Why is the bibliography omitted?

II.1 THE THEORY OF OBSERVATION IN QUANTUM MECHANICS

FRITZ LONDON AND EDMOND BAUER

PREFACE OF PAUL LANGEVIN

NB the 'subject/object' language

Quantum physics has brought an essential advance to science, the finding that in every experiment or measurement there inescapably enters the duality between subject and object, the action and reaction of observer and system observed, the observer and the measuring apparatus being viewable as one entity.

The classical view, disregarding the necessarily limited character of our knowledge and the retroactive effect of the measurement on the system observed, always postulated the possibility of an infinitely precise knowledge of the simultaneous values of all the parameters used for the description of the system. Heisenberg, in giving concrete significance to his principle of indeterminism, has shown how the

Originally published as *La théorie de l'observation en mécanique quantique*, No. 775 of *Actualités scientifiques et industrielles: Exposés de physique générale*, publiés sous la direction de Paul Langevin, Hermann, Paris (1939). English translations—including a new paragraph by Professor Fritz London—done independently by A. Shimony, and by J. A. Wheeler and W. H. Zurek, and by J. McGrath and S. McLean McGrath; reconciled in 1982.

very existence of quanta excludes the possibility of knowing precisely at the same time all the quantities which might be the object of our measurement.

The form in which quantum mechanics is presented today provides an admirable translation of this new situation. The wave function it uses to describe the object no longer depends solely on the object, as was the case in the classical representation, but, above all, states what the observer knows and what, in consequence, are his possibilities for predictions about the evolution of the object. For a given object, this function, consequently, is modified in accordance with the information possessed by the observer. The introduction of the wave function at the very foundation of our representation common-sensibly recognizes what is almost ignored by classical physics, that our possiblities for prediction depend, above all, on our information. It also expresses quite exactly the fact that certain quantities, called noncommutable, cannot be known simultaneously with complete certainty. It characterizes the system by a certain number of observable quantities, different forms of "maximum knowledge" corresponding to different so-called "pure cases."

The present work, where the authors expand lectures given by one of them at the Sorbonne, demonstrates the precision and clarity with which the formalism of quantum theory expresses this representation by the wave function of the information acquired by the observer, and the manner in which each new measurement intervenes to modify this representation.

The act of observation is analyzed here in a particularly penetrating way. The essential character of the new physics emerges with complete clarity in the two stages of change of the wave function: by coupling of the system observed with the measuring device; and by the intervention of the observer, who becomes aware of the result of the measurement and thus determines the new wave function—following the observation—by using the new datum to reconstitute his information bank.

This treatise does a valuable service. It brings out the important finding of the new physics: how we express our knowledge of the external world.

Authors' Preface

The majority of introductions to quantum mechanics follow a rather dogmatic path from the moment that they reach the statistical interpretation of the theory. In general, they are content to show by more or less intuitive considerations how the actual measuring devices always introduce an element of indeterminism, as this interpretation demands. However, care is rarely taken to verify explicitly that the formalism of the theory, applied to that special process which constitutes the measurement, truly implies a transition of the system under study to a state of affairs less fully determined than before. A certain uneasiness arises. One does not see exactly with what right and up to what point one may, in spite of this loss of determinism, attribute to the system an appropriate state of its own. Physicists are

to some extent sleepwalkers, who try to avoid such issues and are accustomed to concentrate on concrete problems. But it is exactly these questions of principle which nevertheless interest nonphysicists and all who wish to understand what modern physics says about the analysis of the act of observation itself.

Although these problems have already been the subject of deep discussions (see especially von Neumann, 1932), there does not yet exist a treatment both concise and simple. This gap we have tried to fill.

Paris, June 1939

INTRODUCTION

It is well known that theoretical physics has been transformed since the beginning of the century into an *essentially statistical* doctrine and that the *discovery of quanta* made this revolution inevitable.

The principal aim of this study will be the *statistical interpretation* of the formalism of quantum theory. Although these questions of interpretation were systematized about ten years ago (Heisenberg, 1927; von Neumann, 1927; Dirac, 1927), one still often meets rather fuzzy ideas about what it means that probabilities appear in modern physics.

According to some, this statistical character shows that our knowledge of laws at the atomic level is still *incomplete*: that there remain to be found some *hidden parameters*, determining those processes which, *provisionally*, we are content to describe in a statistical language. To believe them, one might hope some day to recast the theory in a deterministic mold.

Others would have it understood that the *action of the observer* is involved. They sometimes consider that this would be an action that is *causal*, but *incompletely known*, because one never knows the exact state of the observer. From this circumstance would arise the statistical spread of measurements, the exact results of which might be predictable if one could take better account of the intervention of the observer.

It has also been said that the law of *causality* may be *correct* but *inapplicable* because there is never any way to reproduce identical conditions.

The discussion of these questions is not at all a matter of speculation. It is a definite problem. To treat it one ought to apply quantum theory to—and thereby extract the central features of—the very process of measurement. One can convince oneself that statistical distributions, such as are given by quantum mechanics and verified by experiment, have such a structure that they cannot be reproduced by hidden parameters. It is not, as often claimed, a question of philosophical interpretation; quantum mechanics ought to be *testably false*, if atomic processes in fact were deterministic and only incompletely known. It would be necessary to change the theory fundamentally and give up some battle-tested results, if one wished to

reconstitute it on a deterministic basis. Causality is *no longer applicable*, it is true; but the reason for this fact is not the impossibility, in the last analysis, of reproducing identically the conditions of an experiment. The heart of the matter is the difficulty of separating the object and the observer.

Modern physics often advances only by sacrificing some of our traditional philosophical convictions. The case of quantum mechanics is especially instructive. In all innocence one sought to construct a theory which would contain only relations between the "observable" quantities of Bohr's theory, such as the frequencies and the intensities of spectral lines. Heisenberg followed this route and in this way succeeded in obtaining a formalism which would resolve this problem. But, as often happens in theoretical physics, the formalism of the theory, once established, carried one further than one expected. It implied more relations than its founders had started with, relations between quantities altogether disconnected from the original spectroscopy, but themselves also observable (coordinates, momenta, etc.). One was led quite naturally in this way to try—after the initial shock—to interpret these relations which had been exposed automatically by the theory. In this way the discussion of this formalism taught us that the apparent philosophical point of departure of the theory, the idea of an observable world, totally independent of the observer, was a vacuous idea. Without intending to set up a theory of knowledge, although they were guided by a rather questionable philosophy, physicists were so to speak trapped in spite of themselves into discovering that the formalism of quantum mechanics already implies a well-defined theory of the relation between the object and the observer, a relation quite different from that implicit in naive realism, which had seemed, until then, one of the indispensable foundation stones of every natural science.

To discuss the process of measurement it is necessary to consider at least two systems, the observer and the object. It is therefore necessary to apply the quantum theory of the many-body system. This exists at present only in the nonrelativistic approximation. We are therefore forced to limit ourselves to this approximation, which still neglects all effects of the time delay in the propagation of forces.

It is not possible to give here a detailed introduction to quantum mechanics. We will limit ourselves (§1 and 2) to recalling briefly, and a bit dogmatically, the definitions and the laws that we will need. For a more detailed exposition of quantum theory see, for example, de Broglie, 1930; Bloch, 1930; Kemble, 1937; Dushman, 1938.

§1. Resumé of the Principles of Quantum Physics

In atomic physics the use of statistical concepts came far earlier than wave mechanics. The first step in this direction was probably made at the moment when one described spontaneous radioactive decay by the laws of probability. Of course, at the beginning we thought that this was a provisional approach, forced by our

ignorance of what is going on inside the nucleus. But when Bohr, obviously guided by an analogy with statistical concepts, constructed his model of the atom with its spontaneous quantum jumps, and, above all, when Einstein gave his famous demonstration of Planck's radiation law on the basis of the idea of spontaneous and stimulated *transition probabilities* (Einstein's A and B coefficients), one already had the strong feeling that these probabilities ought to be something basic and primordial. In a world of discontinuous phenomena, the appearance of a statistical form for the elementary laws would seem almost inescapable. The theory of Bohr, although it does not yet furnish a mathematical scheme that is complete and coherent, has already allowed us to state questions of principle to which quantum theory must address itself. In a physics that deals with magnitudes, whose domains of variation are not necessarily continuous, one wants to know:

(1) *What are the possible values* of a physical quantity;
(2) *With what probabilities* are they realized in a given system and under given circumstances.

Quantum mechanics furnishes us with a precise scheme which allows a quantitative treatment of questions of this kind. We can summarize it in the following way:

"The state" of a system—given in classical mechanics at each instant t by the $2f$ values of the variables $q_1(t), q_2(t) \ldots q_f(t), p_1(t), p_2(t) \ldots p_f(t)$—is represented in quantum mechanics by a *complex function* of the f variables $q_1, q_2 \ldots q_f$ and of t, the "Schrödinger wave function"

$$\psi(q_1, q_2 \ldots q_f; t),$$

which is normalized in such a way that*

$$\int \psi\psi^* \, dq = 1. \tag{1}$$

The evolution of the system in time is governed in classical mechanics by a "Hamiltonian function" $H(q, p)$, characteristic of the system in question. This function of the coordinates $q_1, q_2 \ldots q_f$ and of the momenta $p_1, p_2 \ldots p_f$, which is nothing other than the energy, permits one to write the Hamiltonian equations of motion. This is essentially† the same function $H(q, p)$ which in quantum mechanics also gives the law of evolution of the ψ that represents the state of the system: one forms the operator‡ $H(q, -i\hbar\partial/\partial q)$ by replacing p_k in the Hamiltonian

* In what follows the integrals $\int dq$ are always taken over the entire configuration space $q_1, q_2 \ldots q_f$, and ψ^* is the complex conjugate of ψ.

† Completed by a term referring to "spin."

‡ Here the symbol $h/2\pi$ employed by London and Bauer has been replaced for convenience, and in accord with modern practice, by the Dirac notation for the quantum of angular momentum, $\hbar$. —Eds.

by the differential operator $-i\hbar\partial/\partial q_k$ and q_k by the operation of multiplication by q_k. Here a certain ambiguity enters in the ordering of factors, because the operations q_k and $\partial/\partial q_k$ are not commutable. It is enough for us to know that there are prescriptions which ordinarily suffice to determine uniquely the order of the operators, but we will not go into this in detail here. The operator $H(q, -i\hbar\partial/\partial q)$, once given, allows us to write the equation of evolution:[§]

$$H(q, -i\hbar\partial/\partial q)\psi = i\hbar\psi\partial/\partial t. \tag{2}$$

This equation, discovered by Schrödinger, has the important property of leaving invariant the integral $\int \psi\psi^* dq$, which is necessary for our normalization $\int \psi\psi^* dq = 1$ to be possible. Thus ψ, once normalized, always retains its normalization.

$\psi(q, t_0)$ represents a "state" of the system at an instant t_0. Here we take this term in a sense completely analogous to that which it has in classical mechanics, where one says that the data $q_1(t_0) \cdots q_f(t_0)$, $p_1(t_0) \cdots p_f(t_0)$ "represent a state." The knowledge of the representative of the state at a given instant is *necessary* and *sufficient* for an unambiguous calculation, with the aid of the dynamic law, of the representative of the state at every subsequent moment. We cannot forgo a part of these data without losing the ability to calculate the future. Neither can we add to them supplementary data without introducing useless tautologies or contradictions of the data already in hand.

The *stationary states* of Bohr correspond to special solutions of the Schrödinger equation, solutions exactly periodic in time and of the form:

$$\psi = \exp(-iEt/\hbar)u(q). \tag{3}$$

In consequence of (1) and (2) the "amplitude" $u(q)$ obeys a time-independent equation,

$$Hu = Eu, \tag{4}$$

and the condition of normalization,

$$\int uu^* dq = 1. \tag{5}$$

The pair of equations (4) and (5) ordinarily does not have a solution for every value of E. They present a "proper value" or "eigenvalue" or "characteristic value" problem. Only for a "spectrum" of special values $E_1, E_2, E_3 \ldots$, possibly also

[§] In accordance with present-day practice we omit the square brackets in which the authors enclose the left-hand side of this and subsequent similar equations; and we use on the right-hand side the opposite sign of i from that which they use whenever they write the time-evolution operator. —Eds.

containing continuous intervals,* can condition (5) be fulfilled. For other values of E the linear equation (4) of course also possesses solutions u containing an arbitrary factor; but they are not square integrable, ruling out the normalization (5).

The allowed values,

$$E_1, E_2, \ldots E_k, \ldots$$

are called the "eigenvalues" of the operator H. The corresponding solutions u are called "eigenfunctions" and designated by corresponding indices,

$$u_1, u_2, \ldots u_k, \ldots .$$

It was the fundamental idea of Schrödinger to identify the spectrum of eigenvalues, $E_1, E_2, \ldots$, which have the dimensions of energy, with the allowed values of the energy in Bohr's theory, and the success of this ingenious idea is well known.

One knows also that the founders of wave mechanics were guided in the beginning by the conviction that it was necessary to *get rid of discontinuities*, or rather to derive them from an essentially *continuous* substructure, from a field theory, and thus overturn the basically statistical picture of Bohr and Einstein. But this program did not turn out to be realizable.

The statistical interpretation of quantum mechanics may be considered to be a particularly conservative attempt to maintain the picture worked out by Bohr and Einstein and to embody it in a coherent theoretical system.

Now that we know how to interpret the "monochromatic" solutions of the wave equation (2) such as

$$\psi = \exp(-iE_k t/\hbar)u_k(q)$$

as describing a state of energy E_k, we have to *discover the meaning of the most general solutions*. One can show that if one limits oneself to square integrable functions (1), *the most general solution* of equation (2) is written in the form:

$$\psi = \sum_k c_k \exp(-iE_k t/\hbar)u_k(q), \qquad (6)$$

where the c_k are complex constants.

Born, replying to this question, formulated the foundations of the statistical interpretation of quantum theory. He postulated that the quantity $|c_k|^2$ gives the

* In the case of a continuous spectrum condition, (5) has to be imposed on a function u associated with a small *interval of* E of the continuum ("proper differential"). In what follows we will not take account of these sophistications, which do not concern questions of principle, and we will write all formulas as if there were nothing but discrete spectra.

probability of finding the value E_k of the energy when the system in the state (6) is subjected to a measurement of energy.

If we introduce the abbreviation* $\psi_k = c_k \exp(-iE_k t/\hbar)$, we can write

$$\psi = \sum_k \psi_k(t) u_k(q). \tag{6'}$$

The coefficients ψ_k depend only on the time, and, as $|\psi_k|^2 = |c_k|^2$, we can also interpret the quantity $|\psi_k(t)|^2$ as the probability of finding the value E_k for the energy in the state ψ.

The coefficients ψ_k may, moreover, be calculated very easily with the help of two basic properties of eigenfunctions:

(1) "*Orthogonality*":

$$\int u_k^* u_l \, dq = \delta_{kl} \tag{7}$$

where

$$\delta_{kl} \equiv \begin{cases} 1 \text{ if } k = l \\ 0 \text{ if } k \neq l \end{cases};$$

(2) "*Completeness*": for an arbitrary square integrable function $f(q)$ (that is, $\int f^* f \, dq$ = finite) one has the identity:

$$\sum_k \left| \int f u_k \, dq \right|^2 = \int f^* f \, dq. \tag{8}$$

Orthogonality immediately gives the form of the coefficients ψ_k of the expansion (6'). Multiplying (6') by u_l^* and integrating, one obtains:

$$\int \psi(q) u_l^*(q) \, dq = \psi_l. \tag{9}$$

Completeness guarantees that the series thus defined converges to the function $\psi(q)$. Thus, the integral of the square of the absolute value of the difference,

$$\psi(q) - \sum \psi_l u_l(q),$$

goes to zero,

* Not a customary abbreviation and one possibly confusing to the reader, but well adapted to the purpose of the next few sections. —Eds.

$$\int \left|\psi(q)-\sum_l \psi_l u_l(q)\right|^2 dq = \int \psi\psi^* dq - \sum_l \psi_l \int \psi^* u_l\, dq - \sum_l \psi_l^* \int \psi u_l^*\, dq + \sum_l \psi_l \psi_l^*$$

$$= \int \psi\psi^* dq - \sum_l \left|\int \psi u_l^*\, dq\right|^2 = 0,$$

from which also follows the result

$$\sum_l \psi_l \psi_l^* = \int \psi\psi^* dq = 1.$$

In other words, the sum of the probabilities of finding the various values of the energy is unity. Thus if we ask what is the energy, we can be assured of always finding some value. Naturally, it is necessary that things be so if the definitions of Born are to make sense.

Thus we see how, for energy at least, quantum theory answers the two questions which show up (first paragraph of §1) in every theory of the discontinuous:

1. The *possible values of the energy* are the eigenvalues $E_1, E_2, E_3, \ldots$ of the operator $H(q, -i\hbar\partial/\partial q)$.

2. The *probability* of finding the value E_k of the energy in the state represented by ψ is given by

$$Prob_\psi(E_k) = \left|\int \psi u_k^*\, dq\right|^2, \tag{10}$$

where u_k is the eigenfunction associated with the value E_k of the energy.

In particular, if by chance this state is represented by an eigenfunction of the energy; that is, if

$$\psi = u_k,$$

then equation (10) gives the probabilities 1 for the eigenvalue E_k and 0 for all other eigenvalues of the energy.

In the original theory of Bohr one was occupied above all with the energy. However, in our present formalism energy does not play an exceptional part except for the time-evolution of the state as represented by its ψ function (equation 2). If ψ is given at a certain instant, we can also look for statistical predictions for an arbitrary physical quantity $F(q, p)$ at that moment.

The *generalization* of our definitions *for other quantities* $F(q, p)$ (as, for example, $xp_y - yp_x$, x, p_x, etc.) is fully specified as follows.

We form the operator $F(q, -i\hbar\partial/\partial q)$ and define, with the help of the equations of the eigenvalue problem,

$$Fv(q) = fv(q)$$
$$\int vv^* \, dq = 1, \tag{11}$$

the eigenvalues $f_1, f_2, \ldots f_\rho \ldots$ of the operator F and the corresponding eigenfunctions $v_1, v_2, \ldots v_\rho \ldots$. Mathematicians have shown that these eigenfunctions, too, under rather general conditions, form a *complete system* of orthogonal functions; that is, that we have the relations

$$\int v_\rho v_\sigma^* \, dq = \delta_{\rho\sigma} \tag{12}$$

and that we can expand the function ψ at a given moment t_0 in a convergent series of these functions v_ρ

$$\psi(q, t_0) = \sum_\rho \psi_\rho(t_0) v_\rho(q),$$

where

$$\psi_\rho(t_0) = \int \psi(q, t_0) v_\rho^*(q) \, dq.$$

The generalization of our previous definitions is immediate:

1. The *possible values of the physical quantity* F are given by the eigenvalues of the operator $F(q, -i\hbar\partial/\partial q)$.
2. The *probability* for finding the eigenvalue f_ρ of the quantity $F(q, p)$ for the state represented by ψ is given by

$$Prob_\psi(f_\rho) = \left| \int \psi(q, t_0) v_\rho^*(q) \, dq \right|^2 = |\psi_\rho|^2. \tag{13}$$

In particular, if by chance the state is represented by the eigenfunction v_ρ of F; that is, if

$$\psi = v_\rho,$$

one obtains the probability 1 for finding the value f_ρ of the quantity F and 0 for the probability for every other eigenvalue.

From these definitions immediately follows the *mean value of* F *in the state* ψ,

Expectation values

$$Mean_\psi(F) = \sum_\rho |\psi_\rho|^2 f_\rho. \tag{14}$$

This expression can be written in a more convenient form, which allows one to calculate immediately the mean value without having to go back to an explicit

evaluation of the individual eigenvalues f_ρ or the expansion of ψ as a series in the eigenfunctions v_ρ. One easily checks that

$$Mean_\psi(F) = \int \psi^* F\psi \, dq \tag{14'}$$

because, as consequence of (11) and (12) this expression lets itself be written as

$$\int \sum_{\rho\sigma} \psi_\rho^* v_\rho^* F \psi_\sigma v_\sigma \, dq = \sum_{\rho\sigma} \psi_\rho^* \psi_\sigma \int v_\rho^* f_\sigma v_\sigma \, dq = \sum_{\rho\sigma} \psi_\rho^* \psi_\sigma f_\sigma \delta_{\rho\sigma}.$$

§2. Vector Notation

Our definitions are now complete. We have only to introduce a slightly more convenient notation, that of the language of vectors.

Vectors. We will say that the function $\psi(q)$, representing the instantaneous state of a system, is a "vector" in a space of an infinite number of dimensions, the function space of Hilbert. The integral of the product of the two functions ψ^* and ϕ taken over all coordinates $q_1, q_2 \ldots q_f$ will be called the "*scalar product of* ψ *and* ϕ,"

$$(\psi, \phi) = \int \psi^*(q)\phi(q)\, dq = \sum_\rho \psi_\rho^* \phi_\rho. \tag{1}$$

The quantity $(\psi, \psi) = \int \psi^* \psi \, dq = \sum_\rho |\psi_\rho|^2$ will be called the *square of the length of the vector.*

$(\psi, \phi) = 0$ means that the vectors ψ and ϕ are "orthogonal."

The eigenfunctions $v_1, v_2 \ldots v_\rho \ldots$ of an operator F satisfy the relation (§1, equation 12),

$$(v_\rho, v_\sigma) = \begin{Bmatrix} 1 \text{ if } \rho = \sigma \\ 0 \text{ if } \rho \neq \sigma \end{Bmatrix}.$$

They form therefore a system of orthogonal unit vectors that define an "*orthogonal coordinate* system" with the help of which one can represent any vector ψ whatsoever in the form,

$$\psi(q) = \sum_\rho \psi_\rho v_\rho(q). \tag{1}*$$

The "components" ψ_ρ of the vector ψ are defined by the "projection"

* Both equations are numbered (1), presumably an oversight. —Eds.

$$\psi_\rho = (v_\rho, \psi)$$

of ψ in the direction of the unit vector v_ρ.

This decomposition into components is quite analogous to the resolution of a vector in ordinary space into projections. We can consider the set of ψ_ρ as equivalent to the function $\psi(q)$ itself; it is one particular decomposition of the vector ψ into orthogonal components. The coefficients

$$\psi_k = (u_k, \psi)$$

give an analogous decomposition of the same vector with respect to another system of orthogonal axes $u_1, u_2 \ldots u_k, \ldots$.

The representation of ψ by itself, that is, the function $\psi(q)$, may be regarded as a *special case* of representation in terms of orthogonal components—specifically, the orthogonal system composed of the eigenfunctions of the particular operator $F = q$. The eigenvalue problem for this operator has the form,

$$qv_\alpha(q) = q_\alpha v_\alpha(q),$$

or

$$(q - q_\alpha)v_\alpha(q) = 0.$$

The solutions are the "limiting" or symbolic functions of Dirac,

$$v_\alpha(q) = \delta(q - q_\alpha).$$

Such a function by definition vanishes for $q \neq q_\alpha$ but for $q = q_\alpha$ is so singular that

$$\int \delta(q - q_\alpha)\, dq = 1.$$

In terms of these special eigenfunctions one obtains for $\psi(q)$ the trivial expansion,

$$\psi(q) = \sum_\alpha \psi(q_\alpha)\delta(q - q_\alpha),$$

where the

$$\psi(q_\alpha) = \int \psi(q)\delta(q - q_\alpha)$$

are the coefficients in the expansion.

From our general definitions thus follows the particular result,

$$|\psi(q)|^2\, dq = \text{(probability to find } q \text{ in the interval from } q \text{ to } q + dq).$$

Tensors. In the present picture the typical operator $F(q, -i\hbar\partial/\partial q)$, representing a physical quantity, is a *tensor*; that is, a linear transformation of *vectors*. Applied to a vector ψ, it transforms it into another vector,

$$\psi' = F\psi.$$

It is *linear* because, for every combination of multiplications and differentiations, one always has the distributive relation,

$$F(\psi + \chi) = F\psi + F\chi, \tag{2}$$

and because, for every constant c, one has

$$Fc\psi = cF\psi.$$

As it represents a *real* (noncomplex) physical quantity, it has one more important property,

$$(\phi, F\psi) = (\psi, F\phi)^*. \tag{3}$$

Such an operator is termed "*Hermitian.*" Relation (3) is easily demonstrated by an integration by parts that takes account of the facts that every operation of differentiation contained in F brings in a factor i, and that F, considered as a function of q and $-i\hbar\partial/\partial q$, is a real function.

When we are using for ψ a representation in terms of the ψ_k—that is, when we refer the state vector to a system of coordinates $u_1, u_2 \ldots u_k \ldots$, we must also decompose the operator F into components referred to the same coordinates. In this way F evidently finds itself expressed as the linear transformation brought about by the matrix

$$F_{kl} = \int u_k^* F u_l \, dq = (u_k, Fu_l). \tag{4}$$

Thus if one applies this transformation to a vector ψ_l,

$$\sum_l F_{kl}\psi_l = (u_k, F\sum_l \psi_l u_l) = (u_k, F\psi),$$

one obtains the kth component of the function $F\psi$.

From (3) immediately follows the relation

$$F_{kl} = F_{lk}^* \tag{5}$$

of "*Hermitian matrices.*"

There is a system of coordinates in which the matrix representing the operator

F shows an especially simple form. These are the coordinates defined by the eigenfunctions $v_1 \cdots v_\rho \cdots$ of this operator itself. In terms of them one finds,

$$F_{\rho\sigma} = (v_\rho, Fv_\sigma) = f_\sigma(v_\rho, v_\sigma) = f_\sigma \delta_{\rho\sigma} = \begin{cases} f_\sigma \text{ for } \sigma = \rho \\ 0 \text{ for } \sigma \neq \rho \end{cases}.$$

In its "eigen"-basis the matrix F therefore takes a diagonal form and its diagonal elements are the eigenvalues of F,

$$F_{\rho\sigma} = \begin{Vmatrix} f_1 & 0 & 0 & 0 & \cdots \\ 0 & f_2 & 0 & 0 & \cdots \\ 0 & 0 & f_3 & 0 & \cdots \\ 0 & 0 & 0 & f_4 & \cdots \\ & \cdots\cdots\cdots & & & \\ & \cdots\cdots\cdots & & & \end{Vmatrix} \tag{6}$$

In the system of coordinates ψ_ρ used to expand ψ in a series of eigenfunctions $v_1 \cdots v_\rho \cdots$ of an arbitrary operator F, the Hamiltonian operator H takes the form $H_{\rho\sigma} = (v_\rho, Hv_\sigma)$ and the Schrödinger equation becomes

$$\sum_\rho H_{\rho\sigma}\psi_\sigma = i\hbar\dot{\psi}_\rho, \tag{7}$$

the discontinuous form in which Heisenberg, Born, and Jordan found the equations of quantum mechanics in the first place.

Invariants. We use capital letters H, F, etc. for tensors, and Greek letters ψ, ϕ, etc. for vectors. We can then suppress indices (or arguments) associated with any special decomposition into components and write (7) in an *invariant* form, independent of the system of coordinates, or rather encompassing all possible systems (see §1, equation 2),

$$H\psi = i\hbar\dot{\psi}.$$

Two distinct representations of the same vector ψ, for example $\psi_k = (u_k, \psi)$ and $\psi_\rho = (v_\rho, \psi)$, are related by a linear transformation,

$$\psi_k = \sum_\rho S_{k\rho}\psi_\rho \quad \text{with} \quad S_{k\rho} = (u_k, v_\rho). \tag{8}$$

Thus, as $\psi = \sum_\rho v_\rho(v_\rho, \psi)$, one has,

$$\psi_k = (u_k, \psi) = \sum_\rho (u_k, v_\rho)(v_\rho, \psi) = \sum_\rho S_{k\rho}\psi_\rho.$$

For the coefficients $S_{k\rho}$ one easily finds the relations,

$$\sum_{\rho} S_{k\rho} S^*_{l\rho} = \delta_{kl} \quad \text{or} \quad S\tilde{S}^* = 1, \tag{9}$$

where $\tilde{S}_{k\rho} = S_{\rho k}$.

These relations characterize the transformation (8) as a "*unitary*" *transformation.*

Two different representations of the same tensor F, for example $F_{kl} = (u_k, Fu_l)$ and $F_{\rho\sigma} = (u_\rho, Fu_\sigma)$, are related, as one easily verifies, by the relations,

$$F_{kl} = \sum_{\rho\sigma} S_{k\rho} F_{\rho\sigma} \tilde{S}^*_{\sigma l}.$$

Thus, as $u_k = \sum_{\rho} v_\rho (v_\rho u_k) = \sum_{\rho} S^*_{k\rho} v_\rho$, one has

$$F_{kl} = (u_k, Fu_l) = \int \sum_{\rho} S_{k\rho} v^*_\rho F \sum_{\sigma} S^*_{l\sigma} v_\sigma \, dq = \sum_{\rho\sigma} S_{k\rho} \int v^*_\rho F v_\sigma \, dq \, S^*_{l\sigma}$$

$$= \sum_{\rho\sigma} S_{k\rho} F_{\rho\sigma} \tilde{S}^*_{\sigma l}.$$

Physically significant numerical values naturally ought to be scalars *invariant* under these unitary transformations. The only scalar invariants that we will meet are the "*scalar product*" of two vectors ψ and ϕ,

$$(\psi, \phi) = \int \psi^*(q)\phi(q)\, dq = \sum_{k} \psi^*_k \phi_k = \sum_{\rho} \psi^*_\rho \phi_\rho,$$

and the "trace" of a tensor F_{kl},

$$Trace(F) = \sum_{k} F_{kk} = \sum_{\rho} F_{\rho\rho}.$$

Thus, for example, mean value of a physical quantity F in a state ψ is given in invariant form by the scalar product

$$Mean_\psi(F) = \sum_{k} F_{kk} = \sum_{\rho} F_{\rho\rho}.$$

The other results of the theory can also be expressed in invariant form. We will come back to this in §5.

The scalar product can also be considered as the trace of a special matrix $(\psi \times \phi)$ defined by

$$(\psi \times \phi)_{k\rho} = \psi^*_k \phi_\rho, \tag{10}$$

termed the "direct product" of the vectors ψ and ϕ.

§3. Statistics and Objectivity

Already in the classic memoir (Born, 1926) where he proposed foundations for the statistical interpretation of quantum mechanics, Born remarked that the probabilities which he introduced there must have a strange character, quite different from what one normally understands when one speaks of probability.

This feature he expressed in a form a bit paradoxical: "Although the movements of particles are not determined, except by probabilities, these probabilities themselves evolve according to a causal law." What he understands here by "causal law" is a connection between the "states" at different moments, such that a knowledge of the initial state at an arbitrary instant uniquely implies a knowledge of the state at every subsequent time. A "state," on the other hand, is a well-defined collection of data on the system in question at a given moment.

Naturally, there is no way of predicting *a priori* whether, in a given domain of science, there exist causal laws as so defined, nor what are the necessary and sufficient conditions for giving rise to such laws. If one does not end up with unique predictions, if one finds oneself forced to be satisfied with *probabilities*, that may be either because our knowledge of the "state" is not yet complete or because causality does not hold. But conversely, when one has succeeded in establishing causal laws, that is evidently a criterion for deciding that one has attained a complete knowledge of the object in question and thus, in some measure, a maximum description.

But the Schrödinger equation has all the features of a causal connection. If the ψ function is known at a given moment, it is determined at every subsequent time. It therefore seems difficult to believe that this function nevertheless contains a statistical collection. At first sight it seems impossible to avoid the following dilemma.

1. One might imagine that the ψ function has the character of the ordinary probability function such as one uses, for example, to describe Brownian motion. A function of this type contains certain statistical predictions that we can test. We then verify which of the possibilities foreseen in the theory is realized in fact in a given case. After this observation we are naturally entitled to use, for the subsequent predictions, the knowledge thus obtained and to replace our original probability function by a function of the same type, but better tuned. Evidently this is only possible by virtue of the enrichment of our knowledge, which is always partial. Of course we do not claim that the object itself has changed its state as a consequence of our observation. All that has changed is the discrepancy between our knowledge and the object. In this case the ψ function will therefore represent the state of our *partial knowledge* of the object and not the state of the object itself.

2. Imagine, on the contrary, that the ψ function has an "objective" character, as, for example the wave functions of optics. It then claims to represent, in an idealized and simplified form, something *complete*, a *maximum* picture of the state

of the *object*. But if this is the situation, it seems difficult to understand how this ψ function implies a statistic. If one checks experimentally predictions that can be made from it, and if one observes which of the possible outcomes is realized—an outcome predicted by the theory, but only with a certain probability—*by what right can we add this new knowledge* to the supposedly complete knowledge that we already had?

Heisenberg found the solution to this dilemma. He emphasized that it is the *process of measurement* itself which introduces the element of indeterminacy in the state of the object.

Thus the statistical feature would not show up except on the occasion of a measurement. If the ψ function gives us probabilities, it does so only *in anticipation of an eventual measurement*. Thus these are only, so to speak, "*potential*" *probabilities* which come into force only on the occasion of an actual measurement. They do not affect the precision with which the state of the system is currently known; thus it is already maximal when the ψ function is given.

Of course it may happen that there is an additional uncertainty in the state of the system—that is, in the ψ function itself. In this case it is a question of *probabilities* in the *ordinary* sense of the word. They arise from an incomplete knowledge of the state of the object. It is necessary to distinguish clearly between these probabilities and the "potential" probabilities furnished by ψ functions.

§4. Mixtures and Pure States

That an essential distinction is in question can be seen most clearly by an example.

I. Let us consider first the case where the system is represented by a wave function,

$$\psi = \sum_k \psi_k u_k(q),$$

where $u_1, u_2 \cdots u_k \cdots$ are, for example, eigenfunctions of the energy. We know that the quantity $|\psi_k|^2$ gives us the probability of finding the value E_k of the energy when an energy measurement is made on a system in a state ψ.

II. It has often been considered that this case corresponds straightforwardly to a *virtual ensemble* of identical systems in different states possessing respectively energies $E_1, E_2 \ldots E_k \ldots$, each one of them being contained in an ensemble with a relative weight $p_k = |\psi_k|^2$.

However, this latter case (II), which will interest us also, is basically different from the *pure state* (I) represented by a single function ψ. It is a *mixture* of several distinct pure states, each represented by *its wave function*,*

* In what follows the upper indices $(\psi^{(1)}, \psi^{(2)}, \ldots \psi^{(n)})$ always designate distinct pure states.

$$\psi(q) = u_1(q) \quad (1)$$

$$\psi(q) = u_2(q) \quad (2)$$

.....

$$\psi(q) = u_k(q) \quad (k)$$

.....

with the respective fractional abundances

$$p_1 = |\psi_1|^2, \qquad p_2 = |\psi_2|^2, \ldots p_k = |\psi_k|^2, \ldots$$

One easily verifies that case I and case II give two completely different statistical distributions. Of course they are identical for energy, because we have arranged that their statistics should be the same. But let us consider another physical quantity, F for example, not having the same eigenfunctions as energy. In case I the mean value of F takes the form (see §1),

$$Mean_{\mathrm{I}}(F) = (\psi, F\psi) = \sum_{k,l} \psi_k^* \psi_l F_{kl}, \tag{I}$$

while in case II each component $\overset{(k)}{\psi} = u_k$ gives a contribution of this kind:

$$Mean\,(F) = (\overset{(k)}{\psi}, F\overset{(k)}{\psi}) = (u_k, Fu_k) = F_{kk}.$$

The differing statistics of pure cases and of mixtures

When this component appears with the probability p_k in the mixture, one obtains altogether

$$Mean(F) = \sum_k p_k(\overset{(k)}{\psi}, F\overset{(k)}{\psi}) = \sum p_k F_{kk}. \tag{II}$$

If in particular we consider the case of a mixture where $p_k = |\psi_k|^2$, we find,

$$Mean_{\mathrm{II}}(F) = \sum_k |\psi_k|^2 F_{kk}. \tag{II'}$$

Thus, provided that F_{kl} is not by a chance a diagonal matrix (that is, F does not have the same eigenfunctions as H), the two cases I and II are completely different.

But only with same eigenvalues

It is evident that in case II our knowledge of the system is much more restricted than in case I. If our knowledge is limited to the statistics of energy—that is, if we have only the equations

$$p_1 = |\psi_1|^2, \qquad p_2 = |\psi_2|^2 \cdots p_k = |\psi_k|^2 —$$

we do not know the coefficients ψ_k themselves, but only their absolute values. Given that the ψ_k are normally complex quantities, we can write them in the form

$$\psi_k = \sqrt{p_k}\exp(i\alpha_k), \tag{1}$$

where the phases α_k are still indeterminate. One easily verifies that the difference between cases II and I arises from this ignorance about the phases α_k. Thus, in introducing (1) in the expression $Mean_I(F)$, we find

$$Mean_I(F) = \sum_{kl} \sqrt{p_k p_l} F_{kl} \exp[i(\alpha_l - \alpha_k)].$$

If we then average over the unknown phases we find that all of the terms with $k \neq l$ drop out and we get precisely $Mean_{II}(F)$.

One thus sees that it is necessary to make a careful distinction between:

I. A *pure state* described by a single wave function ψ that represents, we see, something irreducible, the probabilities that it implies being only "potential" probabilities;

II. A *mixture*, composed of different pure states

$$\overset{(1)}{\psi}, \overset{(2)}{\psi} \cdots \overset{(n)}{\psi},$$

realized with probabilities $p_1, p_2 \cdots p_n \ldots$. These latter probabilities are understood in the ordinary sense of the word. Naturally, they are all non-negative. We suppose them to be normalized:

$$\sum_n p_n = 1.$$

§5. The Statistical Operator

It will be useful to introduce here a concise notation to describe statistical ensembles in all generality. We will consider a mixture such as we have just defined (§4, case II). The mean value of a physical quantity G in this mixture is

$$Mean(G) = \sum_n p_n(\overset{(n)}{\psi}, G\overset{(n)}{\psi}),$$

or, referred to a concrete coordinate system,

$$Mean(G) = \sum_n p_n \sum_{\rho\sigma} \overset{(n)(n)}{\psi_\sigma \psi_\rho^* G_{\rho\sigma}},$$

which can be written

$$\begin{aligned} Mean(G) &= Trace(PG) \\ &= Trace(GP), \end{aligned} \tag{I}$$

when one introduces a Hermitian matrix P, the *statistical matrix*, defined by

$$P = \sum_n p_n \overset{(n)}{(\psi} \times \overset{(n)}{\psi)}; \tag{1}$$

that is, in some chosen system of coordinates,

$$P_{\sigma\rho} = \sum_n p_n \overset{(n)}{\psi_\rho^*} \overset{(n)}{\psi_\sigma}. \tag{1'}$$

The case of a pure state $\psi = \sum_\rho \psi_\rho u_\rho$ is included in these formulas as the special case of a mixture where all the p_n are zero except for the single one, which is equal to unity. Its statistical matrix takes the form

$$\overset{(\psi)}{P_{\sigma\rho}} = \psi_\rho^* \psi_\sigma, \quad \text{or} \quad \overset{(\psi)}{P} = (\psi \times \psi). \tag{2}$$

Let us call the statistical matrix for a pure state an *elementary matrix* (Einzelmatrix). The matrix P for a general mixture can thus be considered as a linear superposition of "elementary matrices,"

$$P = \sum_n p_n \overset{(n)}{P}.$$

The elementary matrix for a pure state ψ takes an especially simple form in a system of coordinates $v_1, v_2 \cdots v_\rho$ in which the wave function ψ is identical with one of the axes. For example, let ψ be equal to v_2. Then one has

$$\overset{(\psi)}{P}_{\rho\sigma} = \delta_{2\rho} \cdot \delta_{2\sigma} = \begin{Vmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{Vmatrix} \tag{2'}$$

Let us now calculate the probability of finding the value g_α of G in the ensemble characterized by the statistical matrix P.

Let us take as axes the eigenfunctions v_α of the operator G. For the pure component

$$\overset{(n)}{\psi} = \sum \overset{(n)}{\psi}_\alpha v_\alpha$$

with the index n we have (§1, equation 13)

$$Prob\,(g_\alpha) = |\overset{(n)}{\psi}_\alpha|^2 = \overset{(n)}{P}_{\alpha\alpha}.$$

When this component occurs in the ensemble with the probability p_n one has altogether

$$Prob\,(g_\alpha) = \sum_n p_n |\overset{(n)}{\psi}_\alpha|^2 = P_{\alpha\alpha}.$$

If we make use of the elementary matrix $\overset{(\alpha)}{P}$ for the pure state in which $G = g_\alpha$ (see 2'),

$$\overset{(\alpha)}{P}_{\beta\gamma} = \delta_{\alpha\beta}\delta_{\alpha\gamma},$$

we can express *the probability of finding the value g_α in the mixture P* in the invariant form,

$$Prob_P\,(g_\alpha) = P_{\alpha\alpha} = Trace\,(P\,\overset{(\alpha)}{P}). \tag{II}$$

In particular, if $P = \overset{(n)}{P}$ designates the case of the pure state where an arbitrary physical quantity F has the eigenvalue f_n, one sees that

$$Trace(\overset{(n)}{P}\ \overset{(\alpha)}{P}) = Prob(f_n, g_\alpha) \tag{III}$$

is the *probability of finding the value* $G = g_\alpha$ *for the pure state in which* $F = f_n$. One notes that this expression is completely symmetric between F and G. The same expression (III) also gives the *probability of finding* $F = f_n$ *for the pure state in which* $G = g_\alpha$.

The statistical matrices P therefore present an evident advantage. They permit us to express all our definitions so that the very form (I, II, III) already indicates the invariant character.

§6. Some Mathematical Properties of Statistical Matrices

(a) Let us calculate the *trace* of an arbitrary statistical matrix P,

$$Trace\, P = \sum_\alpha P_{\alpha\alpha} = \sum_n p_n \sum_\alpha |\overset{(n)}{\psi_\alpha}|^2 = \sum_n p_n = 1.$$

In this way we obtain the relation

$$Trace\, P = 1, \tag{1}$$

which expresses in brief form the normalization of probabilities.

(b) In particular, the *elementary* statistical matrices P for *pure states possess* the additional property

idempotency

$$\overset{(\psi)}{P^2} = \overset{(\psi)}{P}, \tag{2}$$

which results immediately from the definition (§5, equation 2) of these matrices,

$$(\overset{(\psi)}{P^2})_{\sigma\rho} = \sum_\tau \psi^*_\rho \psi_\tau \psi^*_\tau \psi_\sigma = \psi^*_\rho \psi_\sigma = \overset{(\psi)}{P}_{\sigma\rho}.$$

Relation (2) is in any case evident if one recalls the diagonal representation (§5, equation 2′) of elementary matrices.

projection operators

One perceives immediately that the converse is also true. From $P^2 = P$ and from $Trace\ P = 1$ it follows that P is an elementary matrix. Thus when P is written in its diagonal form, P^2 is likewise diagonal, and $P = P^2$ implies $p_i = p_i^2$. The eigenvalues are therefore zero or unity. From the equation $Trace\ P = \sum_i p_i = 1$

it follows, finally, that a single one of p_i is equal to unity, and that all others vanish. The relation $P^2 - P = 0$ is therefore *necessary and sufficient* for the statistical matrix P to be the matrix for a pure state.

(c) We have not yet placed any restrictions on the choice of the pure states which constitute the mixture. In particular, we have not assumed that they were represented by orthogonal wave functions ψ (later to be distinguished from one another by an index (n) written directly above the ψ). But one can show very easily that an arbitrary mixture composed of arbitrary pure states can always be written in the form of a mixture of *orthogonal* pure states with *non-negative* relative probabilities p_i.

Let us first verify that a matrix P is a *semi-definite* matrix; that is, that for our arbitrary vector ξ one always has

$$(\xi, P\xi) \geq 0. \tag{3}$$

Thus the definition (§5, equation 1') of P gives

$$(\xi, P\xi) = \sum_{n\alpha\beta} p_n \xi^*_\alpha \overset{(n)}{\psi}_\alpha \cdot \xi_\beta \overset{(n)}{\psi^*_\beta} = \sum_n p_n |(\xi, \overset{(n)}{\psi})|^2,$$

an expression which cannot be negative because

$$p_n \geq 0 \quad \text{and} \quad |(\xi, \overset{(n)}{\psi})|^2 \geq 0.$$

But P is a Hermitian matrix. Therefore, there exists an *orthogonal* coordinate system $v_1, v_2 \ldots v_\rho$ in which P takes its diagonal form,

$$P_{\rho\sigma} = \begin{Vmatrix} p'_1 & 0 & 0 & 0 & \cdots \\ 0 & p'_2 & 0 & 0 & \cdots \\ 0 & 0 & p'_3 & 0 & \cdots \\ 0 & 0 & 0 & p'_4 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{Vmatrix}. \tag{4}$$

The values of the diagonal elements $p'_1, p'_2 \ldots p'_\rho$ cannot be negative. Thus, again one has in these coordinates

$$(\xi, P\xi) = \sum_\rho p'_\rho |\xi_\rho|^2 \geq 0,$$

which is not possible for an arbitrary vector ξ unless $p'_\rho \geq 0$ for every ρ. The matrix P therefore can always be written in the form

$$P = \sum_\rho p'_\rho \overset{(\rho)}{P}, \tag{4'}$$

where the $\overset{(\rho)}{P}$ are the elementary matrices for a certain system of *orthogonal* pure states $v_1, v_2 \ldots v_\rho$. The $p'_1, p'_2 \ldots p'_\rho \geq 0$ are the relative probabilities for these states.

As *Trace* $P = 1$, we also have

$$\sum p'_\rho = 1,$$

and $p'_\rho \geq 0$ implies

$$0 \leq p'_\rho \leq 1.$$

(d) From this inequality follows another,

$$p'_\rho - p'^2_\rho \geq 0.$$

From this equation we conclude that for an arbitrary vector ξ,

$$(\xi, (P - P^2)\xi) = \sum_\rho |\xi_\rho|^2 (p'_\rho - p'^2_\rho) \geq 0. \tag{5}$$

The matrix $P - P^2$ is therefore likewise semidefinite.

In particular for a pure state the quantity $P - P^2$ vanishes (see equation 2).

Elementary statistical operators like $P = (\psi \times \psi)$ can be considered as "projectors" or "projection operators." Applied to an arbitrary state vector ξ, P singles out the projection of this vector in the direction of the vector ψ,

$$\sum_\sigma P_{\sigma\rho}\xi_\sigma = \sum_\sigma \psi_\rho \psi^*_\sigma \xi_\sigma = \psi_\rho(\psi, \xi).$$

The magnitude of this vector is (ψ, ξ) and its direction is that of the unit vector ψ. The iteration of the projection produces no further change: $P^2 = P$.

§7. The Statistical Operator and Thermodynamics

The operator P describes an ensemble of identical systems which are distributed in an arbitrary way among the different states. It plays the role analogous to that

of a distribution function in ordinary statistical mechanics. Therefore there ought to exist a connection between operator P and the macroscopic thermodynamic quantities. It will be enough for us to point out briefly the main features of this connection without stopping for proofs (see von Neumann, 1927, p. 273). Everything is summarized in the definition of entropy, S:

$$S = -kN \, Trace(P \ln P), \tag{1}$$

where

k = *Boltzmann's constant*

and

N = *the total number of systems.*

This relation becomes quite plausible when one recalls that in statistical mechanics one has

$$S = -k \sum_{\alpha} (n_\alpha \ln n_\alpha - N \ln N),$$

where n_α is the number of systems in the state α. Therefore if $p_\alpha = n_\alpha/N$ represents the probability for a system to be in the state α, one has again

$$S = -kN\left[\sum_{\alpha} p_\alpha(\ln p_\alpha + \ln N) - \ln N\right],$$

which is identical with (1) in a system of coordinates in which P is diagonal (§6, equation 4). Our definition of entropy is therefore the entirely straightforward generalization of the usual definition. One sees immediately that for a pure state the entropy thus defined is zero. Thus, if one represents P in diagonal form, the entropy becomes

$$S = -kN \sum_{\alpha} p_\alpha \ln p_\alpha.$$

Every term $p_\alpha \ln p_\alpha$ of the sum vanishes, because in a pure state the p_α are all zero except for a single one which is unity. One also sees immediately that the entropy of a mixture is always positive.

To maximize the entropy for a given total energy E is to impose on P the following conditions:

$$-kN \, Trace(P \ln P) \to Max;$$

$$Trace \, P = 1;$$

$$N \, Trace(PH) = E.$$

The solution of this extremum problem is represented by the matrix,

$$P = e^{-\beta H}/Trace(e^{-\beta H}) = e^{-\beta H}/Z(\beta), \tag{2}$$

where

$$Z(\beta) = Trace(e^{-\beta H}).$$

The Lagrange factor β is determined, as always, by equilibrium with a perfect gas; thus, $\beta = 1/kT$, where T is the absolute temperature.

One thus obtains for the entropy of the most probable distribution,

$$\begin{aligned} S &= kN\ Trace[e^{-\beta H}(\beta H + \ln Z)/Z(\beta)] \\ &= -kN[\beta\partial \ln Z/\partial\beta - \ln Z], \end{aligned}$$

and for the energy,

$$E = -N\partial \ln Z/\partial\beta,$$

from which one easily gets all the other thermodynamic quantities, for example the free energy

$$F = E - TS = -kNT \ln Z,$$

et cetera.

§8. The Irreducibility of the Pure Case

Our definitions would be worthless if pure cases were not characterized by some kind of irreducibility. We have to show that *it is not possible to represent a pure case in the form of a mixture.* For this purpose we will show that a statistical matrix P, obtained as a mixture of two statistical matrices Q and R,

$$P = \alpha Q + \beta R, \tag{1}$$

with

$$\alpha + \beta = 1, \qquad \alpha \geq 0, \qquad \beta \geq 0,$$

cannot be an elementary statistical matrix (such that $P = P^2$) except if $Q = R = P$. Let us calculate

$$\begin{aligned} P^2 &= \alpha^2 Q^2 + \beta^2 R^2 + \alpha\beta(QR + RQ) \\ &= \alpha^2 Q^2 + \beta^2 R^2 + \alpha\beta[Q^2 + R^2 - (Q - R)^2] \\ &= \alpha Q^2 + \beta R^2 - \alpha\beta(Q - R)^2, \end{aligned}$$

where we used the condition $\alpha + \beta = 1$. Thus

$$P - P^2 = \alpha(Q - Q^2) + \beta(R - R^2) + \alpha\beta(Q - R)^2.$$

We now recall that the matrices $Q - Q^2$ and $R - R^2$, as well as $(Q - R)^2$, are always semidefinite (§6, equation 5). Therefore it is necessary that all these quantities should vanish for P to be an elementary matrix ($P = P^2$). In particular, we have $(Q - R)^2 = 0$, from which we conclude

$$Q = R,$$

because the square of a Hermitian matrix cannot vanish unless the original matrix vanishes,

$$\left(\sum_k A_{ik}A_{ki} = \sum_k A_{ik}A_{ik}^* = \sum_k |A_{ik}|^2 = 0 \text{ implies } A_{ik} = 0\right).$$

From $Q = R$ and from (1) it follows that $Q = R = P$.

Statistical operators thus form an ensemble of a characteristic structure, called *convex ensemble*. Its boundary is formed by the operators for pure cases. No pure case can be constructed by linear superposition with positive coefficients—that is, by mixture—of two nonidentical pure cases.

Although the statistical operator P of a pure case cannot be decomposed, one might imagine that there perhaps exists some other means to reduce directly the corresponding statistics.

In actuality this problem does not differ from the problem that we have just discussed. However, it would perhaps be useful to consider it explicitly.

Of course we always take as our foundation the statistical distributions as predicted by the formalism of the theory and verified by so many experiments. Therefore we will not discuss the validity of these statistical predictions. We will ask ourselves rather if, once assumed, they might not be represented by mixtures of arbitrary form, but belonging to systems that are well defined in the ordinary sense of classical mechanics.

Let us take a concrete example that is as simple as possible, that of the statistical description associated with a "spin." Let us consider an atom of angular momentum $\hbar/2$ and let us focus attention on the component of this angular momentum in an arbitrary direction. For any given direction there are only two values possible for this component of the spin: $+\hbar/2$ and $-\hbar/2$. Let us fix in space the axis of a system of spherical polar coordinates. Let u_+ and u_- be the eigenfunctions associated with the two possible values for the component along this axis. To be concrete, suppose that

$$\psi = u_+$$

is the wave function of the state in question—that this state is a "pure case."

Let us now consider, for the same state ψ, the component of the spin in another direction, oblique to the polar axis and characterized, for example, by the polar coordinates θ, ϕ. The only possible values for the component of the spin in this direction are again $+\hbar/2$ and $-\hbar/2$. To evaluate the probabilities for finding each of these two values it is necessary to represent ψ in terms of the two eigenfunctions associated with those two possibilities. We will call them u'_+ and u'_-. A calculation gives

$$u_+ = c_+u'_+ + c_-u'_-,$$

with

$$c_+ = e^{i\phi/2} \cos\theta/2,$$
$$c_- = e^{-i\phi/2} \sin\theta/2,$$

The squares of the absolute values of the coefficients u'_+ and u'_- represent the probabilities for finding the one or the other of the two possible components in the direction θ, ϕ; thus,

$$p'_+ = |c_+|^2 = \cos^2\theta/2,$$
$$p'_- = |c_-|^2 = \sin^2\theta/2.$$

But, in the same state, the probabilities of the two values for the component along the original polar axis are, respectively,

$$p_+ = 1$$

and

$$p_- = 0.$$

Evidently it is impossible to decompose these statistics into a mixture of definitely oriented spins. In such a mixture it would be necessary that a fraction $\cos^2\theta/2$ of the atoms should have a component $\hbar/2$ in the direction θ, ϕ and that a fraction $\sin^2\theta/2$ should have the opposite orientation. This, by itself, would be possible. But we ought, in addition, to have 100% of the atoms with the same component $\hbar/2$ along the direction of the polar axis, and yet a fraction $\cos^2\theta'/2$ along any other direction θ'. That would be a juggling trick rather difficult to bring off!

Evidently it is impossible to arrange a virtual ensemble of oriented atoms that meets simultaneously all of these statistical requirements. The mathematics of the probability calculations already precludes this possibility.

We will use the same simple example in section §12 to study in detail how, by

his own intervention, the observer succeeds in doing the juggling trick. However, to keep the discussion clear, it is first necessary for us to bring out an aspect of quantum mechanics which we have not yet mentioned, but which contains the very essence of the theory, the feature responsible for the appearance of probabilities.

§9. Statistics of a System Composed of Two Subsystems

At first sight the mathematical formalism of quantum mechanics seems entirely analogous to that of the theories of "classical" physics: a differential equation uniquely prescribes the evolution of the wave function ψ that describes the state of the system. Therefore it seems as though our task, faced with Schrödinger's equation, is no different from that of Laplace faced with the equations of Newton. The state of a closed system, perhaps the entire universe, is completely determined for all time if it is known at a given instant. According to the Schrödinger equation, a pure case represented by a ψ function remains always a pure case. One does not immediately see any occasion for the introduction of probabilities, and our statistical definitions might appear in the theory as a foreign structure.

We will see that that is not the case. It is true that the state of a closed system, once given pure, always remains pure. But let us study what happens when one puts into contact *two systems*, both originally in pure states, and afterwards separates them.

Let us therefore consider two systems, I and II, originally separated. Let x be the ensemble of the coordinates of I and y the coordinates of II. *Each of the two systems* is assumed to be in a *pure state* given by its wave function:

$$\psi(x) = \sum_k \psi_k u_k(x) \qquad (\textit{system I}),$$

$$\phi(y) = \sum_\rho \phi_\rho v_\rho(y) \qquad (\textit{system II}).$$

We have expressed the functions ψ and ϕ as series built on the orthogonal functions $u_k(x)$ and $v_\rho(y)$. The coefficients ψ_k and ϕ_ρ depend only on time.

Although the two systems are originally taken to be separated, we can nevertheless describe them by a combined wave function $\Psi(x, y)$, whose evolution is governed by a combined Hamiltonian. The fact that the two systems are isolated from each other is expressed in the form of the Hamiltonian. It is the *sum* of two terms, each depending only upon the coordinates of one of the two systems; thus,

$$H(x, p_x, y, p_y) = H_I(x, p_x) + H_{II}(y, p_y).$$

It is easily verified that the combined wave function $\Psi(x, y)$, which unites the statistics contained in $\psi(x)$ and $\phi(y)$, is the simple product of the two separate

wave functions. It obeys the wave equation with Hamiltonian $H_I + H_{II}$. One has therefore *before contact*

$$\Psi(x, y) = \psi(x)\phi(y) = \sum_{k,\rho} \psi_k \phi_\rho u_k(x) v_\rho(y). \tag{1}$$

Let us now bring the two systems into contact. That makes it necessary to add to H an *interaction* term, $H_i(x, p_x, y, p_y)$, containing the two sets of variables x, p_x and y, p_y in a form that is not simply additive. It is evident that the combined function Ψ for the pair of systems will no longer in general keep the form of a product of two functions each depending on a single set of variables. However, it naturally can be expanded *at each instant as a series of products* $u_k(x)v_\rho(y)$ with coefficients that will depend on the instant chosen. Thus, if $u_k(x)$ and $v_\rho(y)$ form *complete* systems of orthogonal functions in their own domains, x and y, the products $u_k(x)v_\rho(y)$ also form an orthogonal system that is complete in the space of functions on the domain (x, y) of the ensemble. *During or after the contact* the wave function will be written in every case in the general form,

$$\Psi(x, y) = \sum_{k,\rho} \Psi_{k\rho} u_k(x) v_\rho(y), \tag{2}$$

where the coefficients $\Psi_{k\rho}$ in the generic case will not have the special form of a product, $\psi_k\phi_\rho$. As we are always dealing with a unique wave function evolving in accord with a Schrödinger equation, we have *for the combined system a constantly pure case.* Its statistical matrix is an elementary matrix,

$$P_{k\rho,l\sigma} = \Psi_{k\rho}\Psi^*_{l\sigma}.$$

As one pair of indices, k, ρ, is needed here to characterize a state of the total system, the elements of the statistical matrix of this system will evidently depend on two pairs of indices, k, ρ and l, σ.

Let us now focus attention on the system I. What is its statistical matrix?

Let F be a function solely of the variables of system I and F_{kl} its representation in the coordinates $u_k(x)$. The mean or expectation value of F in the state Ψ will be given by

$$(\Psi, F\Psi) = \sum_{kl,\rho} \Psi^*_{k\rho}\Psi_{l\rho}F_{kl} = Trace(P^I F).$$

Therefore the matrix

$$P^I_{lk} = \sum_{\rho} \Psi_{l\rho}\Psi^*_{k\rho} \tag{3a}$$

plays the role of statistical matrix for system I. Similarly one obtains for system II the statistical matrix

$$P^{II}_{\sigma\rho} = \sum_k \Psi_{k\sigma}\Psi^*_{k\rho}. \tag{3b}$$

P^I and P^{II} are evidently no longer elementary matrices. We are dealing with mixtures. What are their components? And what are the relative probabilities or concentrations of these components?

We note that the $\Psi_{k\rho}$ are normalized in the space of the indices k, ρ; thus,

$$\sum_{k,\rho} |\Psi_{k\rho}|^2 = 1.$$

Therefore the magnitude $\overset{(\rho)}{\psi}$ with the components

$$\overset{(\rho)}{\psi}_k = \Psi_{k\rho} \Big/ \left[\sum_l |\Psi_{l\rho}|^2\right]^{1/2}$$

is a normalized state vector for system I and represents a pure case.

We can therefore give to P^I the following form,

$$P^I = \sum_\rho p_\rho (\overset{(\rho)}{\psi} \times \overset{(\rho)}{\psi}), \tag{3a}$$

where

$$p_\rho = \sum_l |\Psi_{l\rho}|^2 \tag{4a}$$

is the concentration with which the pure case $\overset{(\rho)}{\psi}$ is contained in the mixture I. In the same way one has

$$P^{II} = \sum_k p_k (\overset{(k)}{\phi} \times \overset{(k)}{\phi}), \tag{3b'}$$

where

$$\overset{(k)}{\phi}_\rho = \Psi_{k\rho} \Big/ \left[\sum_\sigma |\Psi_{k\sigma}|^2\right]^{1/2}$$

is a unit vector, representing a pure case in II; and

$$p_k = \sum_{\sigma} |\Psi_{k\sigma}|^2 \tag{4b}$$

(k)
is the concentration with which ϕ is contained in the mixture II.

While the combined system I + II, which we suppose isolated from the rest of the world, is and remains in a pure state, we see that during the interaction systems I and II individually transform themselves from pure cases into mixtures.

This is a rather strange result. In classical mechanics we are not astonished by the fact that a *maximal knowledge of a composite system* implies a *maximal knowledge of all its parts*. We see that this equivalence, which might have been considered trivial, does not take place in quantum mechanics. There a maximal knowledge of a composite system ordinarily implies only mixtures for the component parts—that is, a knowledge that is not maximal.

The mixtures represented by P^I and P^{II} naturally cannot express all that it is possible to know about the combined system I + II. It is evident that the elementary combination of statistical mixtures for two individual systems cannot by itself reproduce a pure case for the combined system. Thus the function $\Psi(x, y)$ for the combined system contains still other relations, to wit, *statistical correlations between* the components of the two mixtures I and II.

The fact that the description we obtain for each of the two individual systems does not have the character of a pure case warns us that we are *renouncing part of the knowledge contained in* $\Psi(x, y)$ when we calculate probabilities for each of the two individual systems separately. This renunciation expresses itself by the summation over the index ρ in the definition of P^I, where we abstract away from what might be known about the state ρ of system II and about its connection with system I. This loss of knowledge expresses itself by the appearance of probabilities, now understood in the ordinary sense of the word, as expression of the fact that our knowledge about the combined system is not maximal.

§10. Reversible and Irreversible Evolution

It is evidently necessary to make a characteristic distinction between two essentially different modes of evolution of an individual system, a distinction which has no analog in classical mechanics.

I. *Reversible* or "*causal*" transformations. These take place when the system is isolated. They can be described by the change with time of a ψ function (or of a *certain number* of distinct ψ functions when one deals with a mixture). If $\psi(t_0)$ represents a pure case at the instant t_0, its evolution can be written in the form

$$\psi(t) = T\psi(t_0), \tag{1}$$

with the operator

$$T = \exp[-iH(t - t_0)/\hbar] = \sum_n (1/n!)[-iH(t - t_0)/\hbar]^n. \tag{4b}$$

In the case of a mixture, the time dependence of the statistical operator

$$P(t) = \sum_\rho p_\rho \left(\overset{(\rho)}{\psi}(t) \times \overset{(\rho)}{\psi}(t) \right) = \sum_\rho p_\rho \left(\overset{(\rho)}{T\psi}(t_0) \times \overset{(\rho)}{T\psi}(t_0) \right)$$

is thus given by the equation

$$P(t) = TP(t_0)\tilde{T}^*,$$

where $\tilde{T}_{kl}$ is the abbreviation for the matrix T_{lk}. One verifies that

$$\tilde{T}^* = \sum_n (1/n!)[iH(t - t_0)/\hbar]^n$$

and that

$$T\tilde{T}^* = 1.$$

It is therefore a *unitary transformation* that characterizes a causal evolution. It transforms a pure case into a pure case.

A unitary transformation, keeping invariant—as it does—the trace of a tensor, does not change the value of the entropy; the quantity $Trace\,(P, \ln P)$ stays constant.

II. *Irreversible* transformations, which one might also call "*acausal.*" These take place only when the system in question (I) makes physical contact with another system (II). The total system, comprising the two systems (I + II), again in this case undergoes a reversible transformation so long as the combined system I + II is isolated. But if we fix our attention on system I, this system will undergo an irreversible transformation. If it was in a pure state before the contact, it will ordinarily be transformed into a mixture. If it was already a mixture, it will be transformed into another mixture, the entropy of which (§7, equation 1) will be increased. Once thus degraded, the system has no chance in and by itself ever to regain its initial degree of determination.

We shall see specifically that measurement processes bring about an irreversible transformation of the state of the measured object, such that the initial statistical operator for a pure case

$$[\psi(q) = \sum_n \psi_n u_n(q)],$$

$$P = (\psi \times \psi),$$

is transformed, by interaction with the measuring apparatus, into the mixture,

$$P' = \sum |\psi_n|^2 (u_n \times u_n),$$

the u_n being the eigenfunctions of the quantity measured. The transition from P to P' clearly cannot be represented by a unitary transformation. It is associated with an increase of the entropy from 0 to $-k \sum_n |\psi_n|^2 \ln |\psi_n|^2$, which cannot come about by a unitary transformation.

The distinction between these two modes of evolution has no analog in classical mechanics, where it is always possible to give a maximal description of the state of an object by its $2f$ coordinates and momenta $q_1, q_2 \cdots p_f$, whatever the interactions between systems.

§11. Measurement and Observation. The Act of Objectification

We are now ready to analyze what happens in the act of measurement. We will first outline a protocol for this process and then verify in the following section that it describes properly the typical course of a measurement.

Suppose that we want to measure the quantity $F(x, p_x)$ of a system ("the object") given to be in the state $\psi = \sum_k \psi_k u_k(x)$ where u_k is an eigenfunction corresponding to the value f_k of F. We couple it with an apparatus capable of measuring F.

Let $G(y, p_y)$ be the coordinate specifying the position of the "needle" of the measuring device, and $g_0, g_1 \ldots g_\rho$ its eigenvalues, with eigenfunctions $v_0(y)$, $v_1(y) \ldots v_\rho(y)$. The state $v_0(y)$ corresponds to the zero of this apparatus.

Before the coupling we will attribute to the combined system a collective wave function of the form,

$$\Psi(x, y) = v_0(y) \sum_k \psi_k u_k(x). \tag{1}$$

This is a pure case for each of the two individual systems. *After the interaction* the wave function will be of a more general character,

$$\Psi(x, y) = \sum_{k,\rho} \Psi_{k\rho} u_k(x) v_\rho(y). \tag{2}$$

But an arbitrary interaction does not provide a measurement. In order for it to do so, it is necessary that it disturb the state of the object as little as possible and in addition that it should let one deduce from a g_ρ the corresponding f_k. Thus the values of the measurement scale g_ρ should be coordinated one-to-one with the values f_k of the quantity under consideration, so that one can inscribe directly onto the g_ρ scale the corresponding values of F. That occurs by replacing the index $\rho_{(k)}$ by k. Thus $g_k = g_{\rho(k)}$ will correspond to f_k.

We will see (§12) that after a measurement of the quantity the wave function takes the special form,

$$\Psi(x, y) = \sum \psi_k u_k(x) v_k(y). \tag{2'}$$

According to the preceding section, this function represents a state of the combined system that has for each separate system, object, and apparatus, the character of a *mixture*. According to §9, equation (4), the quantity $p_k = |\psi_k|^2$ gives the probability of finding the object in the pure state u_k with $F = f_k$; and the same quantity $p_k = |\psi_k|^2$ also gives the probability of finding the apparatus indication $G = g_k$. Moreover, we have a correlation between the two mixtures: specifically, we know with certainty that if $G = g_k$, then $F = f_k$. But of course quantum mechanics does not allow us to predict *which* value will actually be found in the measurement. The interaction with the apparatus does not put the object into a new pure state. Alone, it does not confer on the object a new wave function. On the contrary, it actually gives nothing but a statistical mixture: It leads to one mixture for the object and one mixture for the apparatus. For either system regarded individually there results uncertainty, incomplete knowledge. Yet nothing prevents our reducing this uncertainty by further observation. And this is our opportunity.

So far we have only coupled one apparatus with one object. But a coupling, even with a measuring device, is not yet a measurement. A measurement is achieved only when the position of the pointer has been *observed*. It is precisely this increase of knowledge, acquired by observation, that gives the observer the right to choose among the different components of the mixture predicted by theory, to reject those which are not observed, and to attribute thenceforth to the object a new wave function, that of the pure case which he has found.

We note the essential role played by the consciousness of the observer in this transition from the mixture to the pure case. Without his effective intervention, one would never obtain a new ψ function. In order to see this point clearly, let us consider the ensemble of three systems, (*object* x) + (*apparatus* y) + (*observer* z), as a combined and unique system. We will describe it by a global wave function with a form analogous to (2),

$$\Psi(x, y, z) = \sum_k \psi_k u_k(x) v_k(y) w_k(z),$$

where the w_k represent the different states of the observer.

"Objectively"—that is, *for us* who consider as "object" the combined system x, y, z—the situation seems little changed compared to what we just met when we were considering only apparatus and object. We now have three mixtures, one for each system, with those statistical correlations between them that are tied to a pure case for the combined system. Thus the function $\Psi(x, y, z)$ represents a maximal description of the combined "object," consisting of the actual object x,

252 LONDON, BAUER

the apparatus y, and the observer z; and nevertheless we do not know in what state the object x is.

The observer has a completely different impression. For him it is only the object x and the apparatus y that belong to the external world, to what he calls "objectivity." By contrast he has *with himself* relations of a very special character. He possesses a characteristic and quite familiar faculty which we can call the "faculty of introspection." He can keep track from moment to moment of his own state. By virtue of this "immanent knowledge" he attributes to himself the right to create his own objectivity—that is, to cut the chain of statistical correlations summarized in $\sum_k \psi_k u_k(x) v_k(y) w_k(z)$ by declaring, "I am in the state w_k" or more simply, "I see $G = g_k$," or even directly, "$F = f_k$."

Accordingly, we will label this creative action as "making objective." By it the observer establishes his own framework of objectivity and acquires a new piece of information about the object in question.*

Thus it is not a mysterious interaction between the apparatus and the object that produces a new ψ for the system during the measurement. It is only the consciousness of an "I" who can separate himself from the former function $\Psi(x, y, z)$ and, by virtue of his observation, *set up a new objectivity* in attributing to the object henceforward a new function $\psi(x) = u_k(x)$.

Neither is it some ignorance as to the state of the observer that creates quantum indeterminacy. *On the contrary*, in assuming a pure case for the combined system, we have implicitly presupposed an equally perfect knowledge of the initial state $w_0(z)$ of the observer and of the apparatus $v_0(y)$, that is, maximal information. Moreover, we have assumed that the observer can keep track perfectly of his own state.

Of course there might also be restrictions on the immanent knowledge of the observer. But these, if they existed, would in any case have nothing to do with quantum indeterminism; they would be additional restrictions of a completely different character. Moreover, it is not ordinarily required for a discussion of the measuring process that one should have an *all-encompassing* knowledge of the state of the observer; for example, there is little chance of making a big mistake if one does not know his age.

§12. An Example of Measurement

It only remains for us now to verify the protocol for measurement that we have just discussed. Let us take as a typical example the determination of the value of

* This paragraph is new. We have translated it from a typed addition inserted by Professor Fritz London in his own copy of the printed book, kindly sent to us October 24, 1980, by Mrs. Fritz London.—Eds.

one component of the magnetic moment of an atom by the method of Stern and Gerlach. The formulas that we will get can be generalized without difficulty to an arbitrary measurement.

This measurement is made, as is well known, by observing the motion of an atom through a nonuniform magnetic field. The field points in the direction along which one wants to determine the component of the magnetic moment of the atom. The coordinates, y, of the center of gravity of the atom play the role of the pointer reading $G(y, \beta_y)$. The internal coordinates of the atom, relative to the center of gravity, serve as object coordinates x. Specifically, we are concerned with the component, $M = M(x, p_x)$, of the magnetic moment in the direction of the field.

Let us write the wave equation for this problem in the form

$$\{-(\hbar^2/2m)\Delta_y + H_0(x, \partial/\partial x) + [M(x, p_x), F(y)]\}\Psi(x, y) = i\hbar\dot{\Psi}(x, y). \quad (1)$$

Here H_0 is the Hamiltonian operator for the field-free atom after one has separated off the center-of-gravity variables, $-(\hbar^2/2m)\Delta_y$ is the operator of the "apparatus" corresponding to the kinetic energy of the center of gravity, (M, F) is the contribution arising from the magnetic field F, and M is the operator for the magnetic moment of the atom in the direction of the field.

So long as the field F is constant—that is, so long as it does not depend on the coordinates y—the variables x and y can be separated in equation (1). For the different states of the "object" we must deal with the eigenvalue problem,

$$(H_0 + MF)u_k(x) = E_k u_k(x). \quad (2)$$

Let us limit ourselves to the lowest state of the atom, a state which we assume to be degenerate, with its components splitting in the field in proportion to the field strength F,

$$E_k = E_0 + (k\mu/j)F. \quad (3)$$

Here μ is the magnetic moment of the atom and k is the magnetic quantum number, $k = j, j - 1, j - 2, j - 3 \ldots -j$, where $j\hbar$ represents the total magnetic moment of the atom. When the field is no longer constant, $F = F(y)$, equation (2) contains the coordinates y as *parameters*. Consequently the eigenvalues of (2),

$$E_k(y) = E_0 + (k\mu/j)F(y), \quad (3a)$$

will also depend on these parameters y; likewise the eigenfunctions $u_k(x)$, which will be written more appropriately as $u_k(x, y)$. In practical terms the perturbation

of the u_k by the nonuniformity of the field is so slight that we can forget this dependence of the u_k on the parameter y.

Let us now develop $\Psi(x, y)$ in a series in terms of the functions $u_k(x)$,*

$$\Psi(x, y) = \sum_k v_k(t, y)u_k(x),$$

and let us introduce this development into the Schrödinger equation (1). After multiplication by u_k and integration over x we find for the coefficients $v_k(t, y)$ the following equations:

$$\{-(\hbar^2/2m)\Delta_y + E_k(y)\}v_k(y) = i\hbar\dot{v}_k(y). \tag{4}$$

They are still of the Schrödinger type but now refer solely to the motion of the center of gravity. In a typical such equation, the eigenvalue (3a) of (2), $E_k(y)$, plays the role of the potential. Let us now consider an atomic beam and let us develop $E_k(y)$ in the neighborhood of this beam,

$$E_k = E_0 + (k\mu/j)[F_0 + (y, grad\ F) + \cdots].$$

In first approximation the "potential" $E_k(y)$ varies linearly; thus it behaves like the potential of gravity near the earth's surface. Consequently equation (4) is nothing other than the equation of free fall. But in it the acceleration is proportional to the quantum number k. Thus the acceleration depends on the value of the component of the magnetic moment in the direction of the field, and can be *positive* or *negative*. Therefore it is easy to foresee in general terms and without detailed calculation the shape of the various trajectories which come from a well-collimated and initially monokinetic source and then travel through the nonuniform magnetic field. A single beam v_{00} splits into separate beams belonging to different values of k. The cross-hatched regions in Figure 1 show where the functions v_k are appreciably different from zero.

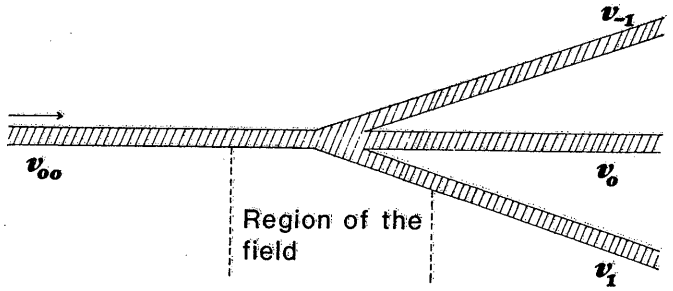

FIGURE 1. Splitting of the atomic beam in a non-uniform field.

* This is certainly permissible. The $v_k(y, t)$ show up in this development, not as given eigenfunctions, but as the still unknown coefficients of the functions $u_k(x)$.

So far we have been discussing the not-yet-normalized solutions of the auxiliary equations (4). A complete solution of the Schrödinger equation (1) will have the form

$$\Psi(x, y) = \sum_k c_k v_k(y) u_k(x). \tag{5}$$

Let

$$\psi = \sum_k \psi_k u_k(x)$$

with

$$\sum_k |\psi_k|^2 = 1$$

be the state of the atom *before the measurement* in the region at the left of the figure, where we can still separate off the motion of the center of gravity and where all the functions $v_k(y)$ are identical $[= v_{00}(y)]$.

The total wave function of object-plus-apparatus *before entry into the region of the field* is

$$\Psi(x, y) = v_{00}(y) \sum_k \psi_k u_k(x).$$

On the other side of the field region, according to (5) and the principle of continuity, it is

$$\Psi'(x, y) = \sum_k \psi_k u_k(x) v_k(y).$$

A measurement of the y coordinate of the center of gravity of an atom that has crossed the field is equivalent to a determination of k because each $v_k(y)$ differs from zero only in a limited region, fixed by k, and because a determination of k is equivalent to a knowledge of the component of the magnetic moment parallel to f. This is exactly the type of statistical link that we presumed in §11.

Let us take, for example, an atom of total angular momentum 1, for which k is restricted to the values 1, 0, -1. The effect of the measurement shows up in the transition of the coefficients $\Psi_{k\rho}$ of the wave function of the combined system from

$$\Psi_{k\rho} = \psi_k \delta_{0\rho} = \begin{Vmatrix} 0 & 0 & 0 \\ \psi_{-1} & \psi_0 & \psi_1 \\ 0 & 0 & 0 \end{Vmatrix} \quad \textit{(before the measurement)}$$

to

$$\Psi'_{k\rho} = \psi_k \delta_{k\rho} = \begin{Vmatrix} \psi_{-1} & 0 & 0 \\ 0 & \psi_0 & 0 \\ 0 & 0 & \psi_1 \end{Vmatrix} \quad (\textit{after the measurement}).$$

Before the measurement the statistical matrices P^I and P^{II} have the form of two pure cases,

$$P^I_{kl} = \begin{Vmatrix} \psi_{-1}\psi^*_{-1} & \psi_{-1}\psi^*_0 & \psi_{-1}\psi^*_1 \\ \psi_0\psi^*_{-1} & \psi_0\psi^*_0 & \psi_0\psi^*_1 \\ \psi_1\psi^*_{-1} & \psi_1\psi^*_0 & \psi_1\psi^*_1 \end{Vmatrix}$$

and

$$P^{II}_{\rho\sigma} = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{Vmatrix}.$$

The action of the magnetic field transforms them into mixtures,

$$(P^I_{kl})' = \begin{Vmatrix} |\psi_{-1}|^2 & 0 & 0 \\ 0 & |\psi_0|^2 & 0 \\ 0 & 0 & |\psi_1|^2 \end{Vmatrix}$$

and

$$(P^{II}_{\rho\sigma})' = \begin{Vmatrix} |\psi_{-1}|^2 & 0 & 0 \\ 0 & |\psi_0|^2 & 0 \\ 0 & 0 & |\psi_1|^2 \end{Vmatrix}.$$

Here the components have a one-to-one correspondence, implying the associated correlations. For the combined system the transition $\Psi \to \Psi'$ is a unitary transformation,

$$\Psi'_{k\rho} = \sum_{l,\sigma} T_{k\rho,l\sigma} \Psi_{l\sigma},$$

where

$$T_{k\rho,l\sigma} = \delta_{kl} \delta_{\rho,l+\sigma},$$

it being understood that

$$\rho = \pm 2 \quad \text{is equivalent to} \quad \rho = \mp 1.$$

§13. Individuality and Pure Case

A measuring arrangement, such as we have just finished describing in the previous section, can be used to "filter" through objects possessing a prescribed value of a certain physical quantity. It is enough to make a suitable slit through the screen on which the atoms fall to transform the Stern-Gerlach apparatus into a source of identically oriented atoms. It thus becomes a good set-up for producing pure cases.

Does not this conclusion contradict what we just learned (§11), to the effect that a pure case is brought about only by an "act of objectifying," accomplished by the appropriate intervention of an observer? It is necessary, however, to be more precise: the filter never puts any *individual* object whatsoever into a new pure state. It can only put it into a mixed state. This is what always happens when one couples one system with another. By contrast, we can evidently say that the states of atoms which have gone through the slit have the desired property. We can attribute to them the wave function of the pure case in question. But this attribution only works out, so to speak, at the expense of the individuality of the object, as one does not know in advance *which* are the atoms that have the property in question. We can easily attribute to the objects that get through the slit the ψ function of the pure case, but we cannot say *which object*, that is, which variable is the argument of this ψ function. Without a supplementary check by an observer, it is not possible to guarantee whether a given atom has gone through the filter or been caught in it. The filter alone thus truly produces pure cases, but in an absolutely *anonymous* form. Of course we can attribute to these cases afresh some names of their own, for example by numbering in sequence the atoms that really get through the slit. But that is no different from a true measurement, and we would be led back in that way to what we have already discussed.

Moreover, anonymous objects are precisely the focus of one's interest in many experiments. The majority of the measurements in atomic physics really do not deal with an *individual system*; rather, they seek to find out the general properties of an *entire species* of atoms—or of molecules, or of elementary particles. Thus for example the Stern-Gerlach set-up just discussed is ordinarily used, not to measure a component of the spin of an *individual atom*, but rather to determine the spin *of the silver atom*.

Quantum mechanics, truly a "theory of species," is perfectly adapted to this experimental task. But given that every measurement contains a macroscopic process, unique and separate, we can hardly escape asking ourselves to what extent and within what limits the everyday concept of an individual object is still recognizable in quantum mechanics.

§14. Scientific Community and Objectivity

At first sight it would appear that in quantum mechanics the concept of scientific objectivity has been strongly shaken. Since the classic period, the idea has become familiar that a physical object is something real, existing outside of the observer, independent of him, and in particular independent of whether or not the object has been subjected to measurement. The situation is not the same in quantum mechanics. Far from it being possible to attribute to a system at every instant its measurable properties, one cannot even claim that to attribute to it so much as a wave function has a well-defined meaning, unless referring explicitly to a definite measurement. Moreover, it looks as if the result of a measurement is intimately linked to the consciousness of the person making it, and as if quantum mechanics thus drives us toward complete solipsism.

Actually, however, we know that the relations between physicists have undergone practically no change since the discovery of quantum mechanics. No physicist has retired into a solipsistic isolation. Physicists use the same means of scientific exchange as in the past and are capable of cooperation in studying the same object. Thus there really exists something like a community of scientific perception, an agreement on what constitutes the object of the investigation, and it is this that still has to be looked into.

First of all, it is easy to recognize that the act of observation, that is, the coupling between the measuring apparatus and the observer (see our example in §11), is truly a *macroscopic* action and not basically quantal. Consequently one always has the right to neglect the effect on the apparatus of the "scrutiny" of the observer. Tracing things back in time, one will obtain definite conclusions about the state of the apparatus (or the photographic plate) and consequently the state of the object before the observation (*but of course after the coupling* is turned off). Moreover, nothing prevents another observer from looking at the same apparatus; and one can predict that, barring errors, his observations will be the same. The possibility of abstracting away from the individuality of the observer and of creating a collective scientific perception therefore in no way comes seriously into question.

It might appear that the scientific community thus created is a kind of *spiritualistic society* which studies imaginary phenomena—that the objects of physics are phantoms produced by the observer himself. In classical physics, one can picture a system at every instant in a unique and continuous way by the set of all of its measurable properties, even when it is not subjected to observation. It is exactly the possibility of this continuity of connection between properties and object that has ordinarily been considered as proof that physics deals with something "real," that is, having in principle an existence "independent of all observers." In quantum mechanics an object is the carrier, not of a definite set of measurable properties, but only of a set of "potential" probability distributions or statistics

(§3) referring to measurable properties, statistics which only come into force on the occasion of an effective, well-defined measurement. If one abstracts away from all acts of measurement, it is meaningless to claim these measurable properties as realized; the very mathematical form of the statistics does not allow it (see §8).

But that does not keep us from predicting or interpreting experimental results. Theory fixes the rules. It teaches us first of all how to filter an object to get a pure case—that is, reproducible conditions; then it suggests how to make measurements, either to check theoretical predictions or to discover new empirical regularities. The theory adapts itself truly marvelously to the realities of experiment. It gives answers on all desired details and is silent on hypothetical questions without experimental meaning.

In present physics the concept of "objectivity" is a little more abstract than the classical idea of a material object. Is it not a guarantee of "*the objectivity*" *of an object* that one can at least formally attribute measurable properties to it in a *continuous manner* even at times when it is not under observation? The answer is No, as this new theory shows by its internal consistency and by its impressive applications. It is enough, evidently, that the properties of the object should be present at the moment they are measured and that they should be predicted by theory in agreement with experiment.

In the limiting case of macroscopic phenomena, quantum theory rejoins classical theory. Thus it justifies the use of the "naive" concept of "objectivity" and at the same time specifies the limitations of this concept.

What has just been said relates to an important philosophical problem that we cannot enter into here: the determination of the necessary and sufficient conditions for an object of thought to possess objectivity and to be an object of science. This problem was perhaps posed for the first time in any general way by such mathematicians as Malebranche, Leibniz, and especially by B. Bolzano. More recently Husserl (1901, 1913; see also the rather similar ideas in Cassirer, 1910, 1936) has systematically studied such questions and has thus created a new method of investigation called "Phenomenology."

Physics insofar as it is an empirical science cannot enter into such problems in all their generality. It is satisfied to use philosophical concepts *sufficient* for its needs; but on occasion it can recognize that some of the concepts that once served it have become *quite unnecessary*, that they contain elements that are useless and even incorrect, actual obstacles to progress. One can doubt the possibility of establishing philosophical truths by the methods of physics, but it is surely not outside the competence of physicists to demonstrate that *certain statements which pretend to have a philosophical validity do not.* And sometimes these "negative" philosophical discoveries by physicists are no less important, no less revolutionary for philosophy than the discoveries of recognized philosophers.