

§ 2. The Nature and Purpose of Physical Theory.

There have been lately a number of new interpretations of quantum mechanics, most of which are equivalent in the sense that they predict the same results for all physical experiments. Therefore no hope of deciding among them on the basis of physical experiments we must turn elsewhere, and enquire into the fundamental question of the nature and purpose of physical theory in general. Only after we have investigated and come to some sort of agreement upon these general questions, i.e. of the role of theories themselves, will we be able to put these alternative interpretations in their proper perspective.

Every theory can be divided into two separate parts, the formal part, and the interpretive part. The formal part consists of a purely logico-mathematical structure, i.e., a collection of symbols together with rules for their manipulation, while the interpretive part consists of a set of "associations", which are rules which put some of the elements of the formal part into correspondence with the perceived world. The essential point of a theory, then, is that it is a mathematical model, together with an isomorphism¹ between the model and the world of experience (i.e., the sense perceptions of the individual, or the "real world" -- depending upon one's choice of epistemology).

¹ By isomorphism we mean a mapping of some elements of the model into elements of the perceived world which has the property that the model is faithful, that is, if in the model a symbol A implies a symbol B, and A corresponds to the happening of an event in the perceived world, then the event corresponding to B must also obtain. The word homomorphism would be technically more correct, since there may not be a one-one correspondence between the model

~~Two types of prediction can be distinguished, the prediction of phenomena already indicated, in which the theory plays simply the role of a device to compare the observed results (the aspect of most interest to the engineer) and to compute the secondary and tertiary, unanticipated before the formulation of the theory effects, often transverse. This fall in which it was the contribution of the engineer.~~

and the external world.

The model nature is quite apparent in the newest theories, as in nuclear physics, and particularly in those fields outside of physics proper, such as the Theory of Games, various economic models, etc., where the degree of applicability of the models is still a matter of considerable doubt. However, when ~~this~~ theory is highly successful and becomes firmly established, the model tends to become identified with "reality" itself, and the model nature of the theory becomes obscured. The rise of classical physics offers an excellent example of this process. [Insert O]

Once we have granted that any physical theory is essentially only a model for the world of experience, we must renounce all hope of finding anything like "the correct theory". There is nothing which prevents any number of quite distinct models from being in correspondence with experience (i.e., all "correct"), and furthermore no way of ever verifying that any model is completely correct, simply because the totality of all experience is never accessible to us.

How, then, does one decide among a number of theories? Beyond the criteria of correctness (correspondence with experience) and logical consistency, we can only call upon such notions as usefulness, simplicity, pictorability, elegance, symmetry properties, etc., which may depend to a large degree upon personal taste. ^{Is there any justification for the adoption of} We must decide first why we want a physical theory at all.

Perhaps the fundamental purpose of a theory is prediction, → ^{the} ~~of the first type of prediction,~~ From this viewpoint we would say that the "best" theory is the one from which the most accurate

predictions can be most easily deduced, -- two not necessarily compatible ideals. Classical physics, for example, permits deductions with far greater ease than the more accurate theories of relativity and quantum mechanics, and in such a case we must retain ~~both~~^{them all}. It would be the worst sort of folly to advocate that the study of classical physics be completely dropped in favor of the newer theories. It can even happen that several quite distinct models can exist which are completely equivalent in their predictions, such that different ones are most applicable in different cases, a situation which seems to be realized in quantum mechanics today. It would seem foolish to attempt to reject all but one in such a situation, where it might be profitable to retain them all.

Nevertheless, we have a strong desire to construct a single all-embracing theory which would be applicable to the entire universe. From what stems this desire? The answer lies in the consideration of inductive inference^{the second type of prediction - the discovery of new phenomena, and involves} and the factors which influence our confidence in a given theory. If we have some data, without any theory connecting the data, it is usually possible to make some sort of crude^{statistical} predictions about the outcome of any further observations. If the data were highly random, however, we would have little confidence in such predictions. If, on the other hand, we had a theory, which had been tested after formulation on some of the data and found successful in all or nearly all cases, then we would ~~make~~^{make} predictions from this theory ~~and~~[↑] have a great deal of confidence in them, depending upon the number of previous successes of the theory.

^{be applicable}
^{the field of its formulation}
^{insert parenthetically}

This is a ~~little~~ difficult subject, and one which is only beginning to be studied seriously. Certain main points are clear, however, for example that is that our confidence increases with the number of successes of a theory. If a new theory replaces several older theories which deal with separate phenomena, i.e., a comprehensive theory of the previously diverse fields, then our confidence in the new theory is very much greater than the confidence in either of the older theories, since the range of success of the new theory is much greater than any of the older ones. It is therefore this factor of confidence which seems to be at the root of the desire for comprehensive theories.

A closely related criterion is simplicity -- by which we refer to conceptual simplicity rather than ease in use, which is of paramount interest to the engineer. A good example of the distinction is the theory of general relativity which is conceptually quite simple, while enormously cumbersome in actual calculations. Conceptual simplicity, like comprehensiveness, has the property of increasing confidence in a theory. A theory containing many ad hoc constants and restrictions^{, or many independent hypotheses,} in no way impresses us as much as one which is largely free of arbitrariness.

In summary, a physical theory is a logical construct (model), consisting of symbols and rules for their manipulation, some of whose elements are associated with elements of the perceived world. The fundamental requirements of a theory are logical consistency and correctness. There is no reason why there cannot be any number of different theories satisfying these requirements, and further criteria such as usefulness, simplicity, comprehensiveness, pictorability, etc., must be resorted to in such cases to further restrict the number. Even so, it may be

5

impossible to give a total ordering of the theories according to "goodness", since different ones may rate highest according to the different criteria, and it may be most advantagous to retain more than one.

As a final note we might comment upon the concept of causality. It should be clearly recognized that causality is a property of a model, and not a property of the world of experience. The concept of causality only makes sense with reference to a theory, in which there are logical dependences among the elements. A theory contains relations of the form "A implies B", which can be read as "A causes B", while our experience, uninterpreted by any theory, gives nothing of the sort, but only a correlation between the event corresponding to B and that corresponding to A.

Footnotes

Chap. II

- page 2 ¹ We regard it as undefined if $P(w_1, \dots, x_j) = 0$. In this case $P(w_1, \dots, x_j, y_k, \dots, z_l)$ is necessarily zero also.
- 3 ¹ This definition corresponds to the negative of the entropy of a probability distribution as defined by Shannon [].
- 9 ¹ A measure is a non-negative, countably additive set function, defined on some subsets of a given set. It is a probability measure if the measure of the entire set is unity. See Halmos [].
- 10 ¹ See Kelley [], pg. 65.
- 16 ¹ See Feller [], or Doob [].
- 2 ² A Markov process is a stochastic process whose future development depends only upon its present state, and not on its past history.
- 18 ¹ See Khinchin [], pg. 15.

Chap. III

- page 3 ¹ More rigorously, one considers only finite sums, then completes the resulting space to arrive at $\mathcal{H}_1 \otimes \mathcal{H}_2$.
- 6 ¹ In case $\sum_i (\phi_i \eta, \psi^S) \phi_i = 0$ (unnormalizable) then choose any function for the relative function. This ambiguity has no consequences of any importance to us. See in this connection the remarks on pg. (8,III).
- 7 ¹ Except if $\sum_i (\phi_i \eta, \psi^S) \phi_i = 0$. There is still, of course, no dependence upon the basis.
- 10 ¹ Also called a statistical operator (von Neumann []).