mechanics, it was possible for the author to prove the corresponding theorem without such assumptions.[206] Since the detailed discussion of this subject, as well as of the ergodic theorem closely connected with it (cf. the reference in Note 206, where this theorem is also proved) would go beyond the scope of this volume, we cannot report on these investigations. The reader who is interested in this problem can refer to the treatments in the references.

have been formulated by W. Pauli (Sommerfeld-Festschrift, 1928), and the H-theorem is proved there with their help. More recently, the author also succeeded in proving the classical-mechanical ergodic theorem, cf. Proc. Nat. Ac., Jan. and March, 1932, as well as the improved treatment of G. D. Birkhoff, Proc. Nat. Ac., Dec. 1931 and March, 1932.

[206] Z. Physik, 57 (1929).

CHAPTER VI

THE MEASURING PROCESS

1. FORMULATION OF THE PROBLEM

In the discussions so far, we have treated the relation of quantum mechanics to the various causal and statistical methods of describing nature. In the course of this we found a peculiar dual nature of the quantum mechanical procedure which could not be satisfactorily explained. Namely, we found that on the one hand, a state ϕ is transformed into the state ϕ' under the action of an energy operator H in the time interval $0 \leq \tau \leq t$:

$$\frac{\partial}{\partial t}\phi_\tau = -\frac{2\pi i}{h} H \phi_\tau \qquad (0 \leq \tau \leq t),$$

so if we write $\phi_0 = \phi$, $\phi_t = \phi'$, then

$$\phi' = e^{-\frac{2\pi i}{h} tH}\phi$$

which is purely causal. A mixture U is correspondingly transformed into

$$U' = e^{-\frac{2\pi i}{h} tH}\, U e^{\frac{2\pi i}{h} tH}.$$

Therefore, as a consequence of the causal change of ϕ

into ϕ', the states $U = P_{[\phi]}$ go over into the states $U' = P_{[\phi']}$ (process **2.** in V.1.). On the other hand, the state ϕ -- which may measure a quantity with a pure discrete spectrum, distinct eigenvalues and eigenfunctions $\phi_1, \phi_2, \ldots$ -- undergoes in a measurement a non-causal change in which each of the states $\phi_1, \phi_2, \ldots$ can result, and in fact does result with the respective probabilities $|(\phi, \phi_1)|^2, |(\phi, \phi_2)|^2, \ldots$. That is, the mixture

$$U' = \sum_{n=1}^{\infty} |(\phi, \phi_n)|^2 P_{[\phi_n]}$$

obtains. More generally, the mixture U goes over into

$$U' = \sum_{n=1}^{\infty} (U\phi_n, \phi_n) P_{[\phi_n]}$$

(process **1.** in V.1.). Since the states go over into mixtures, the process is not causal.

The difference between these two processes $U \longrightarrow U'$ is a very fundamental one: aside from the different behaviors in regard to the principle of causality, they are also different in that the former is (thermodynamically) reversible, while the latter is not (cf. V.3.).

Let us now compare these circumstances with those which actually exist in nature or in its observation. First, it is inherently entirely correct that the measurement or the related process of the subjective perception is a new entity relative to the physical environment and is not reducible to the latter. Indeed, subjective perception leads us into the intellectual inner life of the individual, which is extra-observational by its very nature (since it must be taken for granted by any conceivable observation or experiment). (Cf. the discussion above.) Nevertheless, it is a fundamental requirement of the scientific viewpoint -- the so-called principle of the psycho-physical parallelism -- that it must be possible so to describe the extra-physical process of the subjective perception as if it were in reality in the physical world -- i.e., to assign to its parts equivalent physical processes in the objective environment, in ordinary space. (Of course, in this correlating procedure there arises the frequent necessity of localizing some of these processes at points which lie within the portion of space occupied by our own bodies. But this does not alter the fact of their belonging to the "world about us," the objective environment referred to above.) In a simple example, these concepts might be applied about as follows: We wish to measure a temperature. If we want, we can pursue this process numerically until we have the temperature of the environment of the mercury container of the thermometer, and then say: this temperature is measured by the thermometer. But we can carry the calculation further, and from the properties of the mercury, which can be explained in kinetic and molecular terms, we can calculate its heating, expansion, and the resultant length of the mercury column, and then say: this length is seen by the observer. Going still further, and taking the light source into consideration, we could find out the reflection of the light quanta on the opaque mercury column, and the path of the remaining light quanta into the eye of the observer, their refraction in the eye lens, and the formation of an image on the retina, and then we would say: this image is registered by the retina of the observer. And were our physiological knowledge more precise than it is today, we could go still further, tracing the chemical reactions which produce the impression of this image on the retina, in the optic nerve tract and in the brain, and then in the end say: these chemical changes of his brain cells are perceived by the observer. But in any case, no matter how far we calculate -- to the mercury vessel, to the scale of the thermometer, to the retina, or into the brain, at some time we must say: and this is perceived by the observer. That is, we must always divide the world into two parts, the one being the observed system, the other the observer. In the former, we can follow up all physical processes (in principle at least) arbitrarily precisely. In the latter, this is meaningless. The boundary between the two is arbitrary to a very large extent. In particular we saw in the four different possibilities in the example above, that the observer in this sense needs not to become identified with the body of the actual observer: In one instance in the above example, we included even the thermometer in it, while in another instance, even the eyes and optic nerve tract were not included. That this boundary can be pushed arbitrarily deeply into the interior of the body of the actual observer is the content of the principle of the psycho-physical parallelism -- but this does not change the fact that in each method of description the boundary must be put somewhere, if the method is not to proceed vacuously, i.e., if a comparison with experiment is to be possible. Indeed experience only makes statements of this type: an observer has made a certain (subjective) observation; and never any like this: a physical quantity has a certain value.

Now quantum mechanics describes the events which occur in the observed portions of the world, so long as they do not interact with the observing portion, with the aid of the process **2.** (V.1.), but as soon as such an interaction occurs, i.e., a measurement, it requires the application of process **1.** The dual form is therefore justified.[207] However, the danger lies in the fact that

[207]N. Bohr, Naturwiss. 17 (1929), was the first to point out that the dual description which is necessitated by the formalism of the quantum mechanical description of nature is fully justified by the physical nature of things that it may be connected with the principle of the psycho-physical parallelism.

the principle of the psycho-physical parallelism is violated, so long as it is not shown that the boundary between the observed system and the observer can be displaced arbitrarily in the sense given above.

In order to discuss this, let us divide the world into three parts: I, II, III. Let I be the system actually observed, II the measuring instrument, and III the actual observer.[208] It is to be shown that the boundary can just as well be drawn between I and II + III as between I + II and III. (In our example above, in the comparison of the first and second cases, I was the system to be observed, II the thermometer, and III the light plus the observer; in the comparison of the second and third cases, I was the system to be observed plus the thermometer, II the light plus the eye of the observer, III the observer, from the retina on; in the comparison of the third and fourth cases, I was everything up to the retina of the observer, II his retina, nerve tracts and brain, III his abstract "ego.") That is, in one case **2.** is to be applied to I, and **1.** to the interaction between I and II + III; and in the other case, **2.** to I + II, and **1.** to the interaction between I + II and III. (In each case, III itself remains outside of the calculation.) The proof of this assertion, that both procedures give the same results regarding I (this and only this belongs to the observed part of the world in both cases), is then our problem.

But in order to be able to accomplish this successfully, we must first investigate more closely the process of forming the union of two physical systems (which leads from I and II to I + II).

[208]The discussion which is carried out in the following, as well as that in VI.3., contains essential elements which the author owes to conversations with L. Szilard. Cf. also the similar considerations of Heisenberg, in the reference cited in Note 181.

2. COMPOSITE SYSTEMS

As was stated at the end of the preceding section, we consider two physical systems I, II (which do not necessarily have the meaning of the I, II above), and their combination I + II . In the classical mechanical method of description, I would have k degrees of freedom, and therefore the coordinates $q_1, \ldots, q_k$, in place of which we shall use the one symbol q ; correspondingly, let II have l degrees of freedom, and the coordinates $r_1, \ldots, r_l$ which shall be denoted by r . Therefore, I + II has k + l degrees of freedom and the coordinates $q_1, \ldots, q_k, r_1, \ldots, r_l$, or, more briefly, q, r . In quantum mechanics then, the wave functions of I have the form $\phi(q)$, those of II the form $\xi(r)$ and those of I + II the form $\Phi(q, r)$. In the corresponding Hilbert spaces $\mathfrak{R}^{I}, \mathfrak{R}^{II}, \mathfrak{R}^{I+II}$, the inner product is defined by $\int \phi(q)\overline{\psi(q)}\,dq$, $\int \xi(r)\overline{\eta(r)}\,dr$ and $\iint \Phi(q, r)\Psi(q, r)\,dq\,dr$ respectively. The physical quantities of I, II, I + II are correspondingly the (hypermaximal) Hermitian operators A, A, and $\mathcal{A}$ in $\mathfrak{R}^{I}, \mathfrak{R}^{II}$, and $\mathfrak{R}^{I+II}$ respectively.

Each physical quantity in I is naturally also one in I + II , and in fact its $\mathcal{A}$ is to be obtained from its A in this way: to obtain $\mathcal{A}\,\Phi(q, r)$ consider r as a constant and apply A to the q function $\Phi(q, r)$.[209] This rule of transformation is correct in any case for the coordinate and momentum operators $Q_1, \ldots, Q_k$ and $P_1, \ldots, P_k$, i.e.,

$$q_1, \ldots, q_k, \frac{h}{2\pi i}\frac{\partial}{\partial q_1}, \ldots, \frac{h}{2\pi i}\frac{\partial}{\partial q_k}$$

(cf. I.2.), and it conforms with the principles **I.**, **II.** in

[209]It can easily be shown that if A is Hermitian or hypermaximal, $\mathcal{A}$ is also.

IV.2.[210] We therefore postulate them generally. (This is the customary procedure in quantum mechanics.)

In the same way, each physical quantity in II is also one in I + II , and its A gives its $\mathcal{A}$ by the same rule: $\mathcal{A}\,\Phi(q, r)$ equals $A\Phi(q, r)$ if in the latter expression, q is taken as constant, and $\Phi(q, r)$ is considered as a function of r .

If $\phi_m(q)$ (m = 1,2,...) is a complete orthonormal set in $\mathfrak{R}^{I}$ and $\xi_n(r)$ (n = 1,2,...) one in $\mathfrak{R}^{II}$, then $\Phi_{m|n}(q, r) = \phi_m(q)\xi_n(r)$ (m, n = 1,2,...) is clearly one in $\mathfrak{R}^{I+II}$. The operators A, A, $\mathcal{A}$ can therefore be represented by matrices $\{a_{m|m'}\}$, $\{a_{n|n'}\}$, and $\{\alpha_{mn|m'n'}\}$ respectively (m, n', n, n' = 1,2,...).[211] We shall make frequent use of this. The matrix representation means that

$$A\phi_m(q) = \sum_{m'=1}^{\infty} a_{m|m'}\phi_{m'}(q), \quad A\xi_n(r) = \sum_{n'=1}^{\infty} a_{n|n'}\xi_{n'}(r)$$

and

$$\mathcal{A}\Phi_{mn}(q, r) = \sum_{m',n'=1}^{\infty} \alpha_{mn|m'n'}\Phi_{m'n'}(q, r) ,$$

i.e.,

[210]For **I.** this is clear, and for **II.** also, so long as only polynomials are concerned. For general functions, it can be inferred from the fact that the correspondence of a resolution of the identity and a Hermitian operator is not disturbed in our transition $A \longrightarrow \mathcal{A}$.

[211]Because of the large number and variety of indices, we use this method of denoting the matrices, which differs somewhat from the notation used thus far.

$$\mathcal{A}\,\phi_m(q)\xi_n(r) = \sum_{m'n'=1}^{\infty} \alpha_{mn|m'n'}\phi_{m'}(q)\xi_{n'}(r) .$$

In particular the correspondence $\mathsf{A} \longrightarrow \mathcal{A}$ means that

$$\mathcal{A}\phi_m(q)\xi_n(r) = (\mathsf{A}\,\phi_m(q))\xi_n(r) = \sum_{m'=1}^{\infty} a_{m|m'}\phi_{m'}(q)\xi_n(r) ,$$

i.e.,

$$\alpha_{mn|m'n'} = a_{m|m'}\delta_{n|n'} \quad \left(\delta_{n|n'} \begin{cases} = 1 , \text{ for } n = n' \\ = 0 , \text{ for } n \neq n' \end{cases}\right) .$$

In an analogous fashion, the correspondence $\mathrm{A} \longrightarrow \mathcal{A}$ implies that $\alpha_{mn|m'n'} = a_{n|n'}\delta_{m|m'}$.

A statistical ensemble in I + II is characterized by its statistical operator $\mathcal{U}$ or by its matrix $\{\upsilon_{mn|m'n'}\}$. This also determines the statistical properties of all quantities in I + II , and therefore the properties of the quantities in I also. Consequently there also corresponds to it a statistical ensemble in I alone. In fact, an observer who could perceive only I , and not II , would view the ensemble of systems I + II as one such of systems I . What is now the statistical operator U or its matrix $\{u_{m|m'}\}$, which belongs to this I ensemble? We determine it as follows: The I quantity with the matrix $\{a_{m|m'}\}$ has the matrix $\{a_{m|m'}\delta_{n|n'}\}$ as an I + II quantity, and therefore, by reason of a calculation in I , it has the expectation value

$$\sum_{m,m'=1}^{\infty} u_{m|m'}a_{m'|m} ,$$

while the calculation in I + II gives

$$\sum_{m,n,m',n'=1}^{\infty} \upsilon_{mn|m'n'}a_{m'|m}\delta_{n'|n} = \sum_{m,m',n=1}^{\infty} \upsilon_{mn|m'n'}a_{m'|m}$$

$$= \sum_{m,m'=1}^{\infty}\left(\sum_{n=1}^{\infty} \upsilon_{mn|m'n}\right)a_{m'm} .$$

In order that both expressions be equal, we must have

$$u_{m|m'} = \sum_{n=1}^{\infty} \upsilon_{mn|m'n} .$$

In the same way, our I + II ensemble, if only II is considered and I is ignored, determines a II ensemble, with a statistical operator U and matrix $\{\mu_{n|n'}\}$. By analogy, we obtain

$$u_{n|n'} = \sum_{m=1}^{\infty} \upsilon_{mn|mn'} .$$

We have thus established the rules of correspondence for the statistical operators of I, II, I + II , i.e., U , U, $\mathcal{U}$. They proved to be essentially different from those which control the correspondence between the operators A , A, $\mathcal{A}$ of physical quantities.

It should be mentioned that our U , U, $\mathcal{U}$ correspondence depends only apparently on the choice of the complete orthonormal sets $\phi_m(q)$ and $\xi_n(q)$. Indeed it was derived from an invariant condition (which is satisfied by this arrangement alone): Namely, from the requirement of agreement between the expectation values of A and of $\mathcal{A}$, or of those of A and of $\mathcal{A}$.

$\mathcal{U}$ expresses the statistics in I + II , U and U those statistics restricted to I or II respectively. There now arises the question: do U , U determine $\mathcal{U}$ uniquely or not? In general one will expect a negative answer because all "probability dependencies" which may exist between the two systems disappear as the information is reduced to the sole knowledge of U and U, i.e., of the separated systems I and II. But if one knows the state of I precisely, as also that of II, "probability questions" do not arise, and then I + II, too, is precisely known. An exact mathematical discussion is, however, preferable to these qualitative considerations, and we shall proceed to this.

The problem is, then: For two given definite matrices $\{u_{m|m'}\}$ and $\{u_{n|n'}\}$, find a third definite matrix $\{\upsilon_{mn|m'n'}\}$, such that

$$\sum_{n=1}^{\infty} \upsilon_{mn|m'n} = u_{m|m'}, \quad \sum_{m=1}^{\infty} \upsilon_{mn|mn'} = u_{n|n'} .$$

(From

$$\sum_{m=1}^{\infty} u_{m|m} = 1, \quad \sum_{n=1}^{\infty} u_{n|n} = 1,$$

it then follows directly that

$$\sum_{m,n=1}^{\infty} \upsilon_{mn|mn} = 1,$$

i.e., the correct normalization is obtained.) This problem is always solvable, for example, $\upsilon_{mn|m'n'} = u_{m|m'}u_{n|n'}$ is always a solution (it can easily be seen that this matrix is definite), but the question arises as to whether this is the only solution.

We shall show that this is the case if and only if at least one of the two matrices $\{u_{m|m'}\}$, $\{u_{n|n'}\}$ is a state. First we prove the necessity of this condition, i.e., the existence of several solutions if both matrices correspond to mixtures. In such a case (cf. IV.2.)

$$u_{m|m'} = \alpha v_{m|m'} + \beta w_{m|m'}, \quad u_{n|n'} = \gamma v_{n|n'} + \delta w_{n|n'} .$$

($v_{m|m'}$, $w_{m|m'}$ definite and $v_{n|n'}$, $w_{n|n'}$ also, differing by more than a constant factor,

$$\sum_{m=1}^{\infty} v_{m|m} = \sum_{m=1}^{\infty} w_{m|m} = \sum_{n=1}^{\infty} v_{n|n} = \sum_{n=1}^{\infty} w_{n|n} = 1$$

$\alpha, \beta, \gamma, \delta > 0$, $\alpha + \beta = 1$, $\gamma + \delta = 1$).

We easily verify that each

$$\upsilon_{mn|m'n} = \pi v_{m|m'}v_{n|n'} + \rho w_{m|m'}v_{n|n'} + \sigma v_{m|m'}w_{n|n'} + \tau w_{m|m'}w_{n|n'}$$

with

$$\pi + \sigma = \alpha, \quad \rho + \tau = \beta, \quad \pi + \rho = \gamma, \quad \sigma + \tau = \delta ,$$

$$\pi, \rho, \sigma, \tau > 0 ,$$

is a solution. Then π, ρ, σ, τ can be chosen in an infinite number of ways: Because of $\alpha + \beta = \gamma + \delta$ only three of the four equations are independent; therefore, $\rho = \gamma - \pi$, $\sigma = \alpha - \pi$, $\tau = (\delta - \alpha) + \pi$, and in order that all be > 0, we must require $\alpha - \delta = \gamma - \beta < \pi < \alpha, \gamma$, which is the case for infinitely many π. Now different π, ρ, σ, τ lead to different $\upsilon_{mn|m'n'}$, because the $v_{m|m'} \cdot v_{n|n'}, \ldots, w_{m|m'} \cdot w_{n|n'}$ are linearly independent, since the $v_{m|m'}$, $w_{m|m'}$ are such, as well as the $v_{n|n'}$, $w_{n|n'}$.

Next we prove the sufficiency, and here we may assume that $u_{m|m'}$ corresponds to a state (the other case is disposed of in the same way). Then $U = P_{[\phi]}$ and since the complete orthonormal set $\phi_1, \phi_2, \ldots$ was arbitrary, we can assume $\phi_1 = \phi$. $U = P_{[\phi_1]}$ has the matrix

$$u_{m|m'} \begin{cases} = 1, & \text{for } m = m' = 1 \\ = 0, & \text{otherwise} \end{cases} .$$

Therefore

$$\sum_{n=1}^{\infty} \upsilon_{mn|m'n} \begin{cases} = 1 \ , \text{ for } m = m' = 1 \\ \\ = 0 \ , \text{ otherwise} \end{cases} \Bigg| \ .$$

In particular, for $m \neq 1$,

$$\sum_{n=1}^{\infty} \upsilon_{mn|mn} = 0 \ ,$$

but since all $\upsilon_{mn|mn} \geq 0$ because of the definiteness of $\upsilon_{mn|m'n'}$ $[\upsilon_{mn|mn} = (U \Phi_{mn}, \Phi_{mn})]$, therefore in this case $\upsilon_{mn|mn} = 0$. That is, $(U \Phi_{mn}, \Phi_{mn}) = 0$, and hence, because of the definiteness of U , $(U \Phi_{mn}, \Phi_{m'n'})$ also $= 0$ (cf. II.5., THEOREM 19.), where m', n' are arbitrary. That is, it follows from $m \neq 1$ that $\upsilon_{mn|m'n'} = 0$, and because of the Hermitian nature, this also follows from $m' \neq 1$. For $m = m' = 1$ however, this gives

$$\upsilon_{1n|1n'} = \sum_{m=1}^{\infty} \upsilon_{mn|mn'} = u_{n|n'} \ .$$

Consequently, as was asserted, the solution $\upsilon_{mn|m'n}$ is determined uniquely.

We can thus summarize our result as follows: A statistical ensemble in I + II with the operator $U = \{\upsilon_{mn|m'n'}\}$ is determined uniquely by the statistical ensembles determined by it in I and II individually, with the respective operators $\mathrm{U} = \{u_{m|m'}\}$ and $\mathrm{U} = \{u_{n|n'}\}$, if and only if the following two conditions are satisfied:

1. $\upsilon_{mn|m'n'} = v_{m|m'} v_{n|n'}$. (From

$$\mathrm{Tr}\, U = \sum_{m,n=1}^{\infty} \upsilon_{mn|mn} = \sum_{m=1}^{\infty} v_{m|m} \sum_{n=1}^{\infty} v_{n|n} = 1 \ ,$$

it follows that, by multiplication of $v_{m|m'}$ and $v_{n|n'}$ with two reciprocal constant factors, we can obtain

$$\sum_{m=1}^{\infty} v_{m|m} = 1 \ , \quad \sum_{n=1}^{\infty} v_{n|n} = 1$$

But then we see that $u_{m|m'} = v_{m|m'}$, $u_{n|n'} = v_{n|n'}$.)

2. Either $v_{m|m'} = \bar{x}_m x_{m'}$, or $v_{n|n'} = \bar{x}_n x_n$. (Indeed $\mathrm{U} = P_{[\phi]}$ means that

$$\phi = \sum_{m=1}^{\infty} y_m \phi_m \ ,$$

and therefore $u_{m|m'} = \bar{y}_m y_{m'}$ and correspondingly for $v_{m|m'}$; by analogy the same is true with $\mathrm{U} = P_{[\xi]}$.)

We shall call U and U the projections of U in I and II respectively.[212]

We now apply ourselves to the states of I + II , $U = P_{[\Phi]}$. The corresponding wave functions $\Phi(q, r)$ can be expanded according to the complete orthonormal set $\Phi_{mn}(q, r) = \phi_m(q)\xi_n(r)$:

$$\Phi(q, r) = \sum_{m,n=1}^{\infty} f_{mn} \phi_m(q) \xi_n(r) \ .$$

We can therefore replace them by the coefficients f_{mn} $(m, n = 1,2,\ldots)$ which are subject only to the condition that

$$\sum_{m,n=1}^{\infty} |f_{mn}|^2 = \|\Phi\|^2$$

be finite.

[212] The projections of a state of I + II are in general mixtures in I or II ; cf. above. This circumstance was discovered by Landau, Z. Physik 45 (1927).

We can define two operators F, F^* by

(**F.**)
$$F\phi(q) = \int \overline{\Phi(q,\ r)}\,\phi(q)\,dq$$
$$F^*\xi(r) = \int \Phi(q,\ r)\,\xi(r)\,dr\ .$$

These are linear, but have the peculiarity of being defined in $\mathfrak{R}^{I}$ and $\mathfrak{R}^{II}$ respectively, and of taking on values from $\mathfrak{R}^{II}$ and $\mathfrak{R}^{I}$ respectively. Their relation is that of adjoints, since obviously $(F\phi,\ \xi) = (\phi,\ F^*\xi)$ (the inner product on the left is to be formed in $\mathfrak{R}^{II}$ and that on the right is to be formed in $\mathfrak{R}^{I}$). Since the difference of $\mathfrak{R}^{I}$ and $\mathfrak{R}^{II}$ is mathematically unimportant, we can apply the results of II.11: then, since we are dealing with integral operators, $\Sigma(F)$ and $\Sigma(F^*)$ are equal to

$$\iint |\Phi(q,\ r)|^2 dq\,dr = ||\Phi||^2 = 1 \quad (||\Phi|| \text{ in } \mathfrak{R}^{I+II}!)\ ,$$

and are therefore finite. Consequently F, F^* are continuous, in fact are completely continuous operators, and F^*F as well as FF^* are definite operators, $\mathrm{Tr}\,(F^*F) = \Sigma(F) = 1$, $\mathrm{Tr}\,(FF^*) = \Sigma(F^*) = 1$.

If we again consider the difference between $\mathfrak{R}^{I}$ and $\mathfrak{R}^{II}$ then we see that F^*F is defined and assumes values in $\mathfrak{R}^{I}$, and FF^* similarly in $\mathfrak{R}^{II}$.

Since $F\phi_m(q)$ comes out equal to

$$\sum_{n=1}^{\infty} \overline{f}_{mn}\xi_n(r)\ ,$$

F has the matrix $\{\overline{f}_{mn}\}$ [by use of the complete orthonormal sets $\phi_m(q)$ and $\overline{\xi}_n(r)$ respectively -- note that the latter is a complete orthonormal set along with $\xi_n(r)$], likewise F^* has the matrix $\{f_{mn}\}$ (with the same complete orthonormal systems). Therefore F^*F, FF^* have the matrices

$$\left\{\sum_{n=1}^{\infty} \overline{f}_{mn} f_{m'n}\right\}$$

(using the complete orthonormal set $\phi_m(q)$ in $\mathfrak{R}^{I}$) and

$$\left\{\sum_{n=1}^{\infty} \overline{f}_{mn} f_{mn'}\right\}$$

(using the complete orthonormal set $\overline{\xi_n(r)}$ in $\mathfrak{R}^{II}$).

On the other hand, $U = P_{[\Phi]}$ has the matrix $\{\overline{f}_{mn}f_{m'n'}\}$ (using the complete orthonormal set $\Phi_{mn}(q,\ r) = \phi_m(q)\xi_n(r)$ in $\mathfrak{R}^{I+II}$), so that its projections in I and II , U and U have the matrices

$$\left\{\sum_{n=1}^{\infty} \overline{f}_{mn} f_{m'n}\right\}$$

and

$$\left\{\sum_{m=1}^{\infty} \overline{f}_{mn} f_{mn'}\right\}$$

respectively (with the complete orthonormal sets given above).[213] Consequently

(**U.**)
$$U = F^*F,\ U = FF^*\ .$$

Note that the definitions (**F.**) and the equations (**U.**) make no use of the ϕ_m, ξ_n -- hence they are valid independently of these.

The operators U , U are completely continuous, and by II.11. and IV.3., they can be written in the form

[213] The mathematical discussion is based on a paper by E. Schmidt, Math. Ann. 83 (1907).

$$U = \sum_{k=1}^{\infty} w'_k P_{[\psi_k]}, \quad U = \sum_{k=1}^{\infty} w''_k P_{[\eta_k]},$$

in which the ψ_k form a complete orthonormal set in $\mathfrak{R}^{I}$, the η_k one in $\mathfrak{R}^{II}$ and all $w'_k, w''_k \geq 0$. We now neglect the terms in each of the two formulas with $w'_k = 0$ or $w''_k = 0$ respectively, and number the remaining terms with $k = 1, 2, \ldots$. Then the ψ_k and η_k again form orthonormal, but not necessarily complete sets; the sums

$$\sum_{k=1}^{M'}, \quad \sum_{k=1}^{M''}$$

appear in place of the two

$$\sum_{k=1}^{\infty}$$

where M', M'' can be equal to ∞ or finite. Also, all w'_k, w''_k are now > 0.

Let us now consider a ψ_k. $U\psi_k = w'_k\psi_k$ and therefore $F^*F\psi_k = w'_k\psi_k$, $FF^*F\psi_k = w'_kF\psi_k$, $UF\psi_k = w'_kF\psi_k$. Furthermore

$$(F\psi_k, F\psi_l) = (F^*F\psi_k, \psi_l) = (U\psi_k, \psi_l)$$

$$= w'_k(\psi_k, \psi_l) \begin{cases} = w'_k, & \text{for } k = l \\ = 0, & \text{for } k \neq l \end{cases},$$

therefore, in particular, $\|F\psi_k\|^2 = w'_k$. The $\frac{1}{\sqrt{w'_k}}F\psi_k$ then form an orthonormal set in R^{II} and they are eigenfunctions of U, with the same eigenvalues as the ψ_k for U (i.e., w'_k). That is, each eigenvalue of U is also one of U with at least the same multiplicity. Interchanging U, U shows that they have the same eigenvalues with the same multiplicities. The w'_k and w''_k therefore coincide except for their order. Hence $M' = M'' = M$, and by re-enumeration of the w''_k we can obtain $w'_k = w''_k = w_k$. And if this occurs, then we can clearly choose

$$\eta_k = \frac{1}{\sqrt{w_k}}F\psi_k$$

in general. Then

$$\frac{1}{\sqrt{w_k}}F^*\eta_k = \frac{1}{w_k}F^*F\psi_k = \frac{1}{w_k}U\psi_k = \psi_k.$$

Therefore

$$\textbf{(V.)} \qquad \eta_k = \frac{1}{\sqrt{w_k}}F\psi_k, \quad \psi_k = \frac{1}{\sqrt{w_k}}F^*\eta_k.$$ [212]

Let us now extend the orthonormal set $\psi_1, \psi_2, \ldots$ to a complete $\psi_1, \psi_2, \ldots, \psi'_1, \psi'_2, \ldots$ and likewise $\eta_1, \eta_2, \ldots$ to $\eta_1, \eta_2, \ldots, \eta'_1, \eta'_2, \ldots$ (each of the two sets $\psi'_1, \psi'_2, \ldots$ and $\eta'_1, \eta'_2, \ldots$ can be empty, finite or infinite, and in addition each set independently of the other set). We have observed before, that (**F.**), (**U.**) make no reference to the ϕ_m, ξ_n. We may therefore use (**V.**), as well as the above construction, and let them determine the choice of the complete orthonormal sets $\phi_1, \phi_2, \ldots$ and $\xi_1, \xi_2, \ldots$. Specifically we let these coincide with the $\psi_1, \psi_2, \ldots, \psi'_1, \psi'_2, \ldots$ and $\overline{\eta}_1, \overline{\eta}_2, \ldots, \overline{\eta}'_1, \overline{\eta}'_2, \ldots$ respectively. Now let ψ_k correspond to ϕ_{μ_k}, η_k to ξ_{ν_k} $(k = 1, \ldots, M)$ ($\mu_1, \mu_2, \ldots$ different from one another, $\nu_1, \nu_2, \ldots$ likewise). Then

$$F\phi_{\mu_k} = \sqrt{w_k}\,\xi_{\nu_k},$$

$$F\phi_m = 0 \text{ for } m \neq \mu_1, \mu_2, \ldots.$$

Therefore

$$f_{mn} \begin{cases} = \sqrt{w_k}\,, & \text{for } m = \mu_k,\ n = \nu_k,\ k = 1, 2, \ldots \\ = 0\,, & \text{otherwise} \end{cases} ,$$

or equivalently

$$\phi(q, r) = \sum_{k=1}^{M} \sqrt{w_k}\, \phi_{\mu_k}(q)\xi_{\nu_k}(r) .$$

By suitable choice of the complete orthonormal sets $\phi_m(q)$ and $\xi_n(r)$ we have thus established that each column of the matrix $\{f_{mn}\}$ contains at most one element $\neq 0$ (that this is real and > 0, namely $\sqrt{w_k}$, is unimportant for what follows). What is the physical meaning of this mathematical statement?

Let A be an operator with the eigenfunctions $\phi_1, \phi_2, \ldots$ and with only distinct eigenvalues, say $a_1, a_2, \ldots$; likewise B with $\xi_1, \xi_2, \ldots$ and $b_1, b_2, \ldots$. A corresponds to a physical quantity in I , B to one in II . They are therefore simultaneously measurable. It is easily seen that the statement "A has the value a_m and B has the value b_n" determines the state $\Phi_{mn}(q, r) = \phi_m(q)\xi_n(r)$, and that this has the probability $(P_{[\Phi_{mn}]}\Phi, \Phi) = |(\Phi, \Phi_{mn})|^2 = |f_{mn}|^2$ in the state $\Phi(q, r)$. Consequently, our statement means that A, B are simultaneously measurable, and that if one of them was measured in Φ , then the value of the other is determined by it uniquely. (An a_m with all $f_{mn} = 0$ cannot result, because its total probability

$$\sum_{n=1}^{\infty} |f_{mn}|^2$$

cannot be 0 , if a_m is ever observed -- therefore for

exactly one n, $f_{mn} \neq 0$; likewise for b_n .) That is, there are several possible A values in the state Φ (namely, those a_m for which

$$\sum_{n=1}^{\infty} |f_{mn}|^2 > 0 ,$$

i.e., for which there exists an n with $f_{mn} \neq 0$ -- usually all a_m are such), and an equal number of possible B values (those b_n for which

$$\sum_{n=1}^{\infty} |f_{mn}|^2 > 0 ,$$

i.e., for which there exists an m with $f_{mn} \neq 0$), but Φ establishes a one-to-one correspondence between the possible A values and the possible B values.

If we call the possible m values $\mu_1, \mu_2, \ldots$ and the corresponding possible n values $\nu_1, \nu_2, \ldots$, then

$$f_{mn} \begin{cases} = c_k \neq 0\,, & \text{for } m = \mu_k,\ n = \nu_k,\ k = 1, 2, \ldots \\ = 0\,, & \text{otherwise} \end{cases} ,$$

therefore (M finite or ∞)

$$\Phi(q, r) = \sum_{k=1}^{M} c_k \phi_{\mu_k}(q)\xi_{\nu_k}(r) ,$$

hence

$$u_{mm'} = \sum_{n=1}^{\infty} \overline{f}_{mn} f_{m'n} \begin{cases} = |c_k|^2\,, & \text{for } m = m' = \mu_k,\ k = 1, 2, \ldots \\ = 0\,, & \text{otherwise} \end{cases} ,$$

$$u_{nn'} = \sum_{m=1}^{\infty} \bar{f}_{mn} f_{mn'} \begin{cases} = |c_k|^2, & \text{for } n = n' = \nu_k,\ k = 1, 2, \ldots \\ = 0, & \text{otherwise} \end{cases}$$

and therefore

$$U = \sum_{k=1}^{M} |c_k|^2 P_{[\phi_{\mu_k}]}, \quad U = \sum_{k=1}^{M} |c_k|^2 P_{[\xi_{\nu_k}]}.$$

Hence, when Φ is projected in I or II, it in general becomes a mixture, while it is a state in I + II only. Indeed, it involves certain information regarding I + II which cannot be made use of in I alone or in II alone, namely the one-to-one correspondence of the A and B values with each other.

For each Φ we can therefore so choose A, B, i.e., the ϕ_m and the ξ_n, that our condition is satisfied; for arbitrary A, B, it may of course be violated. Each state Φ then establishes a particular relation between I and II, while the related quantities A, B depend on Φ. How far Φ determines them, i.e., the ϕ_m and the ξ_n, is not difficult to answer. If all $|c_k|$ are different and $\neq 0$, then U, U (which are determined by Φ) determine the respective ϕ_m, ξ_n uniquely (cf. IV.3.). The general discussion is left to the reader.

Finally, let us mention the fact that for $M \neq 1$ neither U nor U is a state (because all $|c_k|^2 > 0$). For $M = 1$ they both are: $U = P_{[\phi_{\mu_1}]}$, $U = P_{[\xi_{\nu_1}]}$. Then $\Phi(q, r) = c_1 \phi_{\mu_1}(q) \xi_{\nu_1}(r)$. We can absorb c_1 in $\phi_{\mu_1}(q)$. Therefore U, U are states if and only if $\Phi(q, r)$ has the form $\phi(q)\xi(r)$, and in that case they are equal to $P_{[\phi]}$ and $P_{[\xi]}$ respectively.

On the basis of the above results, we note: If I is in the state $\phi(q)$ and II in the state $\xi(r)$, then I + II is in the state $\Phi(q, r) = \phi(q)\xi(r)$. If on the other hand I + II is in a state $\Phi(q, r)$ which is not a product $\phi(q)\xi(r)$, then I and II are mixtures and not states, but Φ establishes a one-to-one correspondence between the possible values of certain quantities in I and in II.

3. DISCUSSION OF THE MEASURING PROCESS

Before we complete the discussion of the measuring process in the sense of the ideas developed in VI.1. (with the aid of the formal tools developed in VI.2.), we shall make use of the results of VI.2. to exclude a possible explanation often proposed for the statistical character of the process **1.** (V.1.). This rests on the following idea: Let I be the observed system, II the observer. If I is in a state $U = P_{[\phi]}$ before the measurement, while II on the other hand is in a mixture

$$U = \sum_{n=1}^{\infty} w_n P_{[\xi_n]},$$

then I + II is a uniquely determined mixture $\mathcal{U}$, and in fact, as we can easily calculate from VI.2.,

$$\mathcal{U} = \sum_{n=1}^{\infty} w_n P_{[\Phi_n]}, \quad \Phi_n(q, r) = \phi(q)\xi_n(r)$$

If now a measurement of a quantity A takes place in I, then this is to be regarded as an interaction of I and II. This is a process **2.** (V.1.), with an energy operator H. If it has the time duration t, then we obtain

$$\mathcal{U}' = e^{-\frac{2\pi i}{h} tH}\, \mathcal{U}\, e^{\frac{2\pi i}{h} tH}$$

from U, and in fact,

$$U' = \sum_{n=1}^{\infty} w_n P\left[e^{-\frac{2\pi i}{h} tH}\phi_n\right]$$

If now each

$$e^{-\frac{2\pi i}{h} tH}\phi_n(q, r)$$

were of the form $\psi_n(q)\eta_n(r)$, where the ψ_n are the eigenfunctions of A, and the η_n any fixed complete orthonormal set, then this intervention would have the character of a measurement. For it transforms each state ϕ of I into a mixture of the eigenfunctions ψ_n of A. The statistical character therefore arises in this way: Before the measurement I was in a (unique) state, but II was a mixture -- and the mixture character of II has, in the course of the interaction, associated itself with I + II, and in particular, it has made a mixture of the projection in I. That is, the result of the measurement is indeterminate, because the state of the observer before the measurement is not known exactly. It is conceivable that such a mechanism might function, because the state of information of the observer regarding his own state could have absolute limitations, by the laws of nature. These limitations would be expressed in the values of the w_n, which are characteristic of the observer alone (and therefore independent of ϕ).

At this point, the attempted explanation breaks down. For quantum mechanics requires that $w_n = (P_{\psi_n}\phi, \phi) = |(\phi, \psi_n)|^2$, i.e., w_n dependent on ϕ! There might exist another decomposition

$$U' = \sum_{n=1}^{\infty} w'_n P_{[\phi'_n]},$$

(the $\phi'_n(q, r) = \psi_n(q)\eta_n(r)$ are orthonormal) but this is of no use either; because the w'_n are (except for order) determined uniquely by U' (IV.3.), and are therefore equal to the w_n.[214]

Therefore, the non-causal nature of the process **1.** is not produced by any incomplete knowledge of the state of the observer, and we shall therefore assume in all that follows that this state is completely known.

Let us now apply ourselves again to the problem formulated at the end of VI.1. I, II, III shall have the meanings given there, and, for the quantum mechanical investigation of I, II, we shall use the notation of VI.2., while III remains outside of the calculations (cf. the discussion of this in VI.1.). Let A be the quantity (in I) actually to be measured, $\phi_1(q), \phi_2(q), \ldots$ its eigenfunctions. Let I be in the state $\phi(q)$.

If I is the observed system, II + III the observer, then we must apply the process **1.**, and we find that the measurement transforms I from the state ϕ into one of the states ϕ_n $(n = 1, 2, \ldots)$, the probabilities for which are respectively $|(\phi, \phi_n)|^2$ $(n = 1, 2, \ldots)$. Now, what is the method of description if I + II is the observed system, and only III the observer?

In this case we must say that II is a measuring instrument which shows on a scale the value of A (in I): the position of the pointer on this scale is a physical quantity B (in II) which is actually observed by III (if II is already within the body of the observer, we have the corresponding physiological concepts in place of the scale and pointer, e.g., retina and image on the retina, etc.) Let A have the values $a_1, a_2, \ldots$, B the values $b_1, b_2, \ldots$, and let the numbering be such that a_n is associated with b_n.

[214] This approach is capable of still more variants, which must be rejected for similar reasons.

Initially, I is in the (unknown) state $\phi(q)$ and II in the (known) state $\xi(r)$, therefore I + II is in the state $\Phi(q, r) = \phi(q)\xi(r)$. The measurement (so far as it is performed by II on I) is, as in the earlier example, carried out by an energy operator H (in I + II) in the time t: This is the process **2.**, which transforms the Φ into

$$\Phi' = e^{-\frac{2\pi i}{h} tH} \Phi .$$

Viewed by the observer III, one has a measurement only if the following is the case: If III were to measure (by process **1.**) the simultaneously measurable quantities A, B (in I or II respectively, or both in I + II), then the pair of values a_n, b_n would have the probability 0 for $m \neq n$, and the probability w_n for $m = n$. That is, it suffices "to look at" II, and A is measured in I. Quantum mechanics then requires in addition $w_n = |(\phi, \phi_n)|^2$.

If this is established, then the measuring process so far as it occurs in II, is "explained" theoretically, i.e., the division of I | II + III discussed in VI.1. is shifted to I + II | III.

The mathematical problem is then the following. A complete orthonormal set $\phi_1, \phi_2, \ldots$ is given in I. Such a set $\xi_1, \xi_2, \ldots$ in $\mathfrak{R}^{II}$ as well as a state ξ in R^I, also an (energy) operator H in $\mathfrak{R}^{I+II}$, and a t, are to be found so that the following holds. If ϕ is an arbitrary state in R^I and

$$\Phi(q, r) = \phi(q)\xi(r), \quad \Phi'(q, r) = e^{-\frac{2\pi i}{h} tH} \Phi(q, r) ,$$

then $\Phi'(q, r)$ must have the form

$$\sum_{n=1}^{\infty} c_n \phi_n(q) \xi_n(r)$$

(the c_n are naturally dependent on ϕ). Therefore $|c_n|^2 = |(\phi, \phi_n)|^2$. (That the latter is equivalent to the physical requirement formulated above was discussed in VI.2.)

In the following we shall use a fixed set $\xi_1, \xi_2, \ldots$ and a fixed ξ along with the fixed $\phi_1, \phi_2, \ldots$, and shall investigate the unitary operator

$$\Delta = e^{-\frac{2\pi i}{h} tH}$$

instead of H.

The mathematical problem leads us back to the problem solved in VI.2.: there the quantity corresponding to our present Φ was given, and we showed the existence of c_n, ϕ_n, ξ_n. Now ϕ_n, ξ_n are fixed and Φ, c_n are given dependent on ϕ, and it remains so to determine a fixed Δ that for $\Phi' = \Delta\Phi$ these c_n, ϕ_n, ξ_n result.

We shall show that such a determination of Δ is indeed possible. In this case only the principle is of importance to us, i.e., the existence of any such Δ. The further question, whether the

$$\Delta = e^{-\frac{2\pi i}{h} tH}$$

corresponding to simple and plausible measuring arrangements also have this property, shall not concern us. Indeed, we saw that our requirements coincide with a plausible intuitive criterion of the measurement character in an intervention. Furthermore the arrangements in question are to possess the characteristics of the measurement. Hence quantum mechanics, as applied to observation would be in blatant contradiction with experience, if these Δ did not satisfy the requirements in question (at least approximately).[215] Therefore, in the following, only an abstract

[215] The corresponding calculation for the case of the posi-

Δ which satisfies our conditions exactly, shall be given.

Therefore, let the ϕ_m $(m = 0, \pm 1, \pm 2, \ldots)$ and the ξ_n $(n = 0, \pm 1, \pm 2, \ldots)$ respectively be two given complete orthonormal sets in $\mathfrak{R}^{I}$ and $\mathfrak{R}^{II}$ respectively. (We do not let m, n run over $1, 2, \ldots$, but over $0, \pm 1, \pm 2, \ldots$. This is purely for technical convenience, and is in principle equivalent to the former). Let the state ξ be, for simplicity, ξ_0. We define the operator Δ by

$$\Delta \sum_{m,n=-\infty}^{\infty} x_{mn}\phi_m(q)\xi_n(r) = \sum_{m,n=-\infty}^{\infty} x_{mn}\phi_m(q)\xi_{m+n}(r) ,$$

since the $\phi_m(q)\xi_n(r)$ as well as the $\phi_m(q)\xi_{m+n}(r)$ form a complete orthonormal set in $\mathfrak{R}^{I+II}$, this Δ is unitary. Now

$$\phi(q) = \sum_{m=-\infty}^{\infty} (\phi, \phi_m)\cdot\phi_m(q), \quad \xi(r) = \xi_0(r) ,$$

therefore

$$\Phi(q, r) = \phi(q)\xi(r) = \sum_{m=-\infty}^{\infty} (\phi, \phi_m)\cdot\phi_m(q)\xi_0(r) ,$$

$$\Phi'(q, r) = \Delta\Phi(q, r) = \sum_{m=-\infty}^{\infty} (\phi, \phi_m)\cdot\phi_m(q)\xi_m(r) .$$

Hence our purpose is accomplished. We have in addition $c_n = (\phi, \phi_n)$.

A better overall view of the mechanism of this process can be obtained if we exemplify it by concrete Schrödinger wave functions, and give H in place of Δ.

The observed object, as well as the observer

tion measurement discussed in III.4. is contained in a paper by Weizsäcker, Z. Physik 70 (1931).

(i.e., I and II respectively) may be characterized by a single variable q and r respectively, running continuously from $-\infty$ to $+\infty$. That is, let both be thought of as points which can move along a line. Their wave functions then have always the form $\psi(q)$ and $\eta(r)$ respectively. We assume that their masses m_1 and m_2 are so large that the kinetic energy portion of the energy operator (i.e., $\frac{1}{2m_1}\left(\frac{h}{2\pi i}\frac{\partial}{\partial q}\right)^2 + \frac{1}{2m_2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial r}\right)^2$) can be neglected. Then there remains of H only the interaction energy part which is decisive for the measurement. For this we choose the particular form $\frac{h}{2\pi i} q \frac{\partial}{\partial r}$.

The Schrödinger time dependent differential equation then is (for the I + II wave functions $\psi_t = \psi_t(q, r)$):

$$\frac{h}{2\pi i}\frac{\partial}{\partial t}\psi_t(q, r) = -\frac{h}{2\pi i} q \frac{\partial}{\partial r}\psi_t(q, r) ,$$

$$\left(\frac{\partial}{\partial t} + q\frac{\partial}{\partial r}\right)\psi_t(q, r) = 0 ,$$

i.e.,

$$\psi_t(q, r) = f(q, r - tq) .$$

If, for $t = 0$, $\psi_0(q, r) = \Phi(q, r)$, then we have $f(q, r) = \Phi(q, r)$, and therefore

$$\psi_t(q, r) = \Phi(q, r - tq)$$

In particular, if the initial states of I, II are represented by $\phi(q)$ and $\xi(r)$ respectively, then, in the sense of our calculation scheme (if the time t appearing therein is chosen to be 1)

$$\Phi(q, r) = \phi(q)\xi(r) ,$$

$$\Phi'(q, r) = \psi_1(q, r) = \phi(q)\xi(r - q) .$$

We now wish to show that this can be used by II for a position measurement of I , i.e., that the coordinates are tied to each other. (Since q, r have continuous spectra, they are therefore measurable with only arbitrary precision, but not with absolute precision. Hence this can be accomplished only approximately.)

For this purpose, we wish to assume that $\xi(r)$ is different from 0 only in a very small interval $-\epsilon < r < \epsilon$ (i.e., the coordinate r of the observer before the measurement is very accurately known), in addition ξ should of course be normalized:

$$||\xi|| = 1 \text{ , i.e., } \int |\xi(r)|^2 dr = 1 \text{ .}$$

The probability therefore that q lies in the interval $q_0 - \delta < q < q_0 + \delta$, and r in the interval $r_0 - \delta' < r < r_0 + \delta'$ is

$$\int_{q_0-\delta}^{q_0+\delta}\int_{r_0-\delta}^{r_0+\delta} |\Phi'(q, r)|^2 dq dr = \int_{q_0-\delta}^{q_0+\delta}\int_{r_0-\delta}^{r_0+\delta} |\phi(q)|^2 |\xi(r - q)|^2 dq dr \text{ .}$$

If q_0, r_0 are to differ by more than $\delta + \delta' + \epsilon$, then this is 0 , i.e., q, r are so very closely tied to each other that the difference can never be greater than $\delta + \delta' + \epsilon$. And for $r_0 = q_0$ this is, equal to

$$\int_{q_0-\delta}^{q_0+\delta} |\phi(q)|^2 dq \text{ ,}$$

if we choose $\delta' \geq \delta + \epsilon$, because of the assumptions on ξ . But since we can choose δ, δ', ϵ arbitrarily small (they must be different from zero, however), this means that q, r are tied to each other with arbitrary closeness, and the probability density has the value furnished

by quantum mechanics, $|\phi(q)|^2$.

That is, the relations of the measurement, as we had discussed them in IV.1., and in this section, are realized.

The discussion of more complicated examples, say of an analog to our four-term example of IV.1., or the control determination of the validity of a measurement which II carried out on I , effected by a second observer III , can also be carried out in this fashion. It is left to the reader.