 7); positivists who had already tried to carry through the principle within this domain now made the attempt to apply it to the solution of the problem of induction. They asked: Under what conditions do we apply the inductive principle in order to infer a new statement? They gave the true answer: We apply it when a number of observations is made which concern events of a homogeneous type and which furnish a frequency h^n for a determinate kind of events among them. What is inferred from this? You suppose, they said, that you are able to infer from this a similar future prolongation of the series; but, according to the principle of retrogression, this "prediction of the future" cannot have a meaning which is more than a repetition of the premises of the inference—it means nothing but stating, "There *was* a series of observations of such and such kind." The meaning of a statement about the future is a statement about the past—this is what furnishes the application of the principle of retrogression to inductive inference.

I do not think that such reasoning would convince any

sound intellect. Far from considering it as an analysis of science, I should regard such an interpretation of induction rather as an act of intellectual suicide. The discrepancy between actual thinking and the epistemological result so obtained is too obvious. The only thing to be inferred from this demonstration is that the principle of retrogression does not hold if we want to keep our epistemological construction in correspondence with the actual procedure of science. We know pretty well that science wants to foresee the future; and, if anybody tells us that "foreseeing the future" means "reporting the past," we can only answer that epistemology should be something other than a play with words.

It is the postulate of utilizability which excludes the interpretation of the inductive inference in terms of the principle of retrogression. If scientific statements are to be utilizable for actions, they must pass beyond the statements on which they are based; they must concern future events and not those of the past alone. To prepare for action presupposes—besides a volitional decision concerning the aim of the action—some knowledge about the future. If we were to give a correct form to the reasoning described, it would amount to maintaining that there is no demonstrable knowledge about the future. This was surely the idea of Hume. Instead of any pseudo-solution of the problem of induction, we should then simply confine ourselves to the repetition of Hume's result and admit that the postulate of utilizability cannot be satisfied. The truth theory of meaning leads to a Humean skepticism—this is what follows from the course of the argument.

It was the intention of modern positivism to restore knowledge to absolute certainty; what was proposed with the formalistic interpretation of logic was nothing other than a resumption of the program of Descartes. The great

founder of rationalism wanted to reject all knowledge which could not be considered as absolutely reliable; it was the same principle which led modern logicians to a denial of a priori principles. It is true that this principle led Descartes himself to apriorism; but this difference may be considered as a difference in the stage of historical development—his rationalistic apriorism was to perform the same function of sweeping away all untenable scientific claims as was intended by the later struggle against a priori principles. The refusal to admit any kind of material logic—i.e., any logic furnishing information about some "matter"—springs from the Cartesian source: It is the ineradicable desire of absolutely certain knowledge which stands behind both the rationalism of Descartes and the logicism of positivists.

The answer given to Descartes by Hume holds as well for modern positivism. There is no certainty in any knowledge about the world because knowledge of the world involves predictions of the future. The ideal of absolutely certain knowledge leads into skepticism—it is preferable to admit this than to indulge in reveries about a priori knowledge. Only a lack of intellectual radicalism could prevent the rationalists from seeing this; modern positivists should have the courage to draw this skeptical conclusion, to trace the ideal of absolute certainty to its inescapable implications.

However, instead of such a strict disavowal of the predictive aim of science, there is in modern positivism a tendency to evade this alternative and to underrate the relevance of Hume's skeptical objections. It is true that Hume himself is not guiltless in this respect. He is not ready to realize the tragic consequences of his criticism; his theory of inductive belief as a habit—which surely cannot be called a solution of the problem—is put forward with

the intention of veiling the gap pointed out by him between experience and prediction. He is not alarmed by his discovery; he does not realize that, if there is no escape from the dilemma pointed out by him, science might as well not be continued—there is no use for a system of predictions if it is nothing but a ridiculous self-delusion. There are modern positivists who do not realize this either. They talk about the formation of scientific theories, but they do not see that, if there is no justification for the inductive inference, the working procedure of science sinks to the level of a game and can no longer be justified by the applicability of its results for the purpose of actions. It was the intention of Kant's synthetic a priori to secure this working procedure against Hume's doubts; we know today that Kant's attempt at rescue failed. We owe this critical result to the establishment of the formalistic conception of logic. If, however, we should not be able to find an answer to Hume's objections within the frame of logistic formalism, we ought to admit frankly that the antimetaphysical version of philosophy led to the renunciation of any justification of the predictive methods of science—led to a definitive failure of scientific philosophy.

Inductive inference cannot be dispensed with because we need it for the purpose of action. To deem the inductive assumption unworthy of the assent of a philosopher, to keep a distinguished reserve, and to meet with a condescending smile the attempts of other people to bridge the gap between experience and prediction is cheap self-deceit; at the very moment when the apostles of such a higher philosophy leave the field of theoretical discussion and pass to the simplest actions of daily life, they follow the inductive principle as surely as does every earth-bound mind. In any action there are various means to the realization of our aim; we have to make a choice, and we decide

in accordance with the inductive principle. Although there is no means which will produce with certainty the desired effect, we do not leave the choice to chance but prefer the means indicated by the principle of induction. If we sit at the wheel of a car and want to turn the car to the right, why do we turn the wheel to the right? There is no certainty that the car will follow the wheel; there are indeed cars which do not always so behave. Such cases are fortunately exceptions. But if we should not regard the inductive prescription and consider the effect of a turn of the wheel as entirely unknown to us, we might turn it to the left as well. I do not say this to suggest such an attempt; the effects of skeptical philosophy applied in motor traffic would be rather unpleasant. But I should say a philosopher who is to put aside his principles any time he steers a motorcar is a bad philosopher.

It is no justification of inductive belief to show that it is a habit. It *is* a habit; but the question is whether it is a good habit, where "good" is to mean "useful for the purpose of actions directed to future events." If a person tells me that Socrates is a man, and that all men are mortal, I have the habit of believing that Socrates is mortal. I know, however, that this is a good habit. If anyone had the habit of believing in such a case that Socrates is not mortal, we could demonstrate to him that this was a bad habit. The analogous question must be raised for inductive inference. If we should not be able to demonstrate that it is a good habit, we should either cease using it or admit frankly that our philosophy is a failure.

Science proceeds by induction and not by tautological transformations of reports. Bacon is right about Aristotle; but the *novum organon* needs a justification as good as that of the *organon*. Hume's criticism was the heaviest blow against empiricism; if we do not want to dupe our con-

sconsciousness of this by means of the narcotic drug of aprioristic rationalism, or the soporific of skepticism, we must find a defense for the inductive inference which holds as well as does the formalistic justification of deductive logic.

§ 39. The justification of the principle of induction

We shall now begin to give the justification of induction which Hume thought impossible. In the pursuit of this inquiry, let us ask first what has been proved, strictly speaking, by Hume's objections.

Hume started with the assumption that a justification of inductive inference is only given if we can show that inductive inference must lead to success. In other words, Hume believed that any justified application of the inductive inference presupposes a demonstration that the conclusion is true. It is this assumption on which Hume's criticism is based. His two objections directly concern only the question of the truth of the conclusion; they prove that the truth of the conclusion cannot be demonstrated. The two objections, therefore, are valid only in so far as the Humean assumption is valid. It is this question to which we must turn: Is it necessary, for the justification of inductive inference, to show that its conclusion is true?

A rather simple analysis shows us that this assumption does not hold. Of course, if we were able to prove the truth of the conclusion, inductive inference would be justified; but the converse does not hold: a justification of the inductive inference does not imply a proof of the truth of the conclusion. The proof of the truth of the conclusion is only a sufficient condition for the justification of induction, not a necessary condition.

The inductive inference is a procedure which is to furnish us the best assumption concerning the future. If we do not know the truth about the future, there may be nonetheless

a best assumption about it, i.e., a best assumption relative to what we know. We must ask whether such a characterization may be given for the principle of induction. If this turns out to be possible, the principle of induction will be justified.

An example will show the logical structure of our reasoning. A man may be suffering from a grave disease; the physician tells us: "I do not know whether an operation will save the man, but if there *is* any remedy, it is an operation." In such a case, the operation would be justified. Of course, it would be better to know that the operation will save the man; but, if we do not know this, the knowledge formulated in the statement of the physician is a sufficient justification. If we cannot realize the sufficient conditions of success, we shall at least realize the necessary conditions. If we were able to show that the inductive inference is a necessary condition of success, it would be justified; such a proof would satisfy any demands which may be raised about the justification of induction.

Now obviously there is a great difference between our example and induction. The reasoning of the physician presupposes inductions; his knowledge about an operation as the only possible means of saving a life is based on inductive generalizations, just as are all other statements of empirical character. But we wanted only to illustrate the logical structure of our reasoning. If we want to regard such a reasoning as a justification of the principle of induction, the character of induction as a necessary condition of success must be demonstrated in a way which does not presuppose induction. Such a proof, however, can be given.

If we want to construct this proof, we must begin with a determination of the aim of induction. It is usually said that we perform inductions with the aim of foreseeing the

future. This determination is vague; let us replace it by a formulation more precise in character:

The aim of induction is to find series of events whose frequency of occurrence converges toward a limit.

We choose this formulation because we found that we need probabilities and that a probability is to be defined as the limit of a frequency; thus our determination of the aim of induction is given in such a way that it enables us to apply probability methods. If we compare this determination of the aim of induction with determinations usually given, it turns out to be not a confinement to a narrower aim but an expansion. What we usually call "foreseeing the future" is included in our formulation as a special case; the case of knowing with certainty for every event A the event B following it would correspond in our formulation to a case where the limit of the frequency is of the numerical value 1. Hume thought of this case only. Thus our inquiry differs from that of Hume in so far as it conceives the aim of induction in a generalized form. But we do not omit any possible applications if we determine the principle of induction as the means of obtaining the limit of a frequency. If we have limits of frequency, we have all we want, including the case considered by Hume; we have then the laws of nature in their most general form, including both statistical and so-called causal laws—the latter being nothing but a special case of statistical laws, corresponding to the numerical value 1 of the limit of the frequency. We are entitled, therefore, to consider the determination of the limit of a frequency as the aim of the inductive inference.

Now it is obvious that we have no guaranty that this aim is at all attainable. The world may be so disorderly that it is impossible for us to construct series with a limit. Let us introduce the term "predictable" for a world which

is sufficiently ordered to enable us to construct series with a limit. We must admit, then, that we do not know whether the world is predictable.

But, if the world is predictable, let us ask what the logical function of the principle of induction will be. For this purpose, we must consider the definition of limit. The frequency h^n has a limit at p , if for any given ϵ there is an n such that h^n is within $p \pm \epsilon$ and remains within this interval for all the rest of the series. Comparing our formulation of the principle of induction (§ 38) with this, we may infer from the definition of the limit that, if there is a limit, there is an element of the series from which the principle of induction leads to the true value of the limit. In this sense the principle of induction is a necessary condition for the determination of a limit.

It is true that, if we are faced with the value h^n for the frequency furnished by our statistics, we do not know whether this n is sufficiently large to be identical with, or beyond, the n of the "place of convergence" for ϵ . It may be that our n is not yet large enough, that after n there will be a deviation greater than ϵ from p . To this we may answer: We are not bound to stay at h^n ; we may continue our procedure and shall always consider the last h^n obtained as our best value. This procedure must at sometime lead to the true value p , if there is a limit at all; the applicability of this procedure, as a whole, is a necessary condition of the existence of a limit at p .

To understand this, let us imagine a principle of a contrary sort. Imagine a man who, if h^n is reached, always makes the assumption that the limit of the frequency is at $h^n + a$, where a is a fixed constant. If this man continues his procedure for increasing n , he is sure to miss the limit; this procedure must at sometime become false, if there is a limit at all.

We have found now a better formulation of the necessary condition. We must not consider the individual assumption for an individual h^n ; we must take account of the procedure of continued assumptions of the inductive type. The applicability of this procedure is the necessary condition sought.

If, however, it is only the whole procedure which constitutes the necessary condition, how may we apply this idea to the individual case which stands before us? We want to know whether the individual h^n observed by us differs less than ϵ from the limit of the convergence; this neither can be guaranteed nor can it be called a necessary condition of the existence of a limit. So what does our idea of the necessary condition imply for the individual case? It seems that for our individual case the idea turns out to be without any application.

This difficulty corresponds in a certain sense to the difficulty we found in the application of the frequency interpretation to the single case. It is to be eliminated by the introduction of a concept already used for the other problem: the concept of posit.

If we observe a frequency h^n and assume it to be the approximate value of the limit, this assumption is not maintained in the form of a true statement; it is a posit such as we perform in a wager. We posit h^n as the value of the limit, i.e., we wager on h^n , just as we wager on the side of a die. We know that h^n is our best wager, therefore we posit it. There is, however, a difference as to the type of posit occurring here and in the throw of the die.

In the case of the die, we know the weight belonging to the posit: it is given by the degree of probability. If we posit the case "side other than that numbered 1," the weight of this posit is $5/6$. We speak in this case of a posit with appraised weight, or, in short, of an *appraised posit*.

In the case of our positing h^n , we do not know its weight. We call it, therefore, a *blind posit*. We know it is our best posit, but we do not know how good it is. Perhaps, although our best, it is a rather bad one.

The blind posit, however, may be corrected. By continuing our series, we obtain new values h^n ; we always choose the last h^n . Thus the blind posit is of an approximative type; we know that the method of making and correcting such posits must in time lead to success, in case there is a limit of the frequency. It is this idea which furnishes the justification of the blind posit. The procedure described may be called the *method of anticipation*; in choosing h^n as our posit, we anticipate the case where n is the "place of convergence." It may be that by this anticipation we obtain a false value; we know, however, that a continued anticipation must lead to the true value, if there is a limit at all.

An objection may arise here. It is true that the principle of induction has the quality of leading to the limit, if there is a limit. But is it the only principle with such a property? There might be other methods which also would indicate to us the value of the limit.

Indeed, there might be. There might be even better methods, i.e., methods giving us the right value p of the limit, or at least a value better than ours, at a point in the series where h^n is still rather far from p . Imagine a clairvoyant who is able to foretell the value p of the limit in such an early stage of the series; of course we should be very glad to have such a man at our disposal. We may, however, without knowing anything about the predictions of the clairvoyant, make two general statements concerning them: (1) The indications of the clairvoyant can differ, if they are true, only in the beginning of the series, from those given by the inductive principle. In the end there

must be an asymptotical convergence between the indications of the clairvoyant and those of the inductive principle. This follows from the definition of the limit. (2) The clairvoyant might be an imposter; his prophecies might be false and never lead to the true value p of the limit.

The second statement contains the reason why we cannot admit clairvoyance without control. How gain such control? It is obvious that the control is to consist in an application of the inductive principle: we demand the forecast of the clairvoyant and compare it with later observations; if then there is a good correspondence between the forecasts and the observations, we shall infer, by induction, that the man's prophecies will also be true in the future. Thus it is the principle of induction which is to decide whether the man is a good clairvoyant. This distinctive position of the principle of induction is due to the fact that we know about its function of finally leading to the true value of the limit, whereas we know nothing about the clairvoyant.

These considerations lead us to add a correction to our formulations. There are, of course, many necessary conditions for the existence of a limit; that one which we are to use however must be such that its character of being necessary must be known to us. This is why we must prefer the inductive principle to the indications of the clairvoyant and control the latter by the former: we control the unknown method by a known one.

Hence we must continue our analysis by restricting the search for other methods to those about which we may know that they must lead to the true value of the limit. Now it is easily seen not only that the inductive principle will lead to success but also that every method will do the same if it determines as our wager the value

$$h^n + c_n$$

where c_n is a number which is a function of n , or also of h^n , but bound to the condition

$$\lim_{n \rightarrow \infty} c_n = 0$$

Because of this additional condition, the method must lead to the true value p of the limit; this condition indicates that all such methods, including the inductive principle, must converge asymptotically. The inductive principle is the special case where

$$c_n = 0$$

for all values of n .

Now it is obvious that a system of wagers of the more general type may have advantages. The "correction" c_n may be determined in such a way that the resulting wager furnishes even at an early stage of the series a good approximation of the limit p . The prophecies of a good clairvoyant would be of this type. On the other hand, it may happen also that c_n is badly determined, i.e., that the convergence is delayed by the correction. If the term c_n is arbitrarily formulated, we know nothing about the two possibilities. The value $c_n = 0$ —i.e., the inductive principle—is therefore the value of the smallest risk; any other determination may worsen the convergence. This is a practical reason for preferring the inductive principle.

These considerations lead, however, to a more precise formulation of the logical structure of the inductive inference. We must say that, if there is any method which leads to the limit of the frequency, the inductive principle will do the same; if there is a limit of the frequency, the inductive principle is a sufficient condition to find it. If we omit now the premise that there is a limit of the fre-

quency, we cannot say that the inductive principle is the necessary condition of finding it because there are other methods using a correction c_n . There is a set of equivalent conditions such that the choice of one of the members of the set is necessary if we want to find the limit; and, if there is a limit, each of the members of the set is an appropriate method for finding it. We may say, therefore, that the *applicability* of the inductive principle is a necessary condition of the existence of a limit of the frequency.

The decision in favor of the inductive principle among the members of the set of equivalent means may be substantiated by pointing out its quality of embodying the smallest risk; after all, this decision is not of a great relevance, as all these methods must lead to the same value of the limit if they are sufficiently continued. It must not be forgotten, however, that the method of clairvoyance is not, without further ado, a member of the set because we do not know whether the correction c_n occurring here is submitted to the condition of convergence to zero. This must be proved first, and it can only be proved by using the inductive principle, viz., a method known to be a member of the set: this is why clairvoyance, in spite of all occult pretensions, is to be submitted to the control of scientific methods, i.e., by the principle of induction.

It is in the analysis expounded that we see the solution of Hume's problem.²⁰ Hume demanded too much when he wanted for a justification of the inductive inference a proof that its conclusion is true. What his objections demonstrate is only that such a proof cannot be given. We do not perform, however, an inductive inference with the pretension of obtaining a true statement. What we obtain is a

²⁰ This theory of induction was first published by the author in *Erkenntnis*, III (1933), 421–25. A more detailed exposition was given in the author's *Wahrscheinlichkeitslehre*, § 80.

wager; and it is the best wager we can lay because it corresponds to a procedure the applicability of which is the necessary condition of the possibility of predictions. To fulfil the conditions sufficient for the attainment of true predictions does not lie in our power; let us be glad that we are able to fulfil at least the conditions necessary for the realization of this intrinsic aim of science.

§ 40. Two objections against our justification of induction

Our analysis of the problem of induction is based on our definition of the aim of induction as the evaluation of a limit of the frequency. Certain objections may be raised as to this statement of the aim of induction.

The first objection is based on the idea that our formulation demands too much, that the postulate of the existence of the limit of the frequency is too strong a postulate. It is argued that the world might be predictable even if there are no limits of frequencies, that our definition of predictability would restrict this concept too narrowly, excluding other types of structure which might perhaps be accessible to predictions without involving series of events with limits of their frequencies. Applied to our theory of induction, this objection would shake the cogency of our justification; by keeping strictly to the principle of induction, the man of science might exclude other possibilities of foreseeing the future which might work even if the inductive inference should fail.²¹

To this we must reply that our postulate does not demand the existence of a limit of the frequency for all series of events. It is sufficient if there is a certain number of series of this kind; by means of these we should then be

²¹ This objection has been raised by P. Hertz, *Erkenntnis*, VI (1936), 25; cf. also my answer, *ibid.*, p. 32.

able to determine the other series. We may imagine series which oscillate between two numerical values of the frequency; it can be shown that the description of series of this type is reducible to the indication of determinable subseries having a limit of the frequency. Let us introduce the term *reducible series* for series which are reducible to other series having a limit of their frequency; our definition of predictability then states only that the world is constituted by reducible series. The inductive procedure, the method of anticipation and later correction, will lead automatically to distinguishing series having a limit from other series and to the description of these others by means of the series having a limit. We cannot enter here into the mathematical details of this problem; for an elaboration of this we must refer to another publication.²²

To elude our defense, the objection might be continued by the construction of a world in which there is no series having a limit. In such a world, so our adversary might argue, there might be a clairvoyant who knows every event of a series individually, who could foretell precisely what would happen from event to event—is not this “foreseeing the future” without having a limit of a frequency at one’s disposal?

We cannot admit this. Let us call C the case in which the prediction of the clairvoyant corresponds to the event observed later, \bar{C} (non- C) the opposite case. Now if the clairvoyant should have the faculty supposed, the series of events of the type C and \bar{C} would define a series with a limit of the frequency. If the man should be a perfect prophet, this limit would be the number 1; however we may admit less perfect prophets with a lower limit. Anyway, we have constructed here a series with a limit. We must have such a series if we want to control the prophet; our control

²² Cf. *ibid.*, p. 36.

would consist in nothing but the application of the principle of induction to the series of events C and \bar{C} , i.e., in an inductive inference as to the reliability of the prophet, based on his successes. Only if the reduction to such a series with a limit is possible can we know whether or not the man is a good prophet because only this reduction gives us the means of control.

We see from this consideration that the case imagined is not more general but less general than our world of reducible series. A forecast giving us a true determination of every event is a much more special case than the indication of the limit of the frequency and is therefore included in our inductive procedure. We see, at the same time, that our postulate of the existence of limits of the frequencies is not a restriction of the concept of predictability. Any method of prediction defines by itself a series with a limit of the frequency; therefore, if prediction is possible, there are series with limits of the frequencies.

We are entitled, therefore, to call the applicability of the inductive procedure a necessary condition of predictability. We see at the same time why such a relation holds: *it is a logical consequence of the definition of predictability.* This is why we can give our demonstration of the unique position of the inductive principle by means of tautological relations only. *Although the inductive inference is not a tautology, the proof that it leads to the best posit is based on tautologies only.* The formal conception of logic was placed, by the problem of induction, before the paradox that an inference which leads to something new is to be justified within a conception of logic which allows only empty, i.e., tautological, transformations: this paradox is solved by the recognition that the “something new” furnished by the inference is not maintained as a true statement but as our best posit, and that the demonstration is not directed

toward the truth of the conclusion but to the logical relation of the procedure to the aim of knowledge.

There might be raised, instinctively, an objection against our theory of induction: that there appears some thing like "a necessary condition of knowledge"—a concept which is accompanied since Kant's theory of knowledge by rather an unpleasant flavor. In our theory, however, this quality of the inductive principle does not spring from any a priori qualities of human reason but has its origin in other sources. He who wants something must say what he wants; he who wants to predict must say what he understands by predicting. If we try to find a definition of this term which corresponds, at least to some extent, to the usual practice of language, the definition—*independently of further determination*—will turn out to entail the postulate of the existence of certain series having a limit of the frequency. It is from this component of the definition that the character of the inductive principle as being a necessary condition of predictability is deduced. The application of the principle of induction does not signify, therefore, any restriction or any renunciation of predictability in another form—it signifies nothing but the mathematical interpretation of what we mean by predictability, properly speaking.

We turn now to a second objection. It was the claim of the first objection that our definition of predictability demands too much; the second objection, on the contrary, holds that this definition demands too little, that what we call predictability is not a sufficient condition of actual predictions. This objection arises from the fact that our definition admits infinite series of events; to this conception is opposed the view that a series actually observable is always finite, of even a rather restricted length, determined by the short duration of human lives.

We shall not deny the latter fact. We must admit that there may be a series of events having a limit whose convergence begins so late that the small portion of the series observed by human beings does not reveal any indication of the later convergence. Such a series would have for us the character of a nonconverging series. Applying the principle of induction, we should never have success with our inferences; after a short time, our posits would always turn out false. Although, in such a case, the condition of predictability would be fulfilled, the inductive procedure would not be a practically sufficient means for discovering it.

We shall not deny this consequence either. We do not admit, however, that the case considered raises any objection to our theory. We did not start for our justification of induction from a presupposition that there are series having a limit; in spite of this, we contrived to give the justification sought. This was made possible by the use of the concept of necessary condition; we said that, if we are not sure of the possibility of success, we should at least realize its necessary conditions. The case of convergence coming too late amounts to the same thing as the case of nonconvergence, as far as human abilities are concerned. However, if we succeed in giving a justification of the inductive procedure even if this worst of all cases cannot be excluded a priori, our justification will also have taken account of the other case—the case of a convergence which is too late.

Let us introduce the term *practical limit* for a series showing a sufficient convergence within a domain accessible to human observations; we may add that we may cover by this term the case of a series which, though not converging at infinity, shows an approximate convergence in a segment of the series, accessible in practice and sufficiently long (a so-called "semiconvergent series"). We

may then say that our theory is not concerned with a mathematical limit but with a practical limit. Predictability is to be defined by means of the practical limit, and the inductive procedure is a sufficient condition of success only if the series in question has a practical limit. With these concepts, however, we may carry through our argument just as well. The applicability of the inductive procedure may be shown, even within the domain of these concepts, to be the necessary condition of predictability.

It is the concept of necessary condition on which our reasoning is based. It is true that, if the series in question should have no practical limit—including the case of too late a convergence—this would imply the inefficiency of the inductive procedure. The possibility of this case, however, need not restrain us from at least wagering on success. Only if we knew that the unfavorable case is actual, should we renounce attempts at prediction. But obviously this is not our situation. We do not know whether we shall have success; but we do not know the contrary either. Hume believed that a justification of induction could not be given because *we do not know whether we shall have success*; the correct formulation, instead, would read that a justification of induction could not be given if *we knew that we should have no success*. We are not in the latter situation but in the former; the question of success is for us indeterminate, and we may therefore at least dare a wager. The wager, however, should not be arbitrarily laid but chosen as favorably as possible; we should at least actualize the necessary conditions of success, if the sufficient conditions are not within our reach. The applicability of the inductive procedure being a necessary condition of predictability, this procedure will determine our best wager.

We may compare our situation to that of a man who

wants to fish in an unexplored part of the sea. There is no one to tell him whether or not there are fish in this place. Shall he cast his net? Well, if he wants to fish in that place I should advise him to cast the net, to take the chance at least. It is preferable to try even in uncertainty than not to try and be certain of getting nothing.

§ 41. Concatenated inductions

The considerations concerning the possibility of too slow a convergence of the series could not shake our justification of the inductive procedure, as signifying at least an attempt to find a practically convergent series; they do point out, however, the utility of methods which would lead to a quicker approximation, i.e., which would indicate the true value of the limit at a point in the series where the relative frequency is still rather different from the limiting value. We may want even more; we may want methods which give us the numerical value of the limit before the physical actualization of the series has begun—a problem which may be considered as an extreme case of the first problem. The elaboration of such methods is indeed a question of the greatest relevance; we shall ask now whether or not they exist, and how they are to be found.

We have already met with an example which may be considered as the transition to a method of quicker approximation. We discussed the possibility of a clairvoyant and said that his capacities might be controlled by the inductive principle; we said that, should the control confirm the predictions, the clairvoyant was to be considered a reliable prophet, and his indications as superior to those of the inductive principle. This idea shows an important feature of inductive methods. We may sometimes infer by means of the inductive principle that it is better to apply some other method of prediction; the inductive principle

may lead to its own supersession. This is no contradiction; on the contrary, there is no logical difficulty in such a procedure—it even signifies one of the most useful methods of scientific inquiry.

If we want to study inferences of such a type, we need not trouble clairvoyants or oracles of a mystic kind: science itself has developed such methods to a vast extent. The method of scientific inquiry may be considered as a concatenation of inductive inferences, with the aim of superseding the inductive principle in all those cases in which it would lead to a false result, or in which it would lead us too late to the right result. It is to this procedure of concatenated inductions that the overwhelming success of scientific method is due. The complication of the procedure has become the reason why it has been misinterpreted by many philosophers; the apparent contradiction to a direct application of the inductive principle, in individual cases, has been considered as a proof for the existence of noninductive methods which were to be superior to the "primitive" method of induction. Thus the principle of causal connection has been conceived as a noninductive method which was to furnish us with an "inner connection" of the phenomena instead of the "mere succession" furnished by induction. Such interpretations reveal a profound misunderstanding of the methods of science. There is no difference between causal and inductive laws; the former are nothing but a special case of the latter. They are the case of a limit equal to 1, or at least approximately equal to 1; if we know, in such a case, the value of the limit even before the series has begun, we have the case of the individual prediction of future events happening in novel conditions, such as is demanded within the causal conception of knowledge. This case, therefore, is included in our theory of concatenated inductions.

The connecting link, within all chains of inferences leading to predictions, is always the inductive inference. This is because among all scientific inferences there is only one of an overreaching type: that is the inductive inference. All other inferences are empty, tautological; they do not add anything new to the experiences from which they start. The inductive inference does; that is why it is the elementary form of the method of scientific discovery. However, it is the only form; there are no cases of connections of phenomena assumed by science which do not fit into the inductive scheme. We need only construct this scheme in a sufficiently general form to include all methods of science. For this purpose, we must turn now to an analysis of concatenated inductions.

We begin with a rather simple case which already shows the logical structure by which the inductive inference may be superseded in an individual case. Chemists have found that almost all substances will melt if they are sufficiently heated; only carbon has not been liquefied. Chemists do not believe, however, that carbon is infusible; they are convinced that at a higher temperature carbon will also melt and that it is due only to the imperfection of our technical means that a sufficiently high temperature has not yet been attained. To construe the logical structure of the inferences connected with these experiences, let us denote by A the melted state of the substance, by \bar{A} the contrary state, and arrange the states in a series of ascending temperatures; we then have the scheme

Copper: $\bar{A} \bar{A} A A A A A A A \dots$
Iron: $\bar{A} \bar{A} \bar{A} A A A A A A \dots$

Carbon: $\bar{A} \bar{A} \bar{A} \bar{A} \bar{A} \bar{A} \bar{A} \bar{A} \bar{A} \dots$

To this scheme, which we call a *probability lattice*, we apply the inductive inference in two directions. The first is the horizontal. For the first lines it furnishes the result that above a certain temperature the substance will always be in liquid state. (Our example is a special case of the inductive inference, where the limit of the frequency is equal to 1.) For the last line, the corresponding inference would furnish the result that carbon is infusible. Here, however, an inference in the vertical direction intervenes; it states that in all the other cases the series leads to melting, and infers from this that the same will hold for the last line if the experiment is sufficiently continued. We see that here a cross-induction concerning a series of series occurs, and that this induction of the second level supersedes an induction of the first level.

This procedure may be interpreted in the following way. Applying the inductive principle in the horizontal direction, we proceed to posits concerning the limit of the frequency; these are blind posits, as we do not know a co-ordinated weight. Presupposing the validity of these posits, we then count in the vertical direction and find that the value 1 has a high relative frequency among the horizontal limits, whereas the value 0 furnished by the last line is an exception. In this way we obtain a weight for the horizontal limits; thus the blind posits are transformed into posits with appraised weight. Regarding the weights obtained we now correct the posit of the last line into one with the highest weight. The procedure may therefore be conceived as a transformation of blind posits into posits with appraised weights, combined with corrections following from the weights obtained—a typical probability method, based on the frequency interpretation. It makes use of the existence of probabilities of different levels. The fre-

quency within the horizontal lines determines a probability of the first level; counting the frequency within a series the elements of which are themselves series we obtain a probability of the second level.²³ The probability of the second level determines the weight of the sentence stating a probability of the first level. We must not forget, however, that the transformation into an appraised posit concerns only the posits of the first level, whereas the posits of the second level remain blind. Thus at the end of the transformation there appears a blind posit of higher level. This of course may also be transformed into a posit with appraised weight, if we incorporate it into a higher manifold, the elements of which are series of series; it is obvious that this transformation will again furnish a new blind posit of a still higher level. We may say: Every blind posit may be transformed into a posit with appraised weight, but the transformation introduces new blind posits. Thus there will always be some blind posits on which the whole concatenation is based.

Our example concerns a special case in so far as the limits occurring are 1 and 0 only. If we want to find examples of the general case, we must pass to cases of statistical laws. To have a model of the inferences occurring, let us consider an example of the theory of games of chance, chosen in such a form that simplified inferences occur.

Let there be a set of three urns containing white and black balls in different ratios of combination; suppose we know that the ratios of the white balls to the total number of balls are 1:4, 2:4, and 3:4, but that we do not know to which urn each of these ratios belongs. We choose an urn, then make four draws from it (always putting the drawn

²³ As to the theory of probabilities of higher levels cf. the author's *Wahrscheinlichkeitslehre*, §§ 56–60.

ball back into the urn before the following draw), and obtain three white balls. Relative to further draws from the same urn, we have now two questions:

1. What is the probability of a white ball?

According to the inductive principle, this question will be answered by $3/4$. This is a blind posit. To transform it into a posit with appraised weight, we proceed to the second question:

2. What is the probability that the probability of a white ball is $3/4$?

This question concerns a probability of the second level; it is equivalent to the question as to the probability judged on the basis of the draws already made that the chosen urn contains the ratio $3/4$. The calculus of probability, by means of considerations also involving a problem of a probability lattice, gives to this question a rather complicated answer which we need not here analyze; in our example it furnishes the value $27/46$. We see that though our best posit in the given case will be the limit $3/4$ of the frequency, this posit is not very good; it itself has only the weight $27/46$. Considering the next drawing, as a single case, we have here two weights: the weight $3/4$ for the drawing of a white ball, and the weight $27/46$ for the value $3/4$ of the first weight. The second weight in this case is smaller than the first; if, to obtain a comparison, we write the weights in decimal fractions, we have 0.75 for the first and 0.59 for the second weight.

In this example the original posit is confirmed by the determination of the weight of the second level, this one being greater than $1/2$, and therefore greater than the second level weight belonging to the wagers on the limit $2/4$ or $1/4$. By another choice of the numerical values, a case of correction would result, i.e., a case in which the weight of the second level would incline us to change the

first posit. If there were twenty urns, nineteen of which contained white balls in a ratio of $1/4$, and only one contained white balls in the ratio of $3/4$, the probability at the second level would become $9/28 = 0.32$; in such a case, we should correct the first posit and posit the limit $1/4$, in opposition to the principle of Induction. The occurrence of three white balls among four would then be regarded as a chance exception which could not be considered as a sufficient basis for an inductive inference; this correction would be due to the change of a blind posit into an appraised one.

Our example is, as we said, simplified; this simplification is contained in the following two points. First, we presupposed some knowledge about the possible values of the probabilities of the first level: that there are in dispute only the three values $1/4$, $2/4$, and $3/4$ (in the second case: only the two values $1/4$ and $3/4$). Second, we presupposed that the urns are equally probable for our choice, i.e., we attribute to the urns the initial probabilities $1/3$ (in the second case: $1/20$); this presupposition is also contained in the calculation of the value $27/46$ (in the second case: $9/28$) for the probability of the second level.

In general, we are not entitled to such presuppositions. We are rather obliged to make inquiries as to the possible values of the probabilities of the first level and their corresponding initial probabilities. The structure of these inferences is also to be expressed in a probability lattice, but of a type more general than that used in the example concerning the melting of chemical substances; the limits of the frequencies occurring here are not just 1 or 0. The answers can only be given in the form of posits based on frequency observations, so that the whole calculation involves still further posits and posits of the blind type. This is why we cannot dispense with blind posits; although each can be

transformed into an appraised posit, new blind posits are introduced by the transformation itself.²⁴

Before we enter into an analysis of this process leading to posits and weights of higher levels, we must discuss some objections against our probability interpretation of scientific inferences. It might be alleged that not all scientific inferences are purely of the probability type and so not fully covered by our inductive schema. The objection may run that there are causal assumptions behind our inferences without which we should not venture to place our wagers. In our chemical example, the posit of the limit 1 in the horizontal lines of the figure is not only based on a simple enumeration of the A 's and \bar{A} 's; we know that if a substance is once melted it will not become solid at a higher temperature. Neither is our positing the possibility of liquefying carbon at higher temperatures based on simply counting the lines of the figure; we know from the atomic theory of matter that heat, in increasing the velocity of the atoms, must have the effect of decomposing the structure of the solids. Causal assumptions of this kind play a decisive role in such inferences as furnished by the example.

Although we shall not deny the relevance of considerations of this kind as far as the actual inference of the physicist is concerned, their occurrence, however, does not preclude the possibility that these so-called causal assumptions admit an interpretation of the inductive type. We simplified our analysis to show the inductive structure of the main inferences; what is shown by the objection is that an isolation of some of the inductive chains is not correct, that every case is incorporated in the whole concatenation of knowledge. Our thesis that all inferences occurring are

²⁴ For an exact analysis of these inferences cf. the author's *Wahrscheinlichkeitslehre*, § 77.

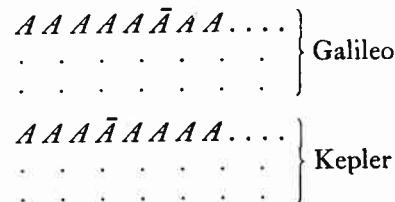
of the inductive type is not thereby shaken. We shall show this by another example which will clarify the inductive nature of so-called causal explanations.

Newton's law of gravitation has always been considered as the prototype of an explicative law. Galileo's law of falling bodies and Kepler's law of the elliptic motion of celestial bodies were inductive generalizations of observed facts; but Newton's law, it is said, was a causal explication of the facts observed. Newton did not observe facts but reflected upon them; his idea of an attractive force explained the motions observed, and the mathematical form he gave to his ideas shows no resemblance to methods of probability such as occur in our scheme. Is not this a proof against our inductive interpretation of scientific inferences?

I cannot admit this. On the contrary, Newton's discovery seems to me to involve typical methods of the probability procedure of science. To show this, let us enter into a more detailed analysis of the example.

The experiments of Galileo were performed on falling bodies whose spatiotemporal positions he observed; he found that the quantities measured fit into the formula $s = gt^2/2$, and inferred, by means of the inductive principle, that the same law holds for similar cases. Let us denote by A the case that the spatiotemporal values measured fulfil the relation $s = gt^2/2$; we have then a series in which A has been observed with a relative frequency almost equal to 1, and for which we maintain a limit of the frequency at 1. Correspondingly, Kepler observed a series of spatiotemporal positions of the planet Mars and found that they may be connected by a mathematical relation which he called the Law of Areas. If we again denote by A the case that the relation is fulfilled by the spatiotemporal values, we also obtain a series in which A has a relative frequency of almost 1, and for which a limit at 1 is

inferred. The contrary cases \bar{A} (non- A) include those cases, never wholly to be eliminated, in which the observations do not fit into the mathematical relation. As the observations of both examples relate to not one but numerous series of experiments, we have to represent them by the following schema:



It is the discovery of Newton that a formula may be given which includes the observations of both Galileo and Kepler; we may therefore consider the preceding scheme, consisting of two parts, as one undivided scheme for which the case A is defined by one mathematical relation only. It is the famous relation $k(m_1 m_2 / r^2)$ which does this; the case A may be regarded as meaning the correspondence of observations to this mathematical law, in both parts of the scheme.

With this recognition, the applicability of probability methods is greatly expanded. We are now able to apply cross-inductions leading from the Galilean lines of the scheme to the Keplerian lines, and inversely; i.e., the validity of Kepler's laws is no longer based on Kepler's observational material alone but jointly on Galileo's material, and conversely, the validity of Galileo's law is jointly supported by Kepler's observational material. Before Newton, similar cross-inductions were only possible within each section of the schema separately. Newton's discovery, therefore, in unifying both theories, involves an increase of certainty for both of them; it links a more com-

prehensive body of observational material together to form one inductive group.

The increase of certainty described corresponds to the conception of the men of science shown on the occasion of theoretical discoveries of this kind. Classical logic and epistemology could not assign any valid argument for this interpretation; it is only probability logic which, by the idea of concatenated inductions, is able to justify such a conception. We see that only in placing the causal structure of knowledge within the framework of probability do we arrive at an understanding of its essential features.

§ 42. The two kinds of simplicity

It might be objected to our interpretation that, logically speaking, Newton's discovery is trivial; if a finite set of observations of very different kinds is given, it is always mathematically possible to construct a formula which simultaneously embraces all the observations. In general such a formula would be very complicated, even so complicated that a human mind would not be able to discover it; it is the advantage of Newton's discovery that in this case a very simple formula suffices. But this, the objection continues, is all Newton did; Newton's theory is simpler, more elegant than others—but progress in the direction of truth is not connected with his discovery. Simplicity is a matter of scientific taste, a postulate of scientific economy, but has no relation to truth.

This kind of reasoning, well known from many a positivistic writer, is the outcome of a profound misunderstanding of the probability character of scientific methods. It is true that for any set of observations a comprehensive formula may be constructed, at least theoretically, and that Newton's formula is distinguished by simplicity from all the others. But this simplicity is not a matter of scien-

tific taste; it has on the contrary an inductive function, i.e., it brings to Newton's formula good predictive qualities. To show this, we must add a remark concerning simplicity.

There are cases in which the simplicity of a theory is nothing but a matter of taste or of economy. These are cases in which the theories compared are logically equivalent, i.e., correspond in all observable facts. A well-known case of this type is the difference of systems of measurement. The metrical system is simpler than the system of yards and inches, but there is no difference in their truth-character; to any indication within the metrical system there is a corresponding indication within the system of yards and inches—if one is true, the other is true also, and conversely. The greater simplicity in this case is really a matter of taste and economy. Calculations within the metrical system permit the application of the rules of decimal fractions; this is indeed a great practical advantage which makes the introduction of the metrical system desirable in those countries which still keep to the yard and inch system—but this is the only difference. For this kind of simplicity which concerns only the description and not the facts co-ordinated to the description, I have proposed the name *descriptive simplicity*. It plays a great role in modern physics in all those places where a choice between definitions is open to us. This is the case in many of Einstein's theorems; it is the reason the theory of relativity offers a great many examples of descriptive simplicity. Thus the choice of a system of reference which is to be called the *system in rest* is a matter of descriptive simplicity; it is one of the results of Einstein's ideas that we have to speak here of descriptive simplicity, that there is no difference of truth-character such as Copernicus believed. The question of the definition of simultaneity or of the

choice of Euclidean or non-Euclidean geometry are also of this type. In all these cases it is a matter of convenience only for which definition we decide.*

However, there are other cases in which simplicity determines a choice between nonequivalent theories. Such cases occur when a diagram is to be drawn which is determined by some physical measurements. Imagine that a physicist found by experiment the points indicated on Figure 6; he wants to draw a curve which passes through

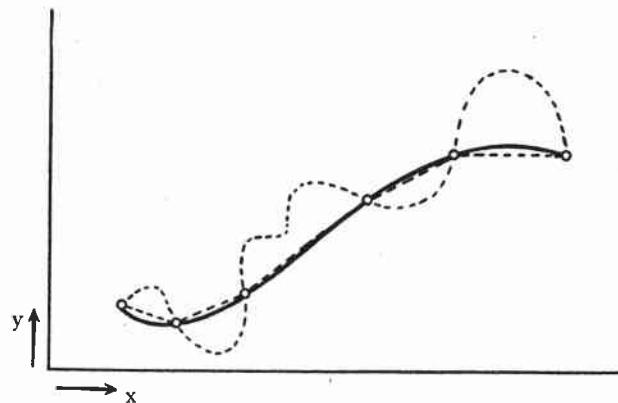


FIG. 6.—The simplest curve: inductive simplicity

the data observed. It is well known that the physicist chooses the simplest curve; this is not to be regarded as a matter of convenience. We have drawn in Figure 6, in addition to the simplest curve, one (the dotted line) which makes many oscillations between the observed points. The two curves correspond as to the measurements observed, but they differ as to future measurements; hence they signify different predictions based on the same observational material. The choice of the simplest curve, consequently, depends on an inductive assumption: we believe that the simplest curve gives the best predictions. In such

a case we speak of *inductive simplicity*; this concept applies to theories which differ in respect to predictions, although they are based on the same observational material. Or, more precisely speaking: The relation "difference as to inductive simplicity" holds between theories which are equivalent in respect to all observed facts, but which are not equivalent in respect to predictions.²⁵

The confusion of both kinds of simplicity has caused much mischief in the field of the philosophy of science. Positivists like Mach have talked of a principle of economy which is to replace the aim of truth, supposedly followed by science; there is, they say, no scientific truth but only a most economical description. This is nothing but a confusion of the two concepts of simplicity. The principle of economy determines the choice between theories which differ in respect to descriptive simplicity; this idea has been erroneously transferred to cases of inductive simplicity, with the result that no truth is left at all but only economy. Actually in cases of inductive simplicity it is not economy which determines our choice. The regulative principle of the construction of scientific theories is the postulate of the best predictive character; all our decisions as to the choice between unequivalent theories are determined by this postulate. If in such cases the question of simplicity plays a certain role for our decision, it is because we make the assumption that the simplest theory furnishes the best predictions. This assumption cannot be justified by convenience; it has a truth-character and demands a justification within the theory of probability and induction.

Our theory of induction enables us to give this justifica-

²⁵ The terms "descriptive simplicity" and "inductive simplicity", have been introduced in the author's *Axiomatik der relativistischen Raum-Zeit-Lehre* (Braunschweig, 1924), p. 9. A further elucidation of these concepts has been given in the author's *Ziele und Wege der physikalischen Erkenntnis* in *Handbuch der Physik*, ed. Geiger-Scheel (Berlin, 1929), IV, 34-36.

tion. We justified the inductive inference by showing that it corresponds to a procedure the continued application of which must lead to success, if success is possible at all. The same idea holds for the principle of the simplest curve.

What we want to construct with the diagram is a continuous function determining both past and future observations, a mathematical law of the phenomena. Keeping this aim before our eyes, we may give a justification of the procedure of the simplest curve, by dividing our reasoning into two steps.

In the first step, let us imagine that we join the observed points by a chain of straight lines, such as drawn in Figure 6. This must be a first approximation; for if there is a function such as we wish to construct, it must be possible to approximate it by a chain of straight lines. It may be that future observation will show too much deviation; then, we shall correct our diagram by drawing a new chain of straight lines, including the newly observed points. This procedure of preliminary drawing and later correction must lead to the true curve, if there is such a curve at all—its applicability is a necessary condition of the existence of a law determining the phenomena.

It is the method of anticipation which is adopted with such a procedure. We do not know whether our observed points are sufficiently dense to admit a linear approximation to the curve; but we anticipate this case, being ready to correct our posit if later observations do not confirm it. At some time we shall have success with this procedure—if success is attainable at all.

But the chain of straight lines does not correspond to the actual procedure applied by the physicist. He prefers a smooth curve, without angles, to the chain of straight lines. The justification of this procedure necessitates a second step in our considerations.

For this purpose we must consider the derivatives of the function represented by the curve. The differential quotients of a function are regarded in physics as physical entities, in the same sense as is the original entity represented by the function; thus, if the original entity is a spatial distance represented as a function of time, the first derivative is a velocity, the second an acceleration, etc. For all these derived entities we aim also to construct mathematical laws; we want to find for them also continuous functions such as are sought in our diagram. Regarding the chain of straight lines from this point of view, it already fails for the first derivative; in this case the first differential quotient, designed as a function of the argument x , is not represented by a continuous curve but by a discontinuous chain of horizontal lines. This may be illustrated by Figure 7, the dotted lines of which correspond to the first derivative of the chain of straight lines of Figure 6; we see that we do not obtain here even a continuous chain of straight lines but a chain broken up into several parts. Thus, if we approximate the original curve by a chain of straight lines, the principle of linear approximation is followed only as to the original curve; for the first derivative it is already violated. This is different, however, for the smooth curve; its derivatives, conceived as functions of x , are smooth curves as well. This may be seen in Figure 7, where the first derivative of the smooth curve of Figure 6 is represented by the continuous smooth line. This is the reason for the preference of the smooth curve. It has, in respect to the set of observed points, qualities similar to those of a linear interpolation and may be justified by the principle of anticipation as well; moreover, it also satisfies the same postulate for its derivatives.

The procedure of the smoothest interpolation may be considered, therefore, as a superposition of linear interpo-

lations carried through for the construction of the original function and of its derivatives. Thus the nonlinear interpolation by the smoothest curve may be justified by a reduction to linear interpolations which determine, on the whole, a nonlinear interpolation to be preferable. The procedure corresponds not to a single induction but to a concatenation of inductions concerning different functions standing in the mutual relation of a function and its derivative; the result is a better induction, as it is based on a repeated application of the inductive principle, and incorporates corrections in the sense defined in § 41.

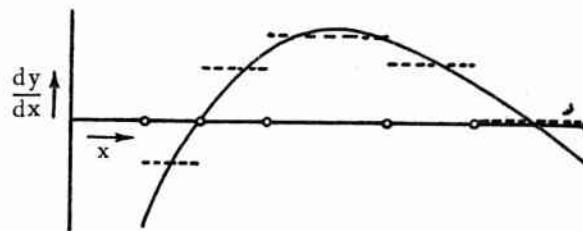


FIG. 7.—Derivatives of the simplest curve, and of the chain of straight lines, developed from Fig. 6.

There remains an objection to our reasoning. We contrived to justify the preference of the smooth curve to the chain of straight lines; but the postulate of the smooth curve is not unambiguous. Though a curve such as drawn in the dotted line of Figure 6 is excluded, there remain other smooth curves very similar to the one drawn; the points observed will not furnish us a clear decision as to the choice between such similar smooth curves. Which are we to choose?

Here we must answer that the choice is not relevant. From the viewpoint of approximation, there is no great difference as to these forms of curves; all of them converge asymptotically; they do not differ essentially as far as

predictions are concerned. The choice between them may therefore be determined from the viewpoint of convenience. The principle of inductive simplicity determines the choice only to a certain extent: it excludes the oscillating curve drawn in Figure 6, but there remains a small domain of indeterminacy within which the principle of descriptive simplicity may be applied. We prefer here a simpler analytical expression because we know better how to handle it in a mathematical context; this is permissible because the functions open to our choice do not relevantly differ as to predictions of further observations between the points observed.

To the latter argument a further objection may be raised. It is true that within the domain of the observed points there is no great difference between all these smooth curves; but this is no longer valid outside this domain. All analytical functions define a prolongation of the curve into a distant domain, and two analytical functions which differ only slightly within the interior domain, may lead to great differences as to extrapolations. Consequently the choice between them can not be justified by descriptive simplicity as far as extrapolations are concerned; how, then, may we justify this choice?

To this we must answer that a set of observations does not at all justify an extrapolation of any considerable length. The desire to know the continuation of the curve far beyond the observed domain may be very strong with the physicist; but, if he has nothing but the observed set at his disposal, he must renounce any hypothesis concerning extrapolations. The inductive principle is the only rule the physicist has at hand; if it does not apply, philosophy cannot provide him with a mysterious principle showing the way where induction fails—in such a case, there remains nothing but to confess a modest *ignoramus*.

Our adversary might object that the man of science does not always comply with this alternative. Only the spirit of mediocrity will submit to renunciation, he will exclaim; the scientific genius does not feel bound to the narrow restrictions of induction—he will guess the law outside the domain of observed facts, even if your principle of induction cannot justify his presentiments. Your theory of induction as an interpolation, as a method of continual approximation by means of anticipations, may be good enough for the subordinate problems of scientific inquiry, for the completion and consolidation of scientific theories. Let us leave this task to the artisans of scientific inquiry—the genius follows other ways, unknown to us, unjustifiable a priori, but justified afterward by the success of his predictions. Is not the discovery of Newton the work of a genius which never would have been achieved by methods of simple induction? Is not Einstein's discovery of new laws of the motion of planets, of the bending of light by gravitation, of the identity of mass and energy, etc., a construction of ideas which has no relation to diagrams of curves of interpolation, to statistics of relative frequencies, to the slow driving of approximations, step by step?

Let me say that I should be the last to discredit the work of the great men of science. I know as well as others that the working of their minds cannot be replaced by directions for use of diagrams and statistics. I shall not venture any description of the ways of thought followed by them in the moments of their great discoveries; the obscurity of the birth of great ideas will never be satisfactorily cleared up by psychological investigation. I do not admit, however, that these facts constitute any objection against my theory of induction as the only means for an expansion of knowledge.

We pointed out in the beginning of our inquiry (§ 1) the

distinction between the context of discovery and the context of justification. We emphasized that epistemology cannot be concerned with the first but only with the latter; we showed that the analysis of science is not directed toward actual thinking processes but toward the rational reconstruction of knowledge. It is this determination of the task of epistemology which we must remember if we want to construct a theory of scientific research.

What we wish to point out with our theory of induction is the logical relation of the new theory to the known facts. We do not insist that the discovery of the new theory is performed by a reflection of a kind similar to our exposition; we do not maintain anything about the question of how it is performed—what we maintain is nothing but a relation of a theory to facts, independent of the man who found the theory. There must be some definite relation of this kind, or there would be nothing to be discovered by the man of science. Why was Einstein's theory of gravitation a great discovery, even before it was confirmed by astronomical observations? Because Einstein saw—as his predecessors had not seen—that the known facts indicate such a theory; i.e., that an inductive expansion of the known facts leads to the new theory. This is just what distinguishes the great scientific discoverer from a clairvoyant. The latter wants to foresee the future without making use of induction; his forecast is a construction in open space, without any bridge to the solid domain of observation, and it is a mere matter of chance whether his predictions will or will not be confirmed. The man of science constructs his forecast in such a way that known facts support it by inductive relations; that is why we trust his prediction. What makes the greatness of his work is that he sees the inductive relations between different elements in the system of knowledge where other people did not see them;

but it is not true that he predicts phenomena which have no inductive relations at all to known facts. Scientific genius does not manifest itself in contemptuously neglecting inductive methods; on the contrary, it shows its supremacy over inferior ways of thought by better handling, by more cleverly using the methods of induction, which always will remain the genuine methods of scientific discovery.

That there is an inductive relation from the known facts to the new theory becomes obvious by the following reflection. The adherents of the contrary opinion believe that the construction of the new theory is due to a kind of mystic presentiment but that later, after a confirmation of the predictions contained in the new theory, it is proved to be true. This is, however, nothing but one of the unwarranted schematizations of two-valued logic. We shall never have a definitive proof of the theory; the so-called confirmation consists in the demonstration of some facts which confer a higher probability upon the theory, i.e., which allow rather simple inductive inferences to the theory. The situation before the confirmation differs from that after it only in degree. This situation is characterized by the occurrence of some facts which confer at least some probability upon the theory and which distinguish it from others as our best posit, according to inductive methods. This is what the good theorist sees. If there were no such inductive relations, his supposition would be a mere guess, and his success due to chance only.

We may add the remark that the distinction of the context of justification from the context of discovery is not restricted to inductive thinking alone. The same distinction applies to deductive operations of thought. If we are faced by a mathematical problem, say, the construction of a triangle from three given parameters, the solution (or the

class of solutions) is entirely determined by the given problem. If any solution is presented to us, we may decide unambiguously and with the use of deductive operations alone whether or not it is correct. The way in which we find the solution, however, remains to a great extent in the unexplored darkness of productive thought and may be influenced by aesthetic considerations, or a "feeling of geometrical harmony." From the reports of great mathematicians it is known that aesthetic considerations may play a decisive role in their discoveries of great mathematical theorems. Yet in spite of this psychological fact, no one would propound a philosophical theory that the solution of mathematical problems is determined by aesthetic points of view. The objective relation from the given entities to the solution and the subjective way of finding it are clearly separated for problems of a deductive character; we must learn to make the same distinction for the problem of the inductive relation from facts to the theories.

There are cases, it is true, in which a clear decision as to the most favorable theory cannot be obtained because there are several theories with equal weights indicated by the facts. This does not mean that we are at a loss with the inductive principle; on the contrary, a great number of theories is always ruled out by this principle. But among the weights of the admissible remainder there may be no maximum, or so flat a maximum that it cannot be considered as furnishing the basis for a clear decision. In such cases, which we may call cases of *differential decision*,²⁶ different men of science will decide for different

²⁶ I choose this name by analogy with the term "differential diagnosis" used by physicians, to denote a case where the observed symptoms of illness indicate several diseases as their possible origin but do not permit a decision among the members of this group unless certain new symptoms can be observed. This differential diagnosis is, logically speaking, a special case of our differential decision.

theories, their decisions being determined by personal taste more than by scientific principles; the final decision will then be made by later experiments of a crucial character. It is a kind of "natural selection," of "struggle for existence," which determines in such a case the final acceptance of a scientific theory; though this case happens, and not too rarely, we must not forget that this is just a case in which scientific prophecy breaks down, the decision in favor of an assumption being possible only after the occurrence of the predicted events. The man who predicted the right theory is then sometimes considered a great prophet because he knew the true prediction even in a case when scientific principles of prediction failed. But we must not forget that his success is the success of a gambler who is proud, having foreseen the *rouge* or the *noir*. This presumed prophetic gift will always expose its spurious nature in a second case of prediction when success will be wanting. A man of science, in the case of differential decision, had better admit that he cannot rationally make his choice.

In the context of our introduction of the concept of inductive simplicity, we illustrated its meaning by a diagram and pointed out a smooth curve as the model of this kind of simplicity. However, this is not the only case of this kind. The inductive connections of modern physics are constructed analytically; this is why the theorist of physics must be a good mathematician.²⁷

The inductive procedure of Newton consisted in his demonstration that a simple mathematical formula covers both Galileo's and Kepler's laws. The simplicity of the formula expresses its character as an interpolation, as a linear,

²⁷ We may add that the graphical interpretation of inductive inferences may be also carried through, for complicated cases, if we pass to a parameter space of a higher number of dimensions (cf. the author's article, "Die Kausalbehauptung und die Möglichkeit ihrer empirischen Nachprüfung," *Erkenntnis*, III [1932], 32).

or almost linear, approximation; it is this quality to which its predictive qualities are due. Newton's theory not only incorporates the observations of Galileo and Kepler but also leads to predictions. The "predictions" may concern phenomena which are already known, but which were neither seen before in connection with the other phenomena nor used as a part of the basis on which the new theory was constructed. Of such a kind was Newton's explanation of the tides. On the other hand, Newton's theory led also to predictions, properly speaking, e.g., the attraction of a ball of lead to other bodies such as observed by Cavendish in the turning of a torsion balance.

We raised the question whether in a diagram an extrapolation is possible which extends to a domain rather far removed from the domain of the points observed. There are examples in which extrapolations of such a kind seem to occur. Such cases, however, are to be otherwise explained; there are facts of another type, not belonging to the domain of the observation points marked in the diagram, which support the extrapolation. Examples of this kind are cases in which the analytical form of the curve is known to the physicist before the observations, and these are made only to determine the numerical constants of the analytical expression. This case, which happens rather frequently in physics, corresponds in our example to a determination of the curve by facts outside the observed domain; for the analytical form of the curve is then determined by reflections connecting the phenomenon in question to other phenomena.

An example of a similar type is Einstein's prediction of the deviation of light rays emitted by stars in the gravitational field of the sun. Had he pursued only the plan of finding a generalization of Newton's law of planetary motion such that the irregularities of the planet Mercury

would have been explained, his hypothesis of the deviation of light would have been an unwarranted extrapolation not justified by inductions. But Einstein saw that a much more comprehensive body of observations was at his disposal, which could be interpolated by means of the idea that a gravitational field and an accelerated motion are always equivalent. From this "equivalence principle" the deviation of light rays followed immediately; thus within the wider context Einstein's prediction was the "smoothest interpolation." It is this quality which is denoted in predicates frequently applied to Einstein's theories, such as "the natural simplicity of his assumptions"; such predicates express the inductive simplicity of a theory, i.e., its character of being a smooth interpolation. This does not diminish the greatness of Einstein's discovery; on the contrary, it is just his having seen this relation which distinguishes him from a clairvoyant and makes him one of the most admirable prophets within the frame of scientific methods. The gift of seeing lines of smooth interpolation within a vast domain of observational facts is a rare gift of fate; let us be glad we have men who are able to perform in respect to the whole domain of knowledge inferences whose structure reappears in the modest inferences which the artisan of science applies in his everyday work.

S 43. The probability structure of knowledge

Our discussion of the methods of scientific research and of the formation of scientific theories has led us to the result that the structure of scientific inferences is to be conceived as a concatenation of inductive inferences. The elementary structure of the concatenation is the probability lattice; we may refer here to the exposition of this form of inference in § 41. As a consequence of idealizations, in which the transition from probability to practical

truth plays a decisive role, the probability character of the inferences is not always easily seen; the short steps of inductive inferences can be combined into long chains forming longer steps of so complicated a structure that it may be difficult to see the inductive inference as the only atomic element in them. To indicate the method of decomposition of such structures, and of their reduction to inductive inferences, we may here add a discussion of some examples.

There are cases in which one experiment may decide the fate of a theory. Such cases of an *experimentum crucis* are often quoted against the inductive conception of science; they seem to prove that it is not the number of instances which decides in favor of a theory but something such as an "immediate insight into the very nature of the phenomenon," opened for us by one single experiment. On a closer consideration, the procedure is revealed as a special case of concatenated inductions. We may know from previous experience that only two possibilities are left for a certain experiment, i.e., we may know, with great probability, that *A* will be followed by *B* or by *C* and, besides, that there is a great probability that *A* will always be followed by the same type of event, not alternately by both. In such a case, if the probabilities occurring are high, one experiment may indeed suffice for the decision. Of such a type was Lavoisier's decisive experiment concerning combustion. There were in practice only two theories left as an explanation of combustion: the first maintained that a specific substance, phlogiston, escaped during the combustion; the second assumed that a substance originating from air entered the burning body during the combustion. Lavoisier showed in a famous experiment that the body was heavier after being burnt than before; thus one experiment could decide in favor of the oxidation theory of

combustion. Yet this was possible only because former inductions had excluded all but two theories and because former inductions had made it very probable that all processes of combustion are of the same type. Thus the *experimentum crucis* finds its explanation in the theory of induction and does not involve further assumptions; it is only the superimposition of a great many elementary inductive inferences which creates logical structures whose form as a whole, if we cling to a schematized conception, suggests the idea of noninductive inference.

It is the great merit of John Stuart Mill to have pointed out that all empirical inferences are reducible to the *inductio per enumerationem simplicem*. The exact proof, however, has been achieved only by the demonstration that the calculus of probability can be reduced to this principle, a demonstration which presupposes an axiomatic construction of the calculus of probability. Physics applies in its inferences, besides logic and mathematics in general, the methods of the calculus of probability; thus an analysis of the latter discipline was as necessary for epistemology as an analysis of logic and the general methods of mathematics.

It is on account of this foundation of probability inferences on the principle of induction that we are entitled to interpret the inferences leading from observations to facts as inductive inferences. Inferences appearing in the form of the schemas developed within the calculus of probability are reducible, for this reason, to inductive inferences. Of this kind are many inferences which, on superficial examination, show no probability character at all but look like a decision concerning an assumption, based on an observation of its "necessary consequences." If a detective infers from some fingerprints on a bloody knife that Mr. X is a murderer, this is usually justified by saying: It is impos-

sible that another man should have the same fingerprints as Mr. X; it is impossible that the bloody knife lying beside the dead body of the victim was not used to kill the man, under the given conditions, and so on. These so-called impossibilities are, however, only very low probabilities, and the whole inference must be considered as falling under the rule of Bayes, one of the well-known schemas of the calculus of probability which is used for inferring from given observations the probabilities of their causes. It furnishes, consequently, not a certainty but only a high probability for the assumption in question.

Scientific inferences from observations to facts are of the same type. If Darwin maintained the theory that the logical order of organisms according to the differentiation of their internal structure, may be interpreted as the historical order of the development of the species, this theory is based on facts such as the correspondence of the time order of geological layers (determined by their lying one above the other) to the occurrence of higher organisms. With the assumption of a theory which considers the higher organisms as old as the lowest ones, this correspondence would appear as a very improbable result. Conversely, according to Bayes's rule the observed fact makes Darwin's theory probable and the other theory improbable. The probability character of this inference is usually veiled by the use of statements such as, "The other theory is incompatible with the observed facts," a statement in which the transition from a low probability to impossibility is performed; and epistemological conceptions have been developed according to which a theory is unambiguously tested by its consequences. A trained eye nevertheless discovers probability structures in all these inferences from facts to theories. With this analysis the reduction of the inferences occurring to inductive inferences is also performed, owing to the

reducibility of the calculus of probability to the inductive principle. This is the reason why we may say that scientific inferences from facts to theories are inductive inferences.

Scientific induction is not of a form "higher" than the ordinary inductions of daily life; but it is better in the sense of a difference of degree. This difference is due to the concatenation of inductions such as expressed in the application of the rules of the calculus of probability; they lead to results which by direct inductions would never be attained. We said that the inductive nature of these inferences is sometimes obscured by a schematization in which probability implications are replaced by strict implications; this may be illustrated by another example. Some philosophers have distinguished a generalizing from an exact induction; the first is to be our poor frequency-bound induction, which is restricted to probabilities only; whereas the second is to be a higher method of cognition which, though based on experience, is to lead to absolute certainty. I may refer here to a discussion I once had with a biologist of high rank, who refused to admit that his science is dependent on so imperfect a principle as *inductio per enumerationem simplicem*. He presented to me an example concerning carnivorous and herbivorous animals. We observe, he argued, that the first have a short intestine, the latter a long one; we infer then by generalizing induction that there is a causal connection between the food and the length of the intestine. This is only a mere supposition, he said; yet it is proved later by exact induction at the moment we succeed in experimentally changing the length of the intestine by the food we give to the animal. Such experiments have indeed been successfully performed with tadpoles.²⁸ But what is overlooked in such reasoning is

²⁸ Cf. Max Hartmann, "Die methodologischen Grundlagen der Biologie," *Erkenntnis*, III (1932-33), 248.

that the difference in question is nothing but a difference of degree. The experiments with tadpoles enlarge the observational material, and precisely in a direction which permits us to make use of certain laws well established by previous inductions, such as the law that food has an influence on the development of the organism, that the other conditions in which the animals were kept do not influence in general their intestines, and the like. I do not say this to depreciate the work of biologists; on the contrary, the progress of knowledge from lower probabilities to higher ones is due to experiments of such a kind. There is no reason though to construct a qualitative difference of methods where quantitative differences are in question. What the experimental scientist does is to construct conditions in which all of the processes occurring except the one which is to be tested are conformable to known cases; by this isolation of the unknown phenomenon from other unknown phenomena he arrives at simpler forms of the inductive inference. As to the interpretation of this procedure we must take care not to confound an idealization with the inferences actually occurring. If we consider those high probabilities occurring as equal to 1, we transform the actual procedure into a schema in which "causal connections" occur, and in which one experiment may demonstrate with certainty some new "causal law." To infer from the applicability of such a scheme the existence of an "exact induction" which is to be of a logical type different from the ordinary induction, means overstraining an approximation and drawing conclusions which are valid for the schema only and not for the real procedure to which it applies.

Any epistemology which forces knowledge into the frame of two-valued logic is exposed to this danger. It was the grave mistake of traditional epistemology to consider knowledge as a system of two-valued propositions; it is to

this conception that all kinds of apriorism are due, these being nothing but an attempt to justify an absolutely certain knowledge of synthetic character. And it is to this conception also that all kinds of skepticism are due, renunciation of truth being the attitude of more critical minds before the problem of such absolute knowledge. The way between Scylla and Charybdis is pointed out by the probability theory of knowledge. There is neither an absolutely certain knowledge nor an absolute ignorance—there is a way between them pointed out by the principle of induction as our best guide.

If we say that the two-valued logic does not apply to actual knowledge, this is not to maintain that it is false. It is to affirm only that the conditions of its application are not realized. Scientific propositions are not used as two-valued entities but as entities having a weight within a continuous scale; hence the presuppositions of two-valued logic are not realized in science. Treating science as a system of two-valued propositions is like playing chess on a board whose squares are smaller than the feet of the pieces; the rules of the game cannot be applied in such a case because it remains indeterminate on which square a piece stands. Similarly, the rules of two-valued logic cannot be applied to scientific propositions, at least not generally, because there is no determinate truth-value corresponding to the propositions, but only a weight. It is therefore probability logic alone which applies to knowledge in its general structure.

Is there no way, we may be asked, to escape this consequence? Is there no way of transforming probability logic into the two-valued logic? As to the answer to this question, we may make use of our inquiries concerning this transformation (§ 36). We showed that there are two ways for making such a transformation. The first one is the way

of dichotomy or trichotomy; we found that this way can only lead to an approximative validity of the two-valued logic. The second way makes use of the frequency interpretation; however it is also restricted to an approximative validity for two reasons: first, because the individual element of the propositional series is not strictly true or false and, second, because the frequency to be asserted cannot be asserted with certainty. It is this latter point of view which we must now analyze more accurately.

The transition under consideration can be conceived, if we use the logical conception of probability (§ 33), as a transition from probability statements to statements about the probability of other statements; but it would be erroneous to believe that in this way we could arrive at a strict logic of two values. A statement about the probability of another statement is in itself not true or false but is only given to us with a determinate weight. Using the transition in question, we shall never arrive at something other than probabilities. We are bound to this flight of steps leading from one probability into another. It is only a schematization if we stop at one of the steps and regard the high probability obtained there as truth. It was a schematization, therefore, when we spoke throughout our inquiry of the predicate of weight; we should have spoken of an infinite set of weights of all levels co-ordinated to a statement. We may refer here to our numerical example (§ 41) in which we calculated the probability 0.75 for a statement of the first level and the probability 0.59 for the statement of the second level that the first statement has the probability 0.75; in this example, we cut off the flight at the second step. This was also a schematization, owing to the simplified conditions in which the problem was given; an exhaustive consideration would have to take into account all probabilities of the infinitely numerous levels.

From the example given we also see another feature of the probability structure of knowledge: that the probabilities occurring are by no means all either of a high or of a low degree. There are intermediate degrees as well; their calculation may be based on the frequency of elementary propositions whose probability is near to the extreme values 1 or 0 (cf. § 36)—but the propositions to which these probabilities are co-ordinated as weights enter into the system of knowledge as propositions of an intermediate degree of weight. This is why for the whole of science two-valued logic does not even apply in the sense of an approximation. An approximative application of two-valued logic obtains only if we consider not the direct propositions of science but those of the second or a higher level—propositions about the probability of direct propositions of science.²⁹

The occurrence of different probabilities of higher levels is a specific feature of probability logic; two-valued logic shows this feature only in a degenerate form. Our probability of the second level would correspond in the two-valued logic to the truth of the sentence, “The sentence a is true”; but if a is true, then “ a is true” is true also. Thus we need not consider the truth-values of higher levels in the two-valued logic; this is why this problem plays no role in traditional logic or logistic. In probability logic, on the other hand, we cannot dispense with considerations of this

²⁹ In our preceding inquiries we frequently made use of the approximate validity of two-valued logic for the second-level language. One schematization of this kind is that we considered statements about the weight of a proposition as being true or false; another one is contained in our use of the concepts of physical and logical possibility, occurring in our definitions of meaning. Strictly speaking, there is, between these types of possibility, a difference of degree only. We were entitled to consider them in a schematized form as qualitatively different because they concern reflections belonging to the second-level language. The approximate validity of two-valued logic for the second-level language also explains why the positivistic language can be conceived as approximately valid in the sense of a second-level language (cf. the remark at the end of § 17).

sort; that is why the application of probability logic to the logical structure of science is a rather complicated matter.

These reflections become relevant if we want to define the probability of a scientific theory. This question has attained some significance in the recent discussion of the probability theory of knowledge. The attempt has been made to show that probability logic is not a sufficiently wide framework to include scientific theories as a whole. Only for simple propositions, it has been said, may a probability be determined; for scientific theories we do not know a definite probability, and we cannot determine it because there are no methods defining a way for such a determination.

This objection originates from underrating the significance of the probabilities of higher levels. We said it was already a schematization if we spoke of *the* probability, or *the weight* of a simple proposition; this schematization, however, is permissible as a sufficient approximation. But this no longer holds if we pass from simple propositions to scientific theories. For example, there is no such thing as *the* probability of the quantum theory. A physical theory is a rather complex aggregate; its different components may have different probabilities which should be determined separately. The probabilities occurring here are not all of the same level. To a scientific theory belongs, consequently, a set of probabilities, including probabilities of the different parts of the theory and of different levels.³⁰

Within the analysis of the problem of the probability of theories, one question in particular has stood in the foreground of discussion. It has been asked whether the probability of a theory concerns the facts predicted by the

³⁰ These different probabilities cannot in general be mathematically combined into one probability; such a simplification presupposes special mathematical conditions which would apply, if at all, only to parts of the theory (cf. *Wahrscheinlichkeitslehre*, § 58).

theory, or whether we have to consider the theory as a sociological phenomenon and to count the number of successful theories produced by mankind. The answer is that both kinds of calculation apply but that they correspond to different levels. The quantum theory predicts a great many phenomena, such as observations on electrometers and light rays, with determinate probabilities; as the theory is to be considered as the logical conjunction of propositions about these phenomena, its probability may be determined as the arithmetical product of these elementary probabilities. This is the probability of the first level belonging to the quantum theory. On the other hand, we may consider the quantum theory as an element in the manifold of theories produced by physicists and ask for the ratio of successful theories within this manifold. The probability obtained in this way is to be interpreted not as the direct probability of the quantum theory but as the probability of the assumption, "The quantum theory is true"; as the truth occurring here is not strict truth but only a high probability, namely, that of the first level, the probability of the second level is independent of that of the first and demands a calculation of its own. We see that at least two probabilities of different levels play a role in questions about theories; we might construct still more, considering other kinds of classification of the theory. If we add a consideration of the fact that the parts of a theory may already belong to different levels, we see that a theory within the probability theory of knowledge is not characterized by a simple weight but by a set of weights partially comprising weights of the same, partially of different, levels.

The practical calculation of the probability of a theory involves difficulties, but it would be erroneous to assume that our conception lacks any practical basis. It is true

that the probability of theories of a high generality is usually not quantitatively calculated; but as soon as determinations of a numerical character occur within science, such as those concerning physical constants, they are combined with calculations which may be interpreted as preliminary steps toward the calculation of the probability of a theory. It is the application of the mathematical theory of errors in which considerations of this kind find their expression. The "average error" of a determination may be interpreted, according to well-known results of the calculus of probability,³¹ as the limits within which the deviation of future observations will remain with the probability 2/3; thus this indication may be conceived as the calculation of a first-level probability of an assumption. If we say that "the velocity of light is 299796 km/sec, with an average error of ± 4 , or of ± 0.0015 per cent"³² this may be read: "The probability that the velocity of light lies between 299792 km/sec and 299800 km/sec, is 2/3." It can easily be shown that we may infer from this a lower limit for the probability (on the first level) of Einstein's hypothesis of the constancy of the velocity of light; passing to somewhat wider limits of precision and applying some properties of the Gaussian law, we may state this result in this form: "The probability of Einstein's hypothesis of the constancy of the velocity of light is greater than 99.99 per cent, if a numerical range of 0.0052 per cent is admitted for the possible value of the constant." Considerations of a similar type may be carried through for theories of a more comprehensive character.

As to probabilities of the second level, we cannot as yet determine their numerical values. It has been objected that we here meet a difficulty of principle because we do

³¹ Cf. *ibid.*, p. 226.

³² A. A. Michelson, *Astrophysical Journal*, LXV (1927), 1.

not know into which class the theory is to be incorporated if we want to determine its probability in the frequency sense; thus if we want to determine the second-level probability of the quantum theory, shall we consider the class of scientific theories in general, or only that of physical theories, or only that of physical theories in modern times? I do not think that this is a serious difficulty, as the same question occurs for the determination of the probability of single events; I have indicated in § 34 the method of procedure in such a case. The narrowest class available is the best; it must, however, be large enough to afford reliable statistics. If the probability of theories (of the second level) is not yet accessible to a quantitative determination, the reason is to be found, I think, in the fact that we have in this field no sufficiently large statistics of uniform cases. That is to say, if we use a class of cases not too small in number, we may easily indicate a subclass in which the probability is considerably different. We know this from general considerations, and thus we do not try to make statistics. Future statistics may perhaps overcome these difficulties, as the similar difficulties of meteorological statistics have been overcome. As long as we have no such statistics, crude appraisals will be used in their place—as in all fields of human knowledge not yet accessible to satisfactory quantitative determinations. Appraisals of this kind (concerning the second-level probability of a theory) may acquire practical importance in cases when we judge a theory by the success obtained with other theories in that domain; if an astronomer propounds a new theory of the evolution of the universe, we hesitate to trust this theory on account of unfortunate experiences with other theories of that kind.

A last objection remains. We said that a theory, and even a simple proposition, is characterized not by a single

weight but by a set of weights infinite in number. We must in any case confine ourselves to a finite number of members. This would be justified if all the following members should be weights of the degree 1; we might then consider the last weight used as truly determined. But, if we know nothing about all the rest of the set, how can we omit all of them? How can we justify using the weights of the lower levels if we do not know anything about the weights of the higher levels?

To see the force of this objection, let us imagine the case that all the rest of the weights are of a very low degree—near zero. This would result in the last weight determined by us being unreliable; the preceding weight would consequently become unreliable as well, and, as this unreliability is equally transferred to the weight of the first level, the whole system of weights would be worthless. How can we justify our theory of weights, and with this the probability procedure of knowledge, before the irrefutable possibility of such a case?

This objection is nothing but the well-known objection to which the procedure of induction is already exposed in its simplest form. We do not know whether we shall have success in laying our wager corresponding to the principle of induction. But we found that, as long as we do not know the contrary, it is advisable to wager—to take our chance at least. We know that the principle of induction determines our best wager, or posit, because this is the only posit of which we know that it must lead to success if success is attainable at all. As to the system of concatenated inductions, we know more: we know that it is better than any single induction. The system, as a whole, will lead to success earlier than a single induction; and it may lead to success even if some single inductions should remain without success. This logical difference, the superi-

ority of the net of concatenated inductions to single inductions, can be demonstrated by purely mathematical considerations, i.e., by means of tautologies; hence our preference for the system of inductions can be justified without any appeal to presuppositions concerning nature. It is very remarkable that such a demonstration can be given; although we do not know whether our means of prediction will have any success, yet we can establish an order between them and distinguish one of them, the system of concatenated inductions, as the best. With this result the application of the system of scientific inductions finds a justification similar to, and even better than, that of the single induction: *the system of scientific inductions is the best posit we know concerning the future.*

We found that the posits of the highest level are always blind posits; thus the system of knowledge, as a whole, is a blind posit. Posits of the lower levels have appraised weights; but their serviceableness depends on the unknown weights of the posits of higher levels. The uncertainty of knowledge as a whole therefore penetrates to the simplest posits we can make—those concerning the events of daily life. Such a result seems unavoidable for any theory of prediction. We have no certainty as to foreseeing the future. We do not know whether the predictions of complicated theories, such as the quantum theory or the theory of albumen molecules, will turn out to be true; we do not even know whether the simplest posits concerning our immediate future will be confirmed, whether they concern the sun's rising or the persistence of the conditions of our personal environment. There is no principle of philosophy to warrant the reliability of such predictions; that is our answer to all attempts made within the history of philosophy to procure for us such certainty, from Plato, through all varieties of theology, to Descartes and Kant.

In spite of that, we do not renounce prediction; the arguments of skeptics like Hume cannot shake our resolution: at least to *try* predictions. We know with certainty that among all procedures for foreseeing the future, known to us as involving success if success is possible, the procedure of concatenated inductions is the best. We try it as our best posit in order to have our chance—if we do not succeed, well, then our trial was in vain.

Is this to say that we are to renounce any belief in success? There is such a belief; everyone has it when he makes inductions; does our solution of the inductive problem oblige us to dissuade him from this firm belief?

This is not a philosophical but a social question. As philosophers we know that such a belief is not justifiable; as sociologists we may be glad that there is such a belief. Not everyone is likely to act according to a principle if he does not believe in success; thus belief may guide him when the postulates of logic turn out to be too weak to direct him.

Yet our admission of this belief is not the attitude of the skeptic who, not knowing a solution of his own, permits everyone to believe what he wants. We may admit the belief because we know that it will determine the same actions that logical analysis would determine. Though we cannot justify the belief, we can justify the logical structure of the inference to which it fortunately corresponds as far as the practical results are concerned. This happy coincidence is certainly to be explained by Darwin's idea of selection; those animals were to survive whose habits of belief corresponded to the most useful instrument for foreseeing the future. There is no reason to dissuade anybody from doing with belief something which he ought to do in the same way if he had no belief.

This remark does not merely apply to the belief in induction as such. There are other kinds of belief which have

crystallized round the methods of expanding knowledge. Men of scientific research are not always of so clear an insight into philosophical problems as logical analysis would require: they have filled up the world of research work with mystic concepts; they talk of "instinctive presentiments," of "natural hypotheses," and one of the best among them told me once that he found his great theories because he was convinced of the harmony of nature. If we were to analyze the discoveries of these men, we would find that their way of proceeding corresponds in a surprisingly high degree to the rules of the principle of induction, applied however to a domain of facts where average minds did not see their traces. In such cases, inductive operations are imbedded within a belief which as to its intension differs from the inductive principle, although its function within the system of operations of knowledge amounts to the same. The mysticism of scientific discovery is nothing but a superstructure of images and wishes; the supporting structure below is determined by the inductive principle.

I do not say this with the intention to discredit the belief—to pull the superstructure down. On the contrary, it seems to be a psychological law that discoveries need a kind of mythology; just as the inductive inference may lead us in certain cases to the preference of methods different from it, it may lead us also to the psychological law that sometimes those men will be best in making inductions who believe they possess other guides. The philosopher should not be astonished at this.

This does not mean that I should advise him to share any of these kinds of belief. It is the philosopher's aim to know what he does; to understand thought operations and not merely to apply them instinctively, automatically. He wants to look through the superstructure and to discover the supporting structure. Belief in induction, belief in a uniformity of the world, belief in a mystic harmony be-

tween nature and reason—they belong, all of them, to the superstructure; the solid foundation below is the system of inductive operations. The difficulty of a logical justification of these operations misled philosophers to seek a justification of the superstructure, to attempt an ontological justification of inductive belief by looking for necessary qualities of the world which would insure the success of inductive inferences. All such attempts will fail—because we shall never be able to give a cogent proof of any material presumption concerning nature. The way toward an understanding of the step from experience to prediction lies in the logical sphere; to find it we have to free ourselves from one deep-rooted prejudice: from the presupposition that the system of knowledge is to be a system of true propositions. If we cross out this assumption within the theory of knowledge, the difficulties dissolve, and with them dissolves the mystical mist lying above the research methods of science. We shall then interpret knowledge as a system of posits, or wagers; with this the question of justification assumes as its form the question whether scientific knowledge is our best wager. Logical analysis shows that this demonstration can be given, that the inductive procedure of science is distinguished from other methods of prediction as leading to the most favorable posits. Thus we wager on the predictions of science and wager on the predictions of practical wisdom: we wager on the sun's rising tomorrow, we wager that food will nourish us tomorrow, we wager that our feet will carry us tomorrow. Our stake is not low; all our personal existence, our life itself, is at stake. To confess ignorance in the face of the future is the tragic duty of all scientific philosophy; but, if we are excluded from knowing true predictions, we shall be glad that at least we know the road toward our best wagers.

INDEX

INDEX

- Abstracta, 93, 211, 235; existence of, 93, 101
Action: and meaning, 70, 80, 309, 344; and weight, 25, 32, 64, 315
Analysis of science, 8
Anticipation, method of, 353, 377
Aristotle, 299, 347
Atom, 215, 263, 267; existence of, 213
Avenarius, Richard, 163
Average error, 398
Bacon, 341, 347
Bases of epistemological construction, 203, 262
Basic statement, 173; in the narrower sense, 181; in the wider sense, 181
Basis: atom, 215, 263, 267; concreta, 263, 275; impression, 263; internal-process, 263; proposition, 263, 268; reaction, 263; stimulus, 263
Bayes, rule of, 124
Behaviorism, 163, 240
Bodily feeling, 235, 259
Bohr, 157
Boltzmann, L., 213
Boole, 299, 334, 342
Bühler, 60
Carnap, R., 5, 38, 60, 76, 145, 163, 171, 204, 226, 269, 335, 338
Causality, 317, 364, 370, 373, 392; homogeneity of, 139
Cavendish, 386
Class; *see* Probability
Complex, 98, 105; disjunctive, 107, 111; existence of the, 111; projective, 111, 130; reducible, 111, 130, 143
Composition, relation of, 99
Concatenation of inductive inferences, 363, 387, 391
Concreta, 93, 98, 210, 214, 265
Constancy of the velocity of light, 398
Constitutive relations, 107
Context: of discovery, 7, 382; of justification, 7, 382
Convention, 9
Conventionalism, 14, 271
Copeland, 298
Cross-induction, 366, 372
Darwin, 390
Decision, 146; differential, 384; entailed, 13, 63
Deducibility relation, 269, 336
Demarcation value, 327
De Morgan, 299
Descartes, 85, 261, 334, 344, 401
Description, 196
Dewey, 49, 163
Discovery, context of, 7, 382
Disparity conception, 300, 302, 325, 338
Dörge, 298
Dream, 92, 102, 139, 144, 165, 202, 205
Economy, principle of, 376
Ego, 152, 259
Einstein, 9, 43, 78, 127, 381, 386
Elements: complete set of, 107; external, 110; internal, 98, 105, 110
Equally possible cases, 300
Equally probable, 305
Euclidean geometry, 12, 271
Evidence, 285

Existence: of abstracta, 93, 101; grammar of the word, 195; of external things, 90, 102, 111, 129, 133; immediate, 199, 218; independent, 115, 132; objective, 199, 204, 218; reducibility of, 105, 114; return to the basis of immediate, 204, 275; subjective, 199, 204
 Existential coupling, 201
Experimentum crucis, 388

Fact: logical, 11, 336; object, 11; physical, 83; single, 84
 Falsification, 88
 Fermat, 298
 Fleck, L., 224
 Frequency interpretation, 300, 304, 329, 337
 Freud, 208, 246
 Galileo, 371, 385
 Gauss, 298
Gestalt, 100, 221

Hartmann, M., 391
 Helmholtz, 9
 Hempel, C. G., 37
 Hertz, P., 357
 Hilbert, 335
 Hume, David, 73, 78, 262, 335, 341, 348, 356, 362, 401

Identity conception, 300, 302, 325, 338
 Illata, 212, 227
 Implication, strict, 269
 Impression, 88, 171; and external things, 101, 115, 129, 132, 144; and form, 173; stereoscopic, 228
Inductio per enumerationem simplicem, 389

Induction, 339, 356, 382, 400; aim of, 350; belief in, 402; concatenation of, 363, 387, 391; cross-, 366, 372; formulation of the principle of, 340; justification of the principle of, 348; as necessary condition, 349

Inductive inferences, concatenation of, 387
 Inner process, 226
 Insufficient reason, principle of, 306
 Interactional quality, 168
 Intuitiveness, 178
 Introspection, 227, 234
 Isomorphism of the two probability concepts, 303
 James, W., 49
 Justification, context of, 7, 382
 Kant, 234, 334, 346, 401
 Kantianism, 12
 Kepler, 371, 385
 Keynes, J. M., 300, 302, 332
 Kokoszynska, Marja, 37
 Kolmogoroff, 298
 Language, 16, 57; analysis of, 270; of chess, 28; communicative function of, 59; egocentric, 135, 140, 147; emotional function of, 60; impression 149, 237; inner-process, 237; objective, 266; reaction, 232; reflexive function of, 60; second-level, 155; stimulus, 231; subjective, 266; suggestive function of, 59; syntax of, 336
 Laplace, 298, 301
 Lavoisier, 388
 Leibnitz, 78, 299, 334
 Lévy-Bruhl, 205
 Lewis, C. I., 151, 269
 Lichtenberg, 261
 Limit, practical, 361
 Localization: of abstracta, 98; of external objects, 223; of psychical phenomena, 234; within our body, 167
 Locke, 164
 Löwy, H., 261
 Logic: alternative, 321, 326; aprioristic conception of, 334; formalistic conception of, 334, 343, 359; probability, 319, 326, 336, 373;

of propositional series, 325; two-valued, 321, 326; of weights, 324
 Lukasiewicz, 300
 Mach, 78, 213, 376
 Meaning, 17, 20, 30, 59, 63, 80; and action, 70; functional conception of, 156; logical, 40, 55, 62, 124, 134; physical, 40, 55, 72; physical-truth, 55, 62, 127, 149; probability, 54, 62, 124, 127, 149, 153, 160; probability theory of, 54, 71, 87, 133, 189; super-empirical, 62, 68; truth, 55, 148; truth theory of, 30, 37, 53, 101, 191; verifiability theory of, 55, 57, 77, 79, 95, 189, 305
 Memory, reliability of, 180
 Michelson, A. A., 84, 398
 Mill, John Stuart, 342, 389
 Mises, v., 298
 Modality, 320
 Mysticism, religious, 58
 Necessary condition, 349, 351, 354, 356, 360
 Necessity, 320, 336
 Neurath, 163
 Newton, 78, 371, 373, 385
 Nominalism, 93
 Ogden, 60
 Ontology, 98, 336, 404
 Overreaching character, 127, 130, 132, 365
 Parallels, convergence of, 223
 Part and whole, 99
 Pascal, 298
 Passivity in observation, 258
 Peirce, C. S., 49, 299, 339
 Perceptual function, constancy of the, 183
 Plato, 334, 401
 Poincaré, 14
 Popper, K., 88, 302, 338
 Posit, 313, 352, 366; appraised, 352; blind, 353, 366
 Positivism, 30, 79, 156, 265; and existence, 101, 112, 129; and meaning, 30, 72, 189; as a problem of language, 145
 Possibility, 38, 320
 Pragmatism, 30, 48, 57, 69, 79, 150, 163
 Predictability, 350, 359
 Predictional value, 26, 190, 315
 Predictions, 339, 360, 381, 401; of a clairvoyant, 353, 358; included in every statement, 85, 131
 Presentation, 89
 Probability, 24, 75, 292, 297; a posteriori, 340; a priori, 124; axiomatic of, 337; backward, 124; class determining the degree of, 316, 399; concatenation of, 274; connection, 52, 104, 109, 192, 244; of disjunction, 173; forward, 124; frequency interpretation of, 300, 304, 329, 337; of higher level, 331, 395; implication 51; inference, 130, 142; initial, 124, 369; lattice, 366; logic, 319, 326, 336, 373; logical concept of, 299, 303; mathematical concept of, 298; of a scientific theory, 396; of the second level, 367, 368; series with changing probabilities, 310; series, order of, 317; of the single case, 302, 305, 309, 352
 Projection, 110, 129, 136, 143, 212; internal, 216
 Proposition, 20; basic, 173; co-ordination of, 95, 108; direct, 47; impression, 89, 169; indirect, 47; molecular, 21; observation, 34, 37, 87; predicates of, 19, 188; religious, 65; report, 276, 282
 Propositional series, 324, 329
 Psychical experience, incomparability of, 248
 Psychoanalysis, 208, 246, 287
 Psychology, 225

- Quale of the sensation, 250
 Quantum mechanics, 139, 187, 317
 Rational belief, 334, 338
 Rational reconstruction, 5
 Reaction, 226
 Realism, 93, 145, 159
 Recollection image, 179
 Reducibility relation, 99, 114
 Reducible series, 358
 Reduction, relation of, 98, 105, 114
 136, 148
 Relativity of motion, 44
 Representation, 209
 Retrogression, principle of, 49, 101,
 130, 148, 343
 Russell, Bertrand, 95, 335
 Schaxel, I., 250
 Schematization of the conception of
 knowledge, 104, 157, 188, 383, 393
 Schiller, 49
 Secondary quality, 167
 Self-observation, 234
 Semiconvergent series, 361
 Sensation, 89
 Sense, 20; inner tactile, 237; internal,
 164, 226; of touch, 166
 Sense data, 89
 Sentence, 21
 Significant, 158
 Similarity: disjunction, 172; immedi-
 ate, 171
 Simplest curve, 375, 379
 Simplicity, 373; descriptive, 374; in-
 ductive, 376
 Simultaneity, 43, 127, 137
 Single case; *see* Probability
 Specimens, collection of, 35, 182, 202
 Statement, 21; *see also* Proposition
 Stimulus, 226; *see also* Basis
 Subjectivism, 290
 Substitute world, 220
- Subvocal speaking, 343
 Sufficient conditions, 355
 Superiority of the immediate present,
 281
 Symbols, 17
 Tarski, A., 37
 Task: advisory, 13; critical, 7; de-
 scriptive, 3
 Tautology, 335
 Things: immediate, 199, 289; ob-
 jective, 199, 276, 289; peremptory
 character of immediate, 275, 290;
 subjective, 199, 276
 Tolman, E. C., 163
 Tornier, 298
 Transport time, 128
 Trichotomy, 327
 Truth, 31, 28, 190; objective, sub-
 jective, and immediate, 280, 287;
 physical theory of, 32, 33
 Truth-value, 21, 321
 Utilizability, 69, 150, 344
 Venn, 300, 334, 342
 Verifiability, 30, 38, 304; absolute, 83,
 125, 187
 Verification, indirect, 46
 Volitional bifurcation, 10, 147
 Volitional decision, 9
 Voltaire, 262
 Wager, 314
 Watson, 163, 243
 Weight, 24, 188, 297, 314, 394; and
 action, 32, 71; appraisal of, 319, 332,
 366, 399; initial, 277; and meaning
 75, 120
 Weights, system of, 273
 Whewell, 342
 William of Ockham, 77
 Wittgenstein, 49, 74, 335
 Zawirski, 300, 321

PREVIOUS PUBLICATIONS BY THE AUTHOR

- Relativitätstheorie und Erkenntnis apriori.* Berlin, 1920.
Axiomatik der relativistischen Raum-Zeit-Lehre. Braunschweig, 1924.
Philosophie der Raum-Zeit-Lehre. Berlin, 1928.
 "Ziele und Wege der physikalischen Erkenntnis," in *Handbuch der Physik*, ed. GEIGER-SCHEEL, Vol. IV. Berlin, 1929.
Atom und Kosmos: Das physikalische Weltbild der Gegenwart. Berlin,
 1930. English ed., London, 1932; American ed., New York, 1933;
 Spanish ed., Madrid, 1931; French ed., Paris, 1934; Hungarian ed.,
 Budapest, 1936.
Ziele und Wege der heutigen Naturphilosophie. Leipzig, 1931. French
 ed., under the title *La Philosophie scientifique*. Paris, 1932.
*Wahrscheinlichkeitslehre: Eine Untersuchung über die logischen und
 mathematischen Grundlagen der Wahrscheinlichkeitsrechnung.* Leiden,
 1935.