This last part of the notes presents the views on measurement by two other important sources

-

Cohen and Jauch

The reduced probability matrix

We consider the possibility of having a system that has interacted with its surrounding.

So we have “system ⊗ environment” or “system ⊗ measurement device” or simply a system which is a part of a larger thing which we can call “universe”.

The question that we would like to ask is as follows: Assuming that we know what is the state of the “universe”, what is the way to calculate the state of the “system”?

The mathematical formulation of the problem is as follows.

The pure states of the ”system” span N_{sys} dimensional Hilbert space, while the states of the ” environment” span N_{env} dimensional Hilbert space.

So the state of the ”universe” is described by $N \times N$ probability matrix $\rho_{i\alpha,j\beta}$, where $N = N_{\text{sys}}N_{\text{env}}$.

This means that if we have operator A which is represented by the matrix $A_{i\alpha,j\beta}$, then it expectation value is

$$\langle A \rangle \quad = \quad \text{trace}(A\rho) \quad = \quad \sum_{i,j,\alpha,\beta} A_{i\alpha,j\beta} \; \rho_{j\beta,i\alpha}$$

The probability matrix of the ”system” is defined in the usual way.

Namely, the matrix element $\rho^{sys}_{j,i}$ is defined as the expectation value of $P^{ji} = |i\rangle\langle j| \otimes \mathbf{1}$.

Hence

$$\rho^{\text{sys}}_{j,i} = \langle P^{ji} \rangle = \text{trace}(P^{ji}\rho) = \sum_{k,\alpha,l,\beta} P^{ji}_{k\alpha,l\beta}\rho_{l\beta,k\alpha} = \sum_{k,\alpha,l,\beta} \delta_{k,i}\delta_{l,j}\delta_{\alpha,\beta} \; \rho_{l\beta,k\alpha} = \sum_{\alpha} \rho_{j\alpha,i\alpha}$$

The common terminology is to say that ρ^{sys} is the reduced probability matrix, which is obtained by tracing out the environmental degrees of freedom.

Just to show mathematical consistency we note that for a general system operator of the type $A = A^{sys} \otimes \mathbf{1}^{env}$ we get as expected

$$\langle A \rangle = \text{trace}(A\rho) = \sum_{i,\alpha,j\beta} A_{i\alpha,j\beta}\rho_{j\beta,i\alpha} = \sum_{i,j} A^{sys}_{i,j}\rho^{sys}_{j,i} = \text{trace}(A^{sys}\rho^{sys})$$

Entangled superposition

Of particular interest is the case where the universe is in a pure state Ψ.

Choosing for the system $\otimes$ environment an arbitrary basis $|i\alpha\rangle = |i\rangle \otimes |\alpha\rangle$, we can expand the wavefunction as

$$|\Psi\rangle = \sum_{i,\alpha} \Psi_{i\alpha}|i\alpha\rangle$$

By summing over α we can write

$$|\Psi\rangle = \sum_{i} \sqrt{p_i}|i\rangle \otimes |\chi^{(i)}\rangle$$

where $\left|\chi^{(i)}\right\rangle \propto \sum_{\alpha} \Psi_{i\alpha}|\alpha\rangle$ is called the ”relative state” of the environment with respect to the i^{th} state of the system, while p_i is the associated normalization factor.

Note that the definition of the relative state implies that $\Psi_{i\alpha} = \sqrt{p_i}\chi^{(i)}_{\alpha}$.

Using these notations it follows that the reduced probability matrix of the system is

$$\rho^{sys}_{j,i} = \sum_{\alpha} \Psi_{j\alpha}\Psi^*_{i\alpha} = \sqrt{p_i p_j}\,\langle\chi^{(i)}|\chi^{(j)}\rangle$$

The prototype example for a system-environment entangled state is described by the superposition

$$|\Psi\rangle \quad = \quad \sqrt{p_1}\;|1\rangle \otimes |\chi^{(1)}\rangle + \sqrt{p_2}\;|2\rangle \otimes |\chi^{(2)}\rangle$$

where $|1\rangle$ and $|2\rangle$ are orthonormal states of the system.

The singlet state of two spin 1/2 particles is possibly the simplest example for an entangled superposition of this type.

Later on we shall see that such entangled superposition may come out as a result of an interaction between the system and the environment.

Namely, depending on the state of the system the environment, or the measurement apparatus, ends up in a different state χ.

Accordingly we do not assume that $\chi^{(1)}$ and $\chi^{(2)}$ are orthogonal, though we normalize each of them and pull out the normalization factors as p_1 and p_2.

The reduced probability matrix of the system is

$$\rho^{\rm sys}_{j,i} \quad = \quad \begin{pmatrix} p_1 & \lambda^* \sqrt{p_1 p_2} \\ \lambda\sqrt{p_1 p_2} & p_2 \end{pmatrix}$$

where $\lambda = \langle\chi^{(1)}|\chi^{(2)}\rangle$.

At the same time the environment is in a mixture of non-orthogonal states:

$$\rho^{\rm env} \quad = \quad p_1\;|\chi^{(1)}\rangle\langle\chi^{(1)}| + p_2\;|\chi^{(2)}\rangle\langle\chi^{(2)}|$$

The purity of the state of the system in the above example is determined by $|\lambda|$, and can be characterized by trace(ρ^2) = $1 - 2p_1 p_2 (1-|\lambda|^2)$.

The value trace(ρ^2) = 1 indicates a pure state, while trace(ρ^2) = 1/N with N = 2 characterizes a 50%-50% mixture.

Schmidt decomposition

If the "universe" is in a pure state we cannot write its ρ as a mixture of product states, but we can write its Ψ as an entangled superposition of product states.

$$|\Psi\rangle \quad = \quad \sum_i \sqrt{p_i}|i\rangle \otimes |B_i\rangle$$

where the $|B_i\rangle$ is the "relative state" of subsystem *B* with respect to the ith state of subsystem *A*, while p_i is the associated normalization factor.

The states $|B_i\rangle$ are in general not orthogonal.

The natural question that arise is whether we can find a decomposition such that the $|B_i\rangle$ are orthonormal.

The answer is positive: Such decomposition exists and it is unique.

It is called Schmidt decomposition, and it is based on singular value decomposition (SVD).

Let us regard $\Psi_{i\alpha} = W_{i,\alpha}$ as an $N_A \times N_B$ matrix.

From linear algebra it is known that any matrix can be written in a unique way as a product:

$$W_{(N_A \times N_B)} = U^A_{(N_A \times N_A)} D_{(N_A \times N_B)} U^B_{(N_B \times N_B)}$$

where U^A and U^B are the so called left and right unitary matrices, while D is a diagonal matrix with so called (positive) singular values.

Thus we can re-write the above matrix multiplication as

$$\Psi_{i\alpha} = \sum_r U^A_{i,r} \sqrt{p_r} \; U^B_{r,\alpha}$$

Substitution of this expression leads to the result

$$|\Psi\rangle = \sum_r \sqrt{p_r} |A_r\rangle \otimes |B_r\rangle$$

where $|A_r\rangle$ and $|B_r\rangle$ are implied by the unitary transformations.

We note that the normalization of Ψ implies $\sum p_r = 1$.

Furthermore the probability matrix is $\rho^{A+B} = W_{i,\alpha} W^*_{j,\beta}$, and therefore the calculation of the reduced probability matrix can be written as:

$$\begin{aligned} \rho^A &= \quad WW^\dagger \quad &&= (U^A) D^2 (U^A)^\dagger \\ \rho^B &= (W^T)(W^T)^\dagger &&= [(U^B)^\dagger D^2 (U^B)]^* \end{aligned}$$

This means that the matrices ρ^A and ρ^B have the same non-zero eigenvalues $\{p_r\}$, or in other words it means that the degree of purity of the two subsystems is the same.

Quantum entanglement

Let us consider a system consisting of two sub-systems, ”A” and ”B”, with no correlation between them. Then, the state of the system can be factorized:

$$\rho^{A+B} = \rho^A \rho^B$$

But in reality the state of the two sub-systems can be correlated.

In classical statistical mechanics ρ^A and ρ^B are probability functions, while ρ^{A+B} is the joint probability function. In the classical state we can always write

$$\rho^{A+B}(x, y) = \sum_{x', y'} \rho^{A+B}(x', y')\delta_{x,x'}\delta_{y,y'}$$

where x and y labels classical definite states of subsystems *A* and *B* respectively.

This means schematically that we can write

$$\rho^{A+B} = \sum_r p_r \rho^{(A_r)} \rho^{(B_r)}$$

where $r = (x', y')$ is an index that distinguish pure classical states of $A \otimes B$, and $p_r = \rho(x', y')$ are probabilities such that $p_r = 1$, and $\varrho^{(Ar)} \mapsto \delta_{x, x'}$ is a pure classical state of subsystem *A*, and $\varrho^{(Br)} \mapsto \delta_{y,y'}$ is a pure classical state of subsystem *B*.

Thus any classical state of $A \otimes B$ can be expressed as a *mixture* of product states.

By definition a quantum state is not *entangled* if it is a product state or a mixture of product states.

Using explicit matrix representation it means that it is possible to write

$$\rho^{A+B}_{i\alpha,j\beta} = \sum_r p_r \; \rho^{(A_r)}_{i,j} \; \rho^{(B_r)}_{\alpha,\beta}$$

It follows that an entangled state, unlike a non-entangled state cannot have a classical interpretation.

Measurements, the notion of collapse

In elementary textbooks the quantum measurement process is described as inducing "collapse" of the wavefunction.

Assume that the system is prepared in state $\rho_{\text{initial}} = |\psi\rangle\langle\psi|$ and that one measures $\hat{P} = |\phi\rangle\langle\phi|$. \If the result of the measurement is $\hat{P} = 1$ then it is said that the system has collapsed into the state $\rho_{\text{final}} = |\phi\rangle\langle\phi|$.

The probability for this "collapse" is given by the projection formula $\text{Prob}(\phi|\psi) = |\langle\phi|\psi\rangle|^2$.

If one regards $\rho(x,x')$ or $\psi(x)$ as representing physical reality, rather than a probability matrix or a probability amplitude, then one immediately gets into puzzles.

Recalling the EPR experiment this world imply that once the state of one spin is measured at Earth, then immediately the state of the other spin (at the Moon) would change from unpolarized to polarized.

This would suggest that some spooky type of "interaction" over distance has occurred.

In fact we shall see that the quantum theory of measurement does not involve any assumption of spooky "collapse" mechanism.

Once we recall that the notion of quantum state has a statistical interpretation the mystery fades away.

In fact we explain (see below) that there is "collapse" also in classical physics!

To avoid potential miss-understanding it should be clear that I do not claim that the classical “collapse” which is described below is an explanation of the the quantum collapse.

The explanation of quantum collapse using a quantum measurement (probabilistic) point of view will be presented in a later section.

The only claim of this section is that in probability theory a correlation is frequently mistaken to be a causal relation: “smokers are less likely to have Alzheimer” not because cigarettes help to their health, but simply because their life span is smaller.

Similarly quantum collapse is frequently mistaken to be a spooky interaction between well separated systems.

Consider the thought experiment which is known as the “Monty Hall Paradox”.

There is a car behind one of three doors.

The car is like a classical ”particle”, and each door is like a ”site”.

The initial classical state is such that the car has equal probability to be behind any of the three doors.

You are asked to make a guess.

Let us say that you pick door #1.

Now the organizer opens door #2 and you see that there is no car behind it.

This is like a measurement.

Now the organizer allows you to change your mind.

The naive reasoning is that now the car has equal probability to be behind either of the two remaining doors.

So you may claim that it does not matter.

But it turns out that this simple answer is very very wrong!

The car is no longer in a state of equal probabilities: Now the probability to find it behind door #3 has increased.

A standard calculation reveals that the probability to find it behind door #3 is twice large compared with the probability to find it behind door #2.

So we have here an example for a classical collapse.

If you are not familiar with this well known ”paradox”, the following may help to understand why we have this collapse.

Imagine that there are billion doors.

You pick door #1.

The organizer opens all the other doors except door #234123.

So now you know that the car is either behind door #1 or behind door #234123. You want the car. What are you going to do?

It is quite obvious that the car is almost definitely behind door #234123.

It is also clear the that the collapse of the car into site #234123 does not imply any physical change in the position of the car.

Quantum measurements, Schroedinger's cat

What do we mean by quantum measurement?

In order to clarify this notion let us consider a system and a detector which are prepared independently as

$$|\Psi\rangle \ = \ \left[\sum_a \psi_a |a\rangle\right] \otimes |q=0\rangle$$

As a result of an interaction we assume that the detector correlates with the system as follows:

$$\hat{U}_{\text{measurement}}\Psi \ = \ \sum \psi_a |a\rangle \otimes |q=a\rangle$$

We call such type of unitary evolution *ideal projective measurement*.

If the system is in a definite a state, then it is not affected by the detector.

Rather, we gain information on the state of the system.

One can think of q as representing a memory device in which the information is stored.

This memory device can be of course the brain of a human observer.

Form the point of view of the observer, the result at the end of the measurement process is to have a definite a.

This is interpreted as a "collapse" of the state of the system.

Some people wrongly think that "collapse" is something that goes beyond unitary evolution.

But in fact this term just makes over dramatization of the above unitary process.

The concept of measurement in quantum mechanics involves psychological difficulties which are best illustrated by considering the "Schrodinger's cat" experiment.

This thought experiment involves a radioactive nucleus, a cat, and a human being.

The half life time of the nucleus is an hour.

If the radioactive nucleus decays it triggers a poison which kills the cat.

The radioactive nucleus and the cat are inside an isolated box.

At some stage the human observer may open the box to see what happens with the cat...

Let us translate the story into a mathematical language.

At time t = 0 the state of the universe (nucleus⊗cat⊗observer) is

$$\Psi \;=\; |\uparrow = \text{radioactive}\rangle \otimes |q = 1 = \text{alive}\rangle \otimes |Q = 0 = \text{ignorant}\rangle$$

where q is the state of the cat, and Q is the state of the memory bit inside the human observer.

If we wait a very long time the nucleus would definitely decay, and as a result we will have a definitely dead cat:

$$U_{\text{waiting}}\Psi \;=\; |\downarrow = \text{decayed}\rangle \otimes |q = -1 = \text{dead}\rangle \otimes |Q = 0 = \text{ignorant}\rangle$$

If the observer opens the box he/she would see a dead cat:

$$U_{\text{seeing}}U_{\text{waiting}}\Psi \;=\; |\uparrow = \text{decayed}\rangle \otimes |q = -1 = \text{dead}\rangle \otimes |Q = -1 = \text{shocked}\rangle$$

But if we wait only one hour then

$$U_{\text{waiting}}\Psi \;=\; \frac{1}{\sqrt{2}}\Big[|\uparrow\rangle \otimes |q = +1\rangle + |\downarrow\rangle \otimes |q = -1\rangle\Big] \otimes |Q = 0 = \text{ignorant}\rangle$$

which means that from the point of view of the observer the system (nucleus+cat) is in a superposition.

The cat at this stage is neither definitely alive nor definitely dead.

But now the observer open the box and we have:

$$U_{\text{seeing}} U_{\text{waiting}} \Psi \; = \; \frac{1}{\sqrt{2}} \Big[|\uparrow\rangle \otimes |q=+1\rangle \otimes |Q=+1\rangle \; + \; |\downarrow\rangle \otimes |q=-1\rangle \otimes |Q=-1\rangle \Big]$$

We see that now, form the point of view of the observer, the cat is in a definite(!) state.

This is regarded by the observer as "collapse" of the superposition.

We have of course two possibilities: one possibility is that the observer sees a definitely dead cat, while the other possibility is that the observer sees a definitely alive cat.

The two possibilities "exist" in parallel, which leads to the "many worlds" interpretation.

Equivalently one may say that only one of the two possible scenarios is realized from the point of view of the observer, which leads to the "relative state" concept of Everett.

Whatever terminology we use, "collapse" or "many worlds" or "relative state", the bottom line is that we have here merely a unitary evolution.

Measurements, formal treatment

In this section we describe mathematically how an ideal projective measurement affects the state of the system.

First of all let us write how the U of a measurement process looks like.

The formal expression is

$$\hat{U}_{\text{measurement}} \; = \; \sum_a \hat{P}^{(a)} \otimes \hat{D}^{(a)}$$

where $\hat{P}^{(a)}$ = |a⟩⟨a| is the projection operator on the state |a⟩, and $\hat{D}^{(a)}$ is a translation operator.

Assuming that the measurement device is prepared in a state of ignorance $|q = 0\rangle$, the effect of $\hat{D}^{(a)}$ is to get $|q = a\rangle$.

Hence

$$\begin{aligned} \hat{U}\Psi &= \left[\sum_a \hat{P}^{(a)} \otimes \hat{D}^{(a)}\right] \left(\sum_{a'} \psi_{a'} |a'\rangle \otimes |q=0\rangle\right) \\ &= \sum_a \psi_a |a\rangle \otimes \hat{D}^{(a)} |q=0\rangle \;=\; \sum_a \psi_a |a\rangle \otimes |q=a\rangle \end{aligned}$$

A more appropriate way to describe the state of the system is using the probability matrix.

Let us describe the above measurement process using this language.

After "reset" the state of the measurement apparatus is $\sigma^{(0)} = |q{=}0\rangle\langle q{=}0|$.

The system is initially in an arbitrary state ρ.

The measurement process correlates that state of the measurement apparatus with the state of the system as follows:

$$\hat{U}\rho \otimes \sigma^{(0)}\hat{U}^\dagger \;=\; \sum_{a,b} \hat{P}^{(a)}\rho\hat{P}^{(b)} \otimes [\hat{D}^{(a)}]\sigma^{(0)}[\hat{D}^{(b)}]^\dagger \;=\; \sum_{a,b} \hat{P}^{(a)}\rho\hat{P}^{(b)} \otimes |q{=}a\rangle\langle q{=}b|$$

Tracing out the measurement apparatus we get

$$\rho^{\text{final}} \;=\; \sum_a \hat{P}^{(a)}\rho\hat{P}^{(a)} \;\equiv\; \sum_a p_a \rho^{(a)}$$

Where p_a is the trace of the projected probability matrix $\hat{P}^{(a)}\rho\hat{P}^{(a)}$, while $\rho^{(a)}$ is its normalized version.

We see that the effect of the measurement is to turn the superposition into a mixture of a states, unlike unitary evolution for which

$$\rho^{\text{final}} = U\,\rho\,U^\dagger$$

So indeed a measurement process looks like a non-unitary process: it turns a pure superposition into a mixture.

A simple example is in order.

Let us assume that the system is a spin 1/2 particle.

The spin is prepared in a pure polarization state $\rho = |\psi\rangle\langle\psi|$ which is represented by the matrix

$$\rho_{ab} = \psi_a\psi_b^* = \begin{pmatrix} |\psi_1|^2 & \psi_1\psi_2^* \\ \psi_2\psi_1^* & |\psi_2|^2 \end{pmatrix}$$

where 1 and 2 are (say) the "up" and "down" states.

Using a Stern-Gerlach apparatus we can measure the polarization of the spin in the up/down direction.

This means that the measurement apparatus projects the state of the spin using

$$P^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad P^{(2)} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

leading after the measurement to the state

$$\rho^{\text{final}} = P^{(1)}\rho P^{(1)} + P^{(2)}\rho P^{(2)} = \begin{pmatrix} |\psi_1|^2 & 0 \\ 0 & |\psi_2|^2 \end{pmatrix}$$

Thus the measurement process has eliminated the off-diagonal terms in ρ and hence turned a pure state into a mixture.

It is important to remember that this non-unitary non-coherent evolution arises because we look only on the state of the system.

On a universal scale the evolution is in fact unitary.

THE MEASURING PROCESS - Jauch

That things have a quality in themselves quite apart from interpretation and subjectivity, is an idle hypothesis: It would presuppose that to interpret and to be a subject are not essential, that a thing detached from all relations is still a thing.

F. NIETZSCHE, in *The Will to Power*

We now present an analysis of the process of measurement.
The characteristic limitations of the accuracy of measurements in the form of the uncertainty relations is traced to the interaction of the measuring device and the system.

In order to analyze this effect we describe the characteristic properties of a measuring device.

The following discussion introduces the notation of equivalent states, events, and data.

In a mathematical interlude we sketch the theory of the tensor product, followed by a discussion on the union and separation of a system.

All the tools are then on hand to analyze in detail the measuring process on a particular and simple model.

Finally, we describe three paradoxes of the measuring process.

UNCERTAINTY RELATIONS

The deeper understanding of quantum phenomena began with the discovery of the uncertainty relations.

From the earliest beginnings in the discussion of the uncertainty relations, it was recognized that the analysis of these relations, which restrict the precision of measurement, must be somehow related to the inevitable interaction between measuring device and the system during the process of measurement.

In our formulation of quantum mechanics, the actual behavior of a physical system under the process of measurement is already incorporated in the empirically given structure of the proposition system.

It is therefore not surprising that the uncertainty relations appear, at this stage of our presentation, as a purely mathematical consequence of the properties of observables.

We now demonstrate this.

Let A and B stand for two observables, which need not be compatible.

They are represented by self-adjoint operators which need not commute.

In order to avoid inessential complications, we shall assume that the two operators are bounded so that their domain of definition is the entire Hilbert space.

Let us consider a general state given by the density operator W, and define the expectation values

$$a \equiv \mathrm{Tr}\, WA, \qquad b \equiv \mathrm{Tr}\, WB. \tag{1}$$

The mean square deviation of the observable A in the state W is then defined by

$$(\Delta a)^2 \equiv \text{Tr } W(A - a)^2. \tag{2}$$

Similarly, we define

$$(\Delta b)^2 \equiv \text{Tr } W(B - b)^2. \tag{3}$$

The uncertainty relation is an inequality for the product of Δa with Δb.

We write $A_0 = A - a$ and $B_0 = B - b$.

Then we have, by definition,

$$\text{Tr } WA_0 = \text{Tr } WB_0 = 0.$$

In order to obtain such an inequality we consider the operator $T \equiv A_0 + i\rho B_0$, with ρ real.

Since A and B are self-adjoint and bounded, $T^* = A_0 - i\rho B_0$.

It follows from this that $\text{Tr } W\, T\, T^* = \text{Tr } T^*\, W\, T > 0$ for all values of ρ.

In order to see this, it suffices to recall that the trace of any operator X, if it exists, can be calculated with the formula

$$\text{Tr } X = \sum_r (\varphi_r, X\varphi_r)$$

for some arbitrary, complete, orthonormal system of vectors φ.

Thus

$$\text{Tr } T^*WT = \sum_r (\varphi_r, T^*WT\varphi_r) = \sum_r (T\varphi_r, WT\varphi_r).$$

Let $\psi_r = T\varphi_r$; then we have finally

$$\mathrm{Tr}\ T^* W T = \sum_r (\psi_r, W\psi_r). \tag{4}$$

The system ψ_r , is in general neither complete nor orthogonal, but this does not matter.

What does matter is that every term $(\psi_r, W\psi_r)$ is positive.

Therefore the trace in question is expressed as a sum of positive terms.

Thus we have found

$$\mathrm{Tr}\ WTT^* \equiv \mathrm{Tr}\ WA_0^2 - i\rho\ \mathrm{Tr}\ W[A_0, B_0] + \rho^2\ \mathrm{Tr}\ WB_0^2 \geq 0. \tag{5}$$

The right-hand side is a quadratic form in ρ which is thus positive definite.

This implies that its determinant must be positive, too.

Since A_0 and B_0 are both self-adjoint, the operator $[A_0, B_0] = [A, B]$ which occurs in Eq. (5) is antiself-adjoint.

We can then define a new self-adjoint operator C by setting $[A, B] = iC$.

We obtain the real quadratic form

$$(\Delta a)^2 + \rho\ \mathrm{Tr}\ WC + \rho^2(\Delta b)^2 \geq 0.$$

The positive definiteness of this quadratic expression in ρ implies that

$$(\Delta a)^2(\Delta b)^2 \geq \tfrac{1}{4}(\operatorname{Tr} WC)^2, \tag{6}$$

or, after taking square roots,

$$(\Delta a)(\Delta b) \geq \tfrac{1}{2}|\operatorname{Tr} WC|. \tag{7}$$

This is the general uncertainty relation.

It gives a limitation on the value of the product of the root-mean-square deviation for any two observables.

In the event that A and B do not commute, then $C \neq 0$, and there exist, in general, states such that the right-hand side is positive (and not zero).

In case C is a positive operator (which will be seen to occur in the most important application), then this will occur for any state.

In this case we have what we might call an *absolute* uncertainty relation: The two observables in question can *never* be measured with an arbitrary simultaneous accuracy.

The difficulty in the interpretation of the uncertainty relation stems from the fact that A and B might be observables which admit an arbitrarily small mean square deviation for certain states.

The uncertainty relation implies, then, that this can never occur for the same state.

For such cases the observable A separately, as well as the observable B separately, can be measured with unrestricted accuracy.

However, if we desire to measure A and B simultaneously, we find that the two measurements must be disturbing each other in precisely such a way that the uncertainty relation is satisfied.

This behavior, which for classical systems is completely unknown, gives rise to the often discussed conjecture whether the actual values of the observables are perhaps determined, but unknown and unknowable because the measuring process interferes with the state of the system.

This point of view is closely related to the possible existence of hidden variables.

But in this case the hidden aspect of the state would even have a subjective character, depending on the state of information which we have about the system.

That this would lead to quite inadmissible consequences can be seen, for instance, by the fact that the state would be different for two different observers who have different information about one and the same system.

This would amount to a denial of objectively valid physical laws throughout the world of microsystems, something for which we have at present no evidence, and no need.

We shall therefore take the position that the correct interpretation of the uncertainty relation is that an observable in a given state need not have a definite value even though the observable is capable of assuming definite values for other states.

The value of an observable may be not only ambiguous, it may even be undetermined, and this is as much an irreducible attribute of a state as specific values were in classical physics.

The situation has often been described by saying that in quantum mechanics it is not possible to ignore the reaction of the measuring device on the system.

This interaction, so we are told, produces an uncontrollable effect on the system, so that the actual values of the observables remain partially hidden from us.

Although this discussion of the interaction of the measuring device and system has been most useful in attracting our attention to the fact that we are dealing with physical objects and not with a mathematical model, the introduction of an uncontrollable reaction of one on the other has given a misleading impression.

It is not an uncontrollable effect which causes the uncertainty relation; on the contrary, the time evolution of the joint system, consisting of measuring device and original system, may be completely controlled and determined by a fixed physical arrangement.

It is the laws of quantum mechanics as they are embodied in the structure properties of the proposition system which cause this result.

A complete understanding of these laws can only be obtained by following them through their effects on the measuring process, and this is precisely what we shall do now.

GENERAL DESCRIPTION OF THE MEASURING PROCESS

All the information which we have about physical systems is obtained from observations and measurements on such systems.

Observations consist in bringing the system under examination in contact with some other system, the observer, or some measuring device *M*, and observing the reaction of the system on the observer.

The general setup of a measurement on a physical system *S* has thus schematically the form indicated in Fig. 1.

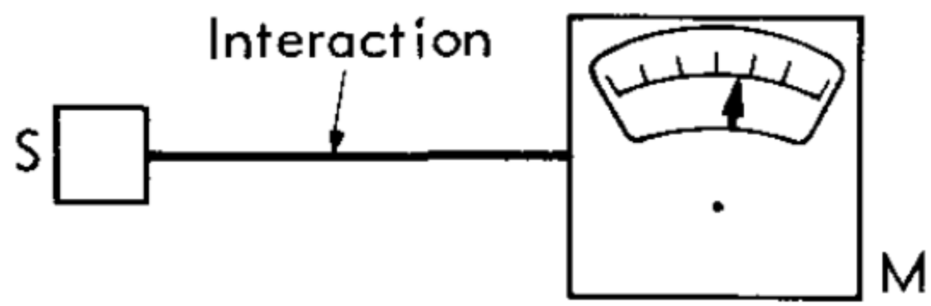

Fig. 1 Schematic representation of the interaction between measuring device *M* and system *S*.

Let us point out at the outset two important features which we shall encounter throughout the discussion on the quantum-mechanical measuring process.

The first of these concerns the fact that the measuring device *M*, if it is to be of any use at all, must interact somehow with the system *S*.

But an interaction always acts both ways.

Not only does *S* influence *M*, thereby producing the desired measurable effect, but also *M* acts on *S*, producing an effect on *S* with no particularly desirable consequences.

In fact, this back-effect on *S* seems to be the cause of much of the difficulty in the interpretation of quantum mechanics.

The second point to be mentioned here is that, in a complete measurement, the schematic arrangement of Fig. 1 is an oversimplification.

If the measurement is to be useful there must be a further observation on *M*, which we shall call "reading the scale."

Such further observations may be made at a later time by examining a permanent record of some sort, but in any case, if they are to be of some use for the construction of a scientific theory they must, at some point, enter the consciousness of a scientific observer.

This appearance of the "conscious observer" in a full description of the measuring process is a disturbing element, since it seems to introduce into the process of reconstructing the objective physical laws a foreign subjective element.

It is essential in the construction of an objective science that it be freed from anthropomorphic elements. This requirement of objectivity of the physical science of micro-objects can actually be satisfied because of the fact that the last stages of observation which we have designated symbolically as "reading the scale" take place on the classical level.

On this level the measuring process also follows the general scheme of Fig. 1, with one important difference.

The interconnection between *S* and *M* acts only in one direction, namely from *S* to *M*.

The reverse action from *M* to *S*, although always present, is not important because its effect can always be reduced to a negligible amount.

This is an essential feature of the classical measuring process which distinguishes it fundamentally from the quantum mechanical measuring process, where this reaction is not negligible.

The insensitivity of the system with respect to its observation when " reading the scale" has the consequence that all the data furnished by a macroscopic measuring device have an objective meaning.

By this we mean that the "scale can be read" by a number of different observers who can communicate and establish that they read concurrent results.

The physical fact is " objectivized," as we might say.

The individual observer, although necessary for completing an actual observation, can now fade into the background; and we retain only the objectively verifiable content of the observation.

These are the building bricks of man's theory of the physical world.

The situation which we encounter here is very similar to the description of a physical event in space and time with respect to some coordinate system.

Although a coordinate description of such events is usually necessary, their objective character and their spatio-temporal interrelations are objective properties quite independent of the particular choice of such a description.

DESCRIPTION OF THE MEASURING PROCESS FOR QUANTUM-MECHANICAL SYSTEMS

As we have pointed out in the preceding discussion, the quantum-mechanical measuring process will affect the system.

We shall now examine the effect of this reaction more carefully.

Let us begin with two examples.

First we consider the measurement of the position of some elementary particle by a counter with a finite sensitive volume.

After the measurement has been performed and the counter has recorded the presence of a particle inside its sensitive volume, we know for certain that the particle, at the instant of the triggering, is actually inside the sensitive volume.

By this we mean the following: Suppose we repeated the measurement immediately after it has occurred (this is of course an idealization, since counters are notorious for having a dead time after they are triggered), then we would with certainty observe the particle inside the volume of the counter.

In the second example, we consider a momentum measurement with a counter which analyzes the pulse height of a recoil particle.

Here the situation is quite different.

The experiment will permit us to determine the value of the momentum only before the collision occurred.

If we repeat the measurement immediately after it has occurred, then we find that the momentum of the particle will have a quite different value from its measured value.

The very act of measurement has changed the momentum, and it is this change which produced the observable effect.

We shall call a measurement which will give the same value when immediately repeated a measurement of the *first kind*.

The second example is then a measurement of the *second kind*.

From now on we shall be primarily concerned with measurements of the first kind.

They are easier to discuss, yet they exhibit the characteristic quantum features which we want to explore here.

Measurements of the first kind can be used for *preparing* a state with definite values for certain observables.

If they are used in this way we speak of *filters*.

For the preparation of such a state, the condition imposed by the measuring device becomes a relevant condition in the preparation of the state.

The states before and after the measurement are not the same, even if the duration of the measurement can be considered negligible.

The change in the state must be attributed to the interaction of the system with the measuring device.

We shall now obtain a formula for the change of the state by a measurement of the first kind.

Let us begin with a simple special case where the observable to be measured is represented by a self-adjoint operator A with nondegenerate eigenvalues.

Let ψ_r , be the eigenvectors of A, and denote by P_r, the projection which contains ψ_r , in its range.

The operator A may then be represented in the form

$$A = \sum a_r P_r, \tag{8}$$

where a_r denotes the eigenvalues of A.

A measurement of the quantity A in a system in the state W will give the result a_r with a probability which is given by

$$p(P_r) = \mathrm{Tr}\, W P_r. \tag{9}$$

If the measurement is of the first kind, then the repetition of a measurement which yielded the value a_r will reproduce this value with certainty.

If the measurement is not used as a filter - that is, if it is not used to select a certain sub-ensemble of systems which have definite values for A - then the state ofthe system after the measurement will be a mixture of the pure states ψ_r, with probabilities as expressed by Eq. (9).

The density operator for this mixture is given by

$$W' = \sum_r P_r W P_r. \tag{10}$$

To see this we verify that for projections P with one-dimensional range we have

$$PWP = (\mathrm{Tr}\, WP)P.$$

Let ψ by a unit vector in the range of P, and let f be an arbitrary vector.

We have then, by definition of the projection operator P,

$$PWPf = (\psi, W\psi)(\psi, f)\psi = (\psi, W\psi)Pf.$$

Thus it suffices to show that

$$(\psi, W\psi) = \mathrm{Tr}\, WP.$$

This can be verified by evaluating the trace in a special orthonormal system ψ_r, which is so chosen that its first vector $\psi_1 = \psi$.

We have then, because P has one-dimensional range, $P\psi_r = 0$ for all $r \neq 1$.

Thus the infinite sum in the evaluation of the trace reduces to one single term, the term on the left-hand side of the last equation.

With this we have established formula Eq. (10) for the case that the ranges of all P_r are one-dimensional.

If the eigenvalues a_r, of the observable A are degenerate, then the state of the system referring to a subspace associated with an a_r depends on the detailed nature of the measuring equipment.

The degeneracy leaves us with a certain freedom of choice for the state after the measurement, which cannot be removed without a detailed knowledge of the actual measuring equipment used.

In such cases the measurement of one and the same observable with different equipment may result in different states after the interaction.

It is therefore convenient to introduce the notion of the ideal measurement which affects the state in a minimal way, and for which the state after the measurement is still given by formula Eq. (10) but without the requirement that the projection operators P_r be one-dimensional.

For the special case of the ideal measurement of a projection operator E, we shall then obtain for the state after measurement the density operator

$$W' = EWE + E'WE', \tag{11}$$

where $E' = I - E$ and W is the density operator before the measurement.

If the projection operator E is used to describe a filter, then we shall call it a passive filter if after the filter process the state is given by

$$W' = \frac{1}{\mathrm{Tr}\, WE} EWE. \tag{12}$$

This state is obtained if we prepare the state W, and add to the conditions which prepare this state the further relevant condition that the measurement of the property represented by E must be true.

When we examine the change of the state as expressed, for instance, in Eq. (11), we observe an important point.

This change is quite different from the change of the state due to the time-evolution of a system, which is expressed by a formula such as

$$W_t = e^{-iHt} W e^{iHt}, \tag{13}$$

with some Hamiltonian operator H.

The fundamental character of the difference between Eq. (11) and Eq. (13) can be seen by the fact that the transformation $W \to W_t$, is unitary, while $W \to W'$ usually is not.

For instance, if the state W is pure and E does not commute with W, then W' is always a mixture.

There is one special case when the observation of a property E does not affect the state, namely, when E commutes with W.

If E commutes with W, we may in fact write $EWE = WE$ and $E'WE' = WE'$; therefore $W' = EWE + E'WE'$ $= WE + WE' = W(E + E') = W$.

This condition is also necessary.

Indeed, if $W' = W$, then $WE = W'E = EWE = EW' = EW$.

Thus we have established:

> *The necessary and sufficient condition such that an ideal measurement of a proposition represented by the projection operator E does not disturb a state W is that E commute with W.*

The measurements which do not disturb the state are thus very special cases; in general we must expect a change which leads to a density operator W' which is not unitarily equivalent to the operator W.

This behavior of the state under measurement is at first sight very strange, because the unitary transformation of states is derived under very general conditions.

Even if the system is subject to variable external forces, we must expect such a unitary transformation of the change.

It seems as if the unity of the description of the physical laws is broken at this point: If we have the system under its own influence the state changes in one way, given by Eq. (13); if we observe or measure something on the state, the state changes in another entirely different way, given by Eq. (11).

We might suppose this difference to be due to the interaction of the system S with the measuring device M.

The difficulty with this explanation is that one is always at liberty to consider the time evolution of the combined system, consisting of S and M.

This combined system, if left undisturbed by outside influences, then evolves in time according to Eq. (13) where the operator H is now the total Hamiltonian of the combined system.

That such an operator must exist follows from the fact that the combined system $S + M$ is also subject to the laws of quantum mechanics, even if it is, as in most practical cases, a very complicated system.

The complication should not detract us from the essence of the question, namely, how to reconcile the two different behaviors of the state vectors without violating the unity of the laws of nature.

We shall analyze and answer this question in the subsequent discussion.

PROPERTIES OF THE MEASURING DEVICE

The general description of the measuring process that we have given in the preceding discussion did not specify the nature of the measuring device M.

This we shall do now.

It is clear that not every system M will be suitable as a measuring apparatus.

In order to fulfill its function of determining the measurable physical properties of the system S, it must satisfy certain conditions which we shall formulate now.

If we examine any measuring apparatus M commonly used for the measurements of quantum systems S, then we observe at once a common feature: Every measuring device is usually a large system, producing macroscopic effects which can be observed or recorded with equipment for which the quantum effects are completely irrelevant.

For instance, if our apparatus M is a counter of some sort, it will contain a large number of molecules in an unstable state which can be triggered by the micro-event to be measured.

The triggering then produces a chain reaction which may lead, for instance, to a potential drop on a capacitor, and eventually recorded in some mechanical counter.

The resulting large-scale effect in the last stage of the measurement can then be read off by any observer without interfering in the least with the state of the apparatus.

The final stage of the observation can therefore be described completely in classical terms alone.

This last remark is very important, and it has played an essential part in the analysis of the measuring process given, for instance, by Bohr.

It is this classical aspect of the measuring device which enables us to establish the objective aspects of the state of a system, and it is from such objectively given ingredients that the physical laws governing the microsystems must be reconstructed.

Thus we must assume that measuring devices are usually constructed out of macroscopic parts which can be adequately described by classical laws; yet this device must be sensitive to the quantum features of the system S.

This is often accomplished by the use of metastable systems which can be triggered by micro-events, which subsequently are amplified to macro-events.

A typical example is the bubble chamber using a superheated liquid as the metastable macrosystem.

A single micro-event in the form of the ionization of an atom may be sufficient to trigger the sequence of events which lead to the formation of a bubble.

The classical aspect of the measuring device which we have stressed here is usually associated with its macroscopic features.

We must, however, be on guard not to confuse the two.

To be sure, most macroscopic objects-that is, objects which consist of a very large number of microsystems do, in most situations, behave according to the classical laws of physics.

But this is not necessarily so.

There are macro-objects which exhibit quantal features, and there are micro-objects which, for certain types of observations, behave classically.

The part of the measuring system which amplifies the triggered event to a macroscopic phenomenon is therefore a convenient device which enables us to observe and register this event, but it is not absolutely indispensable for the completion of a measurement.

This means that the characteristic difficulties in understanding the measuring process are not to be attributed to the inevitable complication of an amplifying device; they are already present in the microscopic part of the measuring equipment.

In order to make this explicit, it is convenient to divide the measuring device M into two parts, $M = m + A$.

Here m denotes the" little" measuring device, that is, that part which is truly responsible for the measurement of the system M; and A represents the amplifying part of M.

Part A may be disconnected in thought or sometimes in reality from the little measuring device, without impairing the essential part of the measuring process.

For instance, if our measuring equipment is the photographic plate, the little measuring device m may be a single silver-halide complex, while the amplifying device A may be an individual grain, containing a large number of atoms.

The record of an event may be triggered by a single ionization process and stored as a latent image long before the amplification A in the form of the development of the plate is put into effect.

The system $S + m$ which really performs the measurement must satisfy further restrictions.

The function of the measurement consists in transferring certain properties of the system S to the system m, in such a way that a mere observation on the system m (by means of the amplifier A) will permit us to draw certain conclusions as to the state of the system S.

There must therefore exist a correlation between the events on M and the states of S.

The measuring device can then distinguish states of S which are correlated with observable events in M.

In a measurement of the first kind, we have the additional requirement that a repetition of the measurement immediately after it has occurred will reproduce the same result.

This means that the system S is left after the measurement in the state registered by M.

We have now listed the essential properties which we must require of the measuring device for a measurement of the first kind.

EQUIVALENT STATES

Now we shall now examine in detail the modification in the description of states which results from a restricted system of observables.

In the preceding discussion we have pointed out that the observables which are measured by a measuring device must be of a classical nature in order that the measurement have an objective character.

This means that one and the same device M can, in general, measure only a restricted class of observables, all of which must commute with one another.

This has the consequence that certain states which may be represented by different density operators may actually be indistinguishable with respect to this system of observables.

The natural way of describing this situation is by introducing a theory of equivalence classes of states.

We shall first do this quite generally, without the requirement that we are dealing with classical observables, and shall specialize later for classical observables.

Let $\mathcal{S}$ be a system of observables.

We say that two states W_1 and W_2 are *equivalent* with respect to $\mathcal{S}$ if

$$\operatorname{Tr} AW_1 = \operatorname{Tr} AW_2 \tag{14}$$

for all $A \in S$.

We shall write for two states equivalent with respect to $\mathcal{S}$: $W_1 \sim W_2(\mathcal{S})$.

If two states are equivalent in this sense, then no measurement with observables from $\mathcal{S}$ can distinguish the two states.

It is not difficult to show that $W_1 \sim W_2$ is an equivalence relation.

We can therefore divide the set of all states into classes of equivalent states.

From the physical point of view, the selection of an individual representative inside a class of equivalent states is irrelevant; any choice will be equally good.

It is therefore quite natural to consider the states with respect to a system $\mathscr{S}$ not as one of the members of the class but as the classes themselves.

By this procedure we remove the redundancy in the description of the state, and we restore the one-to-one correspondence between the physical notion of state and its mathematical description.

Let us denote by $[W]$ the class of the states which are all equivalent to W.

We shall call the redundant states W the *microstates* and the classes of equivalent states $[W]$ the *macrostates*.

The operation of the mixture of states can be transferred from the microstates to the macrostates.

This means we claim the property expressed in the formula

$$[\lambda_1 W_1 + \lambda_2 W_2] = \lambda_1 [W_1] + \lambda_2 [W_2]. \tag{15}$$

This formula implies the following:

$$W_1 \sim W_1' \qquad \text{and} \qquad W_2 \sim W_2'$$

and if

$$W = \lambda_1 W_1 + \lambda_2 W_2, \qquad W' = \lambda_1 W_1' + \lambda_2 W_2',$$

then $W \sim W'$.

Due to this property it is possible to transfer the process of mixing from the microstates to the macrostates, as we have indicated with the notation of the right-hand side of Eq. (15).

In order to prove this property, we mention that it is sufficient to verify it for the projections contained in $\mathscr{S}$.

Let $\mathscr{S}_P$ denote the projections $\mathscr{S}$.

We want to show that

$$\mathrm{Tr}\, EW = \mathrm{Tr}\, EW' \qquad \text{for all } E \in \mathscr{S}_P.$$

This we see by using the linearity property of the trace through the following sequence of steps:

$$\begin{aligned} \mathrm{Tr}\, EW &= \mathrm{Tr}\, E(\lambda_1 W_1 + \lambda_2 W_2) \\ &= \lambda_1 \mathrm{Tr}\, EW_1 + \lambda_2 \mathrm{Tr}\, EW_2 \\ &= \lambda_1 \mathrm{Tr}\, EW_1' + \lambda_2 \mathrm{Tr}\, EW_2' \\ &= \mathrm{Tr}\, E(\lambda_1 W_1' + \lambda_2 W_2') = \mathrm{Tr}\, EW'. \end{aligned}$$

In the third step we have made use of the fact that $W_1 \sim W'_1$ and $W_2 \sim W'_2$ ($\mathscr{S}$).

Thus the property is verified.

We shall have occasion to use a corollary which states that if $W_1 \sim W_2$, then any mixture $W = \lambda_1 W_1 + \lambda_2 W_2$ is in the same equivalence class $[W_1] = [W_2]$.

The equivalence classes are thus closed against the operation of mixing.

This implies that every class which contains more than one micro- state contains mixtures.

An important question which will concern us now is this: What happens to the macrostates during a measurement?

We know what happens to microstates.

If, for instance, the state is W before the measurement, then after the measurement of E it is

$$W^E \equiv EWE + E'WE', \tag{16}$$

where we have written $E' = I - E$.

Let $[W]$ be the class which contains W and $[W^E]$ the class which contains W^E.

We shall now determine under what condition the class $[W^E]$ is independent of the representative W in the class $[W]$.

Thus we choose two members $W_1, W_2 \in [W]$ from the same class, so that $W_1 \sim W_2$.

The condition that $W_1^E \sim W_2^E$ is that for any $F \in \mathscr{S}_P$ we have

$$\mathrm{Tr}\, F(EW_1E + E'W_1E') = \mathrm{Tr}\, F(EW_2E + E'W_2E').$$

By using the invariance property of the trace under cyclic permutations of the operators, we can change this into

$$\mathrm{Tr}\,(EFE + E'FE')W_1 = \mathrm{Tr}\,(EFE + E'FE')W_2.$$

This must be true for all projection operators $F \in \mathscr{S}_P$.

This is true for all equivalent pairs W_1 , W_2 if and only if

$$F^E \equiv EFE + E'FE' \in \mathscr{S}. \tag{17}$$

This, then, is the condition which guarantees that the macrostates are not broken up under the process of measurement.

We shall be particularly interested in the case of an abelian set of observables.

In that case $F^E = F$, and so condition in Eq. (17) is always true.

When this condition is satisfied we may transfer the process of change under measurements from the micro- to the macrostates and we can write a formula

$$[W^E] = [W]^E$$

where the right-hand side denotes the class which contains the element W^E.

We are now also in a position to answer the question under what condition a state is left invariant under all the measurements.

The condition for this is that

$$[W]^E = [W] \qquad \text{or} \qquad W^E \sim W \qquad \text{for all } W \in [W].$$

This means that for all $E \in \mathscr{S}_P$ and all $F \in \mathscr{S}_P$ we must have

$$\operatorname{Tr} FW^E = \operatorname{Tr} FW = \operatorname{Tr} F^E W.$$

A sufficient condition for this to hold is

$$F^E \equiv EFE + E'FE' = F \tag{18}$$

for all $E, F \in \mathscr{S}_P$.

If we require invariance for every state, then this condition is also necessary.

Since invariance of the states under measurements is a property of idealized classical systems, we shall call a state which is invariant under all measurements a *classical state*.

For a classical system the relation in Eq. (18) is always satisfied since E and F commute.

It follows that in a classical system every state is a classical state.

EVENTS AND DATA

One of the chief difficulties in the epistemology of quantum mechanics is its apparent inadequacy for describing events.

The fact that there are systems which do not admit dispersion-free states leads to the inevitable and irreducible probability statements regarding the occurrence of certain events.

Such events may be the measurements associated with yes-no experiments, and as such they may be macroscopic phenomena.

The individual occurrence of such phenomena is then completely outside the scope of the theory; only the probabilities for such events can be accounted for in our description of the state.

In order to understand this problem better we may compare it with the classical situation.

The occurrence of probability statements in the description of states is not unknown in classical mechanics.

Most systems, with a large number of degrees of freedom, are much better described with states which are not dispersion-free.

To illustrate this, suppose we want to describe the state of a thrown die before it is examined to learn what number it shows.

Such a system is in a state determined by the conditions of the throw, but this state can only be described by a probability function as to the occurrence on top of one of the six sides.

The initial condition or preparation which determined this state could be described by a rule such as: Throw die in the air not higher than so and so, and let it come to rest on the table.

From our experience with classical objects such as dice, we know that this description results in a probability statement for the outcome of the throw merely because it is not precise enough.

We know from experience with similar systems that the specification of the preparation (throwing a die) can be made more precise, enough so as to determine the outcome of the throw with certainty.

It is possible to add other relevant conditions to the prescription for preparing the state.

For instance, we might specify the exact position, direction of throw, air currents in the vicinity, angular momentum, and many other variables, to such a degree of precision that the result of the throw will always be the same under the same conditions and we can predict it with certainty.

In this case the state has no dispersion any more.

If we examine this question for quantum systems, then we find that it is not always possible to add conditions which permit the preparation of states without dispersion.

Adding further conditions may indeed change the state but it will not make it dispersion-free.

The classical states show dispersion not because of any intrinsic occurrence of probabilities, but because the prescriptions for preparing the states already involve probabilities.

This gives us justification for considering the probabilities as expressions of our ignorance of the finer features of the state.

We do not have any doubt that a die, when it has come to rest, does indeed show one of the numbers before we look at it; the final act of observation does not produce this number--it merely uncovers a fact which has already occurred.

It is convenient to have a special technical terminology for the description of such facts to which we want to ascribe reality before they are observed.

We shall call them *events*.

The important property of events is that they represent objectively given phenomena capable of being determined by observations which in no way interfere with the state of the system.

When an event has been observed we call it a *datum*.

While in classical mechanics every physical property which, after an observation, is found to be true, may be called an event (because we can always arrange it so that such an observation does not affect the state), the situation is more complex in quantum mechanics.

It was pointed out by Einstein that we may have events in quantum mechanics, too.

Einstein used the term "element of reality" for the description of properties which we have called events.

There are, however, also properties which do not have this element of reality and which cannot be called events.

It is of the utmost importance in the analysis of the measuring process to be able to distinguish between the two kinds of properties.

Let W be a state and E a projection representing a yes-no experiment.

The property E is an event if and only if the measurement of E does not affect the state W.

We have previously shown that this is the case if and only if W commutes with E.

We may then affirm that each individual system of an ensemble of identically prepared systems in the state W realizes one or several of the events; which of these events are in fact realized for an individual system is determined in principle by making measurements of all the events E on that individual system.

Since none of the measurements change the state, all the results which are obtained pertain to that one system in the unchangeable state W.

A little more delicate is the question concerning events if the system of observables is a restricted class of operators $\mathcal{S}$.

As we have seen in the preceding section, the proper description of states in this case consists of the classes of equivalent states, that is, the macrostates.

We are then entitled to affirm that a projection E is an event in the state W if $[W^E] = [W]^E = [W]$.

Thus the observation E may very well change the states in one and the same class but it does not change the class.

This condition is weaker than the one we had for microstates.

W need not commute with E; it is only necessary that it leave the macrostate unchanged, to be an event.

This remark will be very important in the analysis of the measuring process, since the necessary classical feature of the measuring equipment implies the restriction of the observables of the measuring device to an abelian subset of all the observables.

MATHEMATICAL INTERLUDE: THE TENSOR PRODUCT

In the following parts of the analysis of the measuring process, we need to apply the theory of the tensor product of Hilbert spaces, which we shall develop now.

The tensor product is involved whenever we consider the union or separation of two subsystems; a process which occurs precisely during a measurement; the two systems in this case are the system to be measured on the one hand and the measuring equipment on the other.

The reason for the occurrence of the tensor product may be seen from the following remarks.

Let S_1 and S_2 be two systems, and let $\mathscr{S}_1$ be a complete set of commuting observables of S_1, and $\mathscr{S}_2$, such a set for $S2$.

Every observable $A_1 \in \mathscr{S}_1$, is naturally also an observable on the joint system $S_1 + S_2$.

The same is true for every observable $A_2 \in \mathscr{S}_2$.

Furthermore every observable A_1 commutes with every observable A_2, and $\{\mathscr{S}_1, \mathscr{S}_2\}$ a complete set of commuting observables for $S_1 + S_2$.

We must thus find a description which will incorporate these characteristic properties of the union of two systems.

This can be done in the spectral representation for the two systems $\mathscr{S}_1$ and $\mathscr{S}_2$, for instance, as follows:

Let Λ_1 be the Cartesian product of the spectra of a complete commuting set of some of the observables $\mathscr{S}_1$; and similarly, let Λ_2 be the corresponding product of the spectra of some of the observables $\mathscr{S}_2$.

The Hilbert space $\mathscr{H}_1$, of the spectral representation for $\mathscr{S}_1$, consists of square-integrable functions $\varphi_1(\lambda_1)$ with $\lambda_1 \in \Lambda_1$.

Similarly the Hilbert space $\mathscr{H}_2$ of the spectral representation for $\mathscr{S}_2$, consists of square-integrable functions $\varphi_2(\lambda_2)$ with $\lambda_2 \in \Lambda_2$.

The Hilbert space $\mathcal{G}$ of the spectral representation for the set $\{\mathscr{S}_1, \mathscr{S}_2\}$ is then a set of functions $\varphi(\lambda_1, \lambda_2)$, square-integrable over the Cartesian product space $\Lambda_1 \times \Lambda_2$.

In this way we arrive quite naturally at the notion of the *tensor product*.

Associated with the pair of spaces $\mathscr{H}_1$ and $\mathscr{H}_2$, we have constructed another space $\mathcal{G}$, the tensor product of $\mathscr{H}_1$ and $\mathscr{H}_2$, which in some respect generalizes the notion of product.

The measure ρ for the space $\Lambda_1 \times \Lambda_2$ may be taken as the product measure $\rho_1\rho_2$, of the measures for the spaces Λ_1 and Λ_2, separately.

In this way we arrive quite naturally at the notion of the *tensor product*.

Associated with the pair of spaces $\mathscr{H}_1$ and $\mathscr{H}_2$, we have constructed another space $\mathcal{G}$, the tensor product of $\mathscr{H}_1$, and $\mathscr{H}_2$, which in some respect generalizes the notion of product.

The same $\mathcal{G}$ contains a certain subset of vectors, those of the form $\psi(\lambda_1, \lambda_2) = \psi_1(\lambda_1)\psi_2(\lambda_2)$.

This subset is the image of a bilinear mapping of pairs of vectors $\psi_1 \in \mathcal{H}_1$, and $\psi_2 \in \mathcal{H}_2$ into $\mathcal{G}$.

Furthermore, this image set is a complete set of vectors in $\mathcal{G}$ in the sense that every vector in $\mathcal{G}$ can be written as a linear combination of vectors of the form $\psi_1(\lambda_1)\psi_2(\lambda_2)$.

These last two remarks will permit us to free ourselves from the particular construction of the tensor product which we have adopted here for illustrative purposes.

We shall now proceed to a formal and abstract definition of the tensor products.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces. The tensor product $\mathcal{G} = \mathcal{H}_1 \otimes \mathcal{H}_2$ *is a Hilbert space together with a bilinear mapping φ from the topological product $\mathcal{H}_1 \times \mathcal{H}_2$ into $\mathcal{G}$, such that*

1) *the set of all vectors* $\varphi(f_1, f_2)$ (where $f_1 \in \mathcal{H}_1, f_2 \in \mathcal{H}_2$) spans $\mathcal{G}$;
2) $(\varphi(f_1, f_2), \varphi(g_1, g_2)) = (f_1, g_1)(f_2, g_2)$ *for all* $f_1, g_1 \in \mathcal{H}_1$, *and for all* $f_2, g_2 \in \mathcal{H}_2$.

A few remarks on this definition and the notation may enhance the reader's understanding of them:

We are using two different kinds of products, the *topological* product $\mathcal{H}_1 \times \mathcal{H}_2$, and the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$, which should not be confused.

The first consists simply of the pairs of vectors $\{f_1, f_2\}$ with $f_1 \in \mathcal{H}_1$, and $f_2 \in \mathcal{H}_2$.
The second is a Hilbert space $\mathcal{G}$ together with a bilinear mapping φ which satisfies the two conditions indicated above.

We have designated the scalar product in $\mathcal{G}$ with the same bracket notation as the scalar product in $\mathcal{H}_1$ and $\mathcal{H}_2$.
The reader should not confuse the two.

It is important to note that the image of the bilinear mapping φ is not the entire Hilbert space $\mathcal{G}$; it is only a proper subset of $\mathcal{G}$.

However, this subset is total in the sense that it spans the entire space $\mathcal{G}$.
The definition need not be restricted to the product of two spaces.

Indeed, we may define the tensor product of a finite number n of spaces $\mathcal{G} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n$ as a Hilbert space $\mathcal{G}$ together with a multilinear mapping from $\mathcal{H}_1 \times \mathcal{H}_2 \times \cdots \times \mathcal{H}_n$ into $\mathcal{G}$ which satisfies

1′) the set of all vectors $\varphi(f_1, f_2, \ldots, f_n)(f_r \in \mathcal{H}_r)$ spans $\mathcal{G}$;

2′) $(\varphi(f_1, f_2, \ldots, f_n), \varphi(g_1, g_2, \ldots, g_n)) = (f_1, g_1)(f_2, g_2)\cdots(f_n, g_n)$ for all $f_r, g_r \in \mathcal{H}_r$ $(r = 1, 2, \ldots, n)$.

We have already proved the existence of the tensor product for $n = 2$ by the construction employed in our example at the beginning of this discussion.

We now need only verify that this construction does indeed satisfy the conditions (1) and (2) of the definition.

The tensor product is unique in the following precise sense: Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces and denote by $\mathcal{G}$ and $\mathcal{G}'$ two different tensor products with the associated bilinear mappings φ and φ' respectively; then there exists a unique isometric operator U with domain $\mathcal{G}$ and range $\mathcal{G}'$, such that

$$U\varphi(f_1, f_2) = \varphi'(f_1, f_2) \qquad \text{for all } f_1 \in \mathscr{H}'_1, f_2 \in \mathscr{H}'_2. \tag{19}$$

This uniqueness property of the tensor product is very important for its physical interpretation.

The projection operators in $\mathcal{G}$ are the yes-no experiments for the joint system.

The representation of the Hilbert space is determined only up to unitary equivalence by the lattice structure of the propositions.

Then the uniqueness of the tensor product means that the physical property of the joint system, insofar as it is contained in the algebraic structure ofthe lattice of propositions, is entirely determined by the structure of the component systems.

An explicit construction of the tensor product, independent of any particular reference system, can be given as follows:

We define first the notion of the conjugate linear transformation T from $\mathscr{H}_2$ into $\mathscr{H}_1$.

Such a T is required to satisfy

$$\begin{aligned} T(f_2 + g_2) &= Tf_2 + Tg_2, \\ T(\lambda f_2) &= \lambda^* T f_2, \end{aligned} \tag{20}$$

for all $f_2, g_2 \in \mathscr{H}_2$.

Let $\{\psi_r\}$ be a complete orthonormal system in $\mathscr{H}_2$, and define the norm

$$\|T\|^2 \equiv \sum_{r=1}^{\infty} \|T\psi_r\|^2. \tag{21}$$

The sum on the right-hand side need not be finite, but it is independent of the choice of the orthonormal system $\{\psi_r\}$.

This norm satisfies the parallelogram identity, so that it can be derived from a scalar product as follows:

$$(T, S) \equiv \sum_{r=1}^{\infty} (T\psi_r, S\psi_r). \tag{22}$$

The set of all conjugate linear mappings T, with $\|T\| < \infty$, are therefore a Hilbert space .

Moreover there exists a bilinear mapping φ from $\mathscr{H}_1 \times \mathscr{H}_2$ into $\mathcal{G}$ by setting, for each pair of vectors f_1, f_2,

$$Tg_2 = (g_2, f_2) f_1. \tag{23}$$

We denote this particular T by $f_1 \otimes f_2$.

Its norm is $\|f_1 \otimes f_2\| = \|f_1\| \; \|f_2\|$.

Now let $T = f_1 \otimes f_2$ and $S = g_1 \otimes g_2$; then one verifies easily that

$$(T, S) = (f_1, g_1)(f_2, g_2). \tag{24}$$

Thus we have completed the construction of the tensor product: The elements of the space $\mathcal{G}$ are the conjugate linear mappings of $\mathcal{H}_2$ into $\mathcal{H}_1$ with finite norm and the bilinear mapping φ is given by Eq.(23).

To every $T \in \mathcal{G}$ we can uniquely associate another conjugate linear mapping $T^{\#}$ from $\mathcal{H}_1$ into $\mathcal{H}_2$ by the formula

$$(f_1, Tf_2) = (f_2, T^{\#} f_1).$$

One verifies the following rules of this correspondence:

$$T^{\#\#} = T, \qquad \|T^{\#}\| = \|T\|,$$

$$(T + S)^{\#} = T^{\#} + S^{\#}, \qquad (\lambda T)^{\#} = \lambda T^{\#}.$$

If $T = f_1 \otimes f_2$, so that, according to Eq.(23), $Tg_2 = (g_2, f_2)\, f_1$, then we find that

$$(g_1, Tg_2) = (g_1, f_1)(g_2, f_2) = (g_2, T^{\#} g_1);$$

therefore we must have $T^{\#} g_1 = (g_1, f_1)\, f_2$.

It is suggestive to denote this particular $T^{\#}$ by $T^{\#} = f_2 \otimes f_1$.

The notion of the product can be extended from the vectors to the linear operators.

Thus, if we have two bounded operators A_1 and A_2, where A_1 operates in $\mathcal{H}_1$, and A_2 operates in $\mathcal{H}_2$, then we may define for every $T = f_1 \otimes f_2$ the operation

$$(A_1 \otimes A_2)(f_1 \otimes f_2) = Af_1 \otimes Af_2.$$

This definition of $A_1 \otimes A_2$ can be extended by linearity to the entire space $\mathcal{G}$.

Then we have, for every $T \in \mathcal{G}$, the formula

$$(A_1 \otimes A_2)T = A_1 T A_2^*, \tag{25}$$

where A_2^* is the Hermitian conjugate of A_2.

This completes the construction of the tensor product.

THE UNION AND SEPARATION OF SYSTEMS

We shall now consider two coherent systems S_1 and S_2.

The joint system we denote by $S_1 + S_2$.

The principal question that we want to answer here is how the states of the component systems are related to the states of the joint system, and vice versa.

Let us begin with the first part of the question.

The states of the component system shall be given by their respective density operators W_1 and W_2.

We wish to know what the state is if we consider the two components together as a joint system.

The criterion for answering this question is a physical one: If we measure observables which refer to only one of the components we must obtain the same result whether we consider them measured on the joint system or on the component system.

We shall denote by A_1 an observable which refers to the component S_1.

As an observable on S_1 it is a self-adjoint operator in the Hilbert space $\mathcal{H}_1$, pertaining to the system S_1.

But as an observable on the joint system $S_1 + S_2$, it is a self-adjoint operator in the Hilbert space $\mathcal{G} = \mathcal{H}_1$ @ $\mathcal{H}_2$ pertaining to the joint system.

Since it is an observable on S_1 alone, it must be the identity operator I_2 in $\mathcal{H}_2$.

That is, it must have the form $\mathbf{A}_1 = A_1 \otimes I_2$.

Similarly an observable on the joint system which refers only to system S_2 must have the form $\mathbf{A}_2 = I_1 \otimes A_2$.

Let W_1 be the state of system S_1 and W_2 the state of system S_2.

The joint system $S_1 + S_2$ is then in a state W for which we now want to determine the manner of its dependence on W_1 and W_2.

The conditions which W must satisfy are thus

$$\mathrm{Tr}\, \mathbf{A}_1 W = \mathrm{Tr}_1 A_1 W_1 \qquad \mathrm{Tr}\, \mathbf{A}_2 W = \mathrm{Tr}_2 A_2 W_2. \tag{26}$$

Here we have, for greater clarity, introduced the notation Tr_1 and Tr_2 for traces which refer only to $\mathscr{H}_1$ and $\mathscr{H}_2$, respectively.

One possible solution of Eq. (26) is given by $W = W_1 \otimes W_2$, as one may verify immediately.

In general this is not the only solution.

Physically this means that the state of the compound system $S_1 + S_2$ cannot be completely determined by measurements on the component systems alone.

There are thus physically distinguishable properties which express themselves as correlations between observation on S_1 and S_2.

Such correlations are nonexistent for the state $W_1 \otimes W_2$.

There is one exception: If one of the states W_1 or W_2 is pure, then the states W_1 and W_2 determine the state of the compound system uniquely.

Let us now examine the reverse problem: Given W, determine W_1 and W_2 such that Eq. (26) holds.

This problem always has a unique solution.

Let us first demonstrate the uniqueness of the solution.

Let W, W_1, and W_2 be three density operators satisfying Eq. (26) for all A_1 and A_2.

Let us assume that W, W'_1, and W'_2 is another set of such operators satisfying also the conditions of Eq. (26).

We find then that

$$\mathrm{Tr}_1\, A_1 W_1 = \mathrm{Tr}_1\, A_1 W'_1 \tag{27}$$

for all observables $A_1 \in \mathscr{S}_1$.

Since S_1 was assumed to be a coherent system, this means that Eq. (27) must be true for all projections P, in particular one-dimensional ones.

Let P be such a projection and φ a unit vector in its range.

Then we obtain from Eq. (27)

$$(\varphi,\ W_1\varphi) = (\varphi,\ W'_1\varphi)$$

for all φ.

This is possible only if $W_1 = W'_1$.

Similarly one proves $W_2 = W'_2$; This proves uniqueness.

Let us next assume that W is a mixture, for instance, $W = \lambda U + \mu V$.

Let U_1 and U_2 be the component states determined by U, and similarly let V_1 and V_2 be the component states determined by V.

The linearity of the connection Eq. (26) between W and W_1, W_2 then results immediately in the statement: The component states W_1 and W_2 corresponding to the state W are given by

$$W_1 = \lambda U_1 + \mu V_1, \qquad W_2 = \lambda U_2 + \mu V_2. \tag{28}$$

This shows in particular that the component states can be pure only if W is pure.

In any case we need only determine the component states for pure states W; for the mixtures they can, by means of Eq. (28), be calculated immediately in terms of the pure states contained in the mixture.

Let us then assume that W is pure and denote by $\boldsymbol{\Phi}$ the vector in $\mathcal{G}$ contained in the range of W.

$\boldsymbol{\Phi}$ is thus an antilinear mapping of $\mathscr{H}_2$ into $\mathscr{H}_1$.

We assume $\boldsymbol{\Phi}$ normalized, so that $\|\boldsymbol{\Phi}\| = 1$.

Let us choose a complete orthonormal system $\Phi_n \in \mathcal{G}$ such that $\Phi_1 = \Phi$. We then obtain

$$\operatorname{Tr} \mathbf{A}_2 W = \sum_n (\Phi_n, \mathbf{A}_2 W \Phi_n) = (\Phi, \mathbf{A}_2 \Phi).$$

Since $\mathbf{A}_2 = I_1 \otimes A_2$ we use Eq. (25) and obtain

$$\mathbf{A}_2 \Phi = \Phi A_2^* = \Phi A_2.$$

Therefore

$$(\Phi, \mathbf{A}_2 \Phi) = \operatorname{Tr}_1 \Phi^* \Phi A_2$$

which shows that $W_2 = \Phi^{\#}\Phi$.

In a similar way we evaluate W.

The result is

$$W_1 = \Phi\Phi^{\#}, \qquad W_2 = \Phi^{\#}\Phi. \tag{29}$$

In the general case in which W is a mixture of orthogonal states Φ_n with weights λ_n, the component states are

$$W_1 = \sum \lambda_n \Phi_n \Phi_n^{\#} \qquad \text{and} \qquad W_2 = \sum \lambda_n \Phi_n^{\#} \Phi_n. \tag{30}$$

We shall refer to Eqs. (29) and (30) as the *reduction formulas*, and the states W_1 , W_2 are called the *reduced* (or component) *states*.

Let us now discuss the reduction formula Eq. (29) in a little more detail.

Consider first the case that $\boldsymbol{\Phi} = \varphi \otimes \psi$ where φ and W are both normalized and $\varphi \in \mathscr{H}_1, \psi \in \mathscr{H}_2$.

From the definition of $\boldsymbol{\Phi}$ and $\boldsymbol{\Phi}^{\#}$ it follows that

$$\Phi^{\#}\varphi = (\varphi, \varphi)\psi = \psi,$$

$$\Phi\psi = (\psi, \psi)\varphi = \varphi.$$

Thus $\boldsymbol{\Phi}\,\boldsymbol{\Phi}^{\#}\,\varphi = \varphi$ and $\boldsymbol{\Phi}^{\#}\,\boldsymbol{\Phi}\,\psi = \psi$.

Furthermore, if φ_1 is orthogonal to ψ, so that $(\varphi_1, \varphi) = 0$, then

$$\Phi^{\#}\varphi_1 = (\varphi_1, \varphi)\psi = 0$$

so that $\boldsymbol{\Phi}\,\boldsymbol{\Phi}^{\#}\,\varphi_1 = 0$.

Similarly if ψ1 is orthogonal to ψ, then $\boldsymbol{\Phi}^{\#}\boldsymbol{\Phi}\psi_1 = 0$.

In this case we find that $\boldsymbol{\Phi}\,\boldsymbol{\Phi}^{\#} = P$ is a projection in $\mathcal{H}_1$, with one-dimensional range containing φ.

Similarly $\boldsymbol{\Phi}^{\#}\,\boldsymbol{\Phi} = Q$ is a projection in $\mathcal{H}_2$, with one-dimensional range, containing ψ.

This means if the compound state has the form $\boldsymbol{\Phi} = \varphi \otimes \psi$, then the reduced states are pure.

The converse we have already seen, and so we have established:

The reduced states are pure if and only if the pure state $\boldsymbol{\Phi}$ *is of the form* $\varphi \otimes \psi$.

Let us now consider the case in which the compound state is still pure, but not of this form.

Then we know from the previous discussion that neither W_1 nor W_2 can be pure.

Let $\boldsymbol{\Phi}\,\boldsymbol{\Phi}^{\#} = W_1 = \sum_r \alpha_r P_r$ with P_r projections with one-dimensional range and $\alpha_r > 0, \sum_r \alpha_r = 1$.

Define $\psi_r = (1/\sqrt{\alpha_r})\,\boldsymbol{\Phi}^{\#}\,\varphi_r$, where φ_r is a normalized vector in the range of P_r.

We have then

$$\|\psi_r\|^2 = \frac{1}{\alpha_r}\|\Phi^{\#}\varphi_r\|^2 = \frac{1}{\alpha_r}(\varphi_r, \Phi\Phi^{\#}\ \varphi_r) = 1,$$

and

$$W_2\psi_r = \frac{1}{\sqrt{\alpha_r}}\Phi^{\#}\Phi\Phi^{\#}\varphi_r = \frac{1}{\sqrt{\alpha_r}}\Phi^{\#}W_1\varphi_r = \sqrt{\alpha_r}\,\Phi^{\#}\varphi_r = \alpha_r\psi_r.$$

Thus ψ_r is a normalized eigenvector of W_2 with eigenvalue α_r.

Furthermore every eigenvector of W_2 is of this form.

It follows from this that $W_2 = \boldsymbol{\Phi}^{\#}\boldsymbol{\Phi}$ has the form $W_2 = \sum_r \alpha_r Q_r$ with $Q_r\,\psi_r = \psi_r$.

If we complete the vectors φ_r and ψ_s to complete orthonormal systems in $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively, we obtain such a system in $\mathcal{G}$ in the form $\varphi_r \otimes \psi_s$.

By substituting the definition of $\varphi_r \otimes \psi_s$ as an antilinear mapping of $\mathscr{H}_2$, into $\mathscr{H}_1$, we find

$$(\Phi, \varphi_r \otimes \psi_s) = \sum_t (\Phi\psi_t, (\varphi_r \otimes \psi_s)\,\psi_t) = (\Phi\psi_s, \varphi_r) = \sqrt{\alpha_r}\,\delta_{rs}.$$

This means $\boldsymbol{\Phi}$ has the development

$$\Phi = \sum_r \sqrt{\alpha_r}\,\varphi_r \otimes \psi_r. \tag{31}$$

We have thus established the following result:

If $\boldsymbol{\Phi}$ is a general vector in $\mathcal{G}$, then there exists an orthonormal system $\{\varphi_r\}$ in $\mathcal{H}_1$, a similar system $\{\psi_s\}$ in $\mathcal{H}_2$, and positive numbers α_r such that

$$\Phi = \sum_r \sqrt{\alpha_r}\, \varphi_r \otimes \psi_r, \qquad \Phi^\# \Phi = \sum_r \alpha_r P_r,$$

$$\Phi\Phi^\# = \sum_s \alpha_s Q_s, \qquad P_r \varphi_r = \varphi_r, \qquad Q_s \psi_s = \psi_s.$$

With this result we have established the *normal form* of the reduction of the pure state $\boldsymbol{\Phi}$ of the compound state to its component states.

A MODEL OF THE MEASURING PROCESS

Now we shall return to physics and complete the analysis of the process of measurement on a model which is so designed that it permits an explicit description of the state of the system and the measuring device during the entire measuring process.

The measuring device is here assumed in the simplest form, so that its quantum character is not obscured by the complexity of a large system.

The purpose of the model is to show the consistency of measurement in the quantum description of the entire system.

In order for the model to represent the essential features of the measurement, it must satisfy the conditions laid down earlier.

The system thus consists of two parts, the system $S_1 = S$ on which a state is to be measured, and a microscopic but classical measuring device $S_2 = m$.

For the system S_1 we take a system represented by a two-dimensional state vector.

Let φ_+ and φ_- be two orthogonal vectors in this space which are the eigenstates of the quantity to be measured.

The system S_2 is assumed to be described by a three-dimensional Hilbert space.

It contains the vector ψ_0 which we shall call the *neutral state* of m and two more orthogonal vectors ψ_+ and ψ_- .

The vector ψ_+ represents the “pointer reading" indicating system S_1 to be in state φ_+.

Similarly ψ_- is the indicator for state φ_-.

If before interaction the system $S + m$ is in the pure state

$$\Phi_{+0} \equiv \varphi_+ \otimes \psi_0,$$

then after the measurement is completed the system $S + m$ is in the pure state

$$\Phi_+ = \mathscr{U}\Phi_{+0} = \varphi_+ \otimes \psi_+, \tag{32}$$

where U is some unitary operator.

Similarly the initial state $\boldsymbol{\Phi}_{-0} = \varphi_{-} \otimes \psi_0$ is, after the interaction, given by

$$\Phi_{-} = \mathscr{U}\Phi_{-0} = \varphi_{-} \otimes \psi_{-}. \tag{33}$$

These formulas describe the characteristic behavior of a measurement of the first kind in the context of this model.

We now take the general initial state of the form

$$\Phi_0 = \alpha_{+}\Phi_{+0} + \alpha_{-}\Phi_{-0}.$$

Since the transformation U is linear, we can obtain the final state after the measurement in this case by linear superposition of the final states in Eqs. (32) and (33):

$$\Phi = \mathscr{U}\Phi_0 = \alpha_{+}\varphi_{+} \otimes \psi_{+} + \alpha_{-}\varphi_{-} \otimes \psi_{-}. \tag{34}$$

After the measurement is completed we can imagine the interaction between S and m removed.

The reading of the scale consists in amplifying the record contained in m and deducing from it the state of S.

The state of m, to be read with the amplifier, is obtained from the pure state in Eq. (34) by reducing that state to the system m.

We use the reduction formulas of the previous discussion.

In this case their application is especially easy since Eq. (34) is already in the normal form of Eq. (31).

The only slight generalization here is that the coefficients $\alpha_{\pm}$ are complex numbers (in Eq. 31 they were real); however, an adjustment of the phases of $\varphi_{\pm}$, $\psi_{\pm}$ will reduce them to reals.

Thus we may write for the reduced states

$$W_1 = |\alpha_+|^2 P_+ + |\alpha_-|^2 P_-,$$
$$W_2 = |\alpha_+|^2 Q_+ + |\alpha_-|^2 Q_-, \tag{35}$$

where P+ and Q+ are the projections containing + , W+ respectively.

We where P+ and Q+ are the projections containing + , W+ respectively.

We see that both states have become mixtures.

Since m (as well as m + A) is a classical system, the state is a classical state.

No further observation on m will modify the state, and the measurement has become an objective record.

According to earlier discussion, each individual system m which may be used in a statistic of the measurement realizes one of the two alternatives.

These alternatives are thus events in the sense of earlier discussion, and their amplification will make them data.

There is no question of any superposition here.

The reduction of the state to the system m has wiped out any phase relations.

Moreover, we have a measurement since the events in *m* and those in S are correlated.

If *m* is in the state ψ_+, then S is necessarily in the state φ_+.

In order to see this, we calculate the expectation value for the cross correlation expressed by the proposition P_+ *and* Q_- and represented by the projection $P_+ \otimes Q_-$.
It is given by

$$(\Phi, P_+ \otimes Q_- \Phi) = 0. \tag{36}$$

Therefore the event ψ_+ is strictly correlated with φ_+ , and ψ_- is similarly correlated with φ_- .

Thus in the reduction formulas we see the true origin of the probability statements in the quantum-mechanical measuring process.

A difficulty seems to appear for this interpretation if we include the system *S* in the measuring device.

In this case there is no occasion to reduce the pure state $\boldsymbol{\Phi}$ to that of a mixture.

If we do this, then the observable projections of the joint system are $\mathbf{\Pi}_+$ with range $\varphi_+ \otimes \psi_+$ and $\mathbf{\Pi}_-$ with range $\varphi_- \otimes \psi_-$ and, in our model, there are no others.

The abelian system S of observables on this system consists of all linear combinations of these two projections.

In this case we are in the situation where the pure state $\boldsymbol{\Phi}$ is contained in a class of equivalent microstates.

These states contain the state

$$W = |\alpha_+|^2\Pi_+ + |\alpha_-|^2\Pi_-, \tag{37}$$

since

$$\mathrm{Tr}\, W\Pi_+ = (\Phi, \Pi_+\Phi) = |\alpha_+|^2$$

and

$$\mathrm{Tr}\, W\Pi_- = (\Phi, \Pi_-\Phi) = |\alpha_-|^2.$$

Thus the final state is again a probability distribution of events, this time described by the projections $\boldsymbol{\Pi}_\pm$.

In this case the probability statement comes in through the theory of equivalent states.

THREE PARADOXES

(a) Schrödinger's cat. In a paper entitled "The Present State of Quantum Mechanics," Schrödinger wrote a criticism of the orthodox view of quantum mechanics.

He pointed out that this view would imply rather grotesque situations for macroscopic events, and he illustrated it with an example involving a cat.

This example has been reformulated by many other authors in more or less equivalent forms, and it has to this day been considered by many an unsolved paradox.

Here we shall give a literal translation of Schrödinger's cat paradox.

Schrödinger writes:

> A cat is placed in a steel chamber, together with the following hellish contraption (which must be protected against direct interference by the cat): In a Geiger counter there is a tiny amount of radioactive substance, so tiny that maybe within an hour one of the atoms decays, but equally probably none of them decays.
>
> If one decays then the counter triggers and via a relay activates a little hammer which breaks a container of cyanide.
>
> If one has left this entire system for an hour, then one would say that the cat is still living if no atom has decayed. The first decay would have poisoned it.
>
> The ψ-function of the entire system would express this by containing equal parts of the living and dead cat.
>
> The typical feature in these cases is that an indeterminacy is transferred from the atomic to the crude macroscopic level, which then can be decided by direct observation.
>
> This prevents us from accepting a "blurred model" so naively as a picture of reality.

By itself it is not at all unclear or contradictory.

There is a difference between a blurred or poorly focussed photograph and a picture of clouds or fog patches.

The paradoxical aspect of this example is to be found in the supposed reduction of the state from a superposition of macroscopically distinct alternatives to one of the events during the act of observation.

(b) Einstein's element of physical reality. Einstein has been not only one of the founders of quantum mechanics, but also one of its strongest critics.

His critique does not concern the existing theory as such, which he recognizes as satisfactory as far as it goes in the description of physical phenomena.

He questions its completeness.

The paradox of Einstein, Podolsky, and Rosen is one of the most striking forms in which this question is expressed.

The authors take the position that physics is concerned with the description of "physical reality" and they affirm that an objective reality exists which does not depend on our observation.

A priori we do not know what it is, so they say, but this precisely is the task of physics: to establish the properties of the existing physical reality.

They are aware that this position requires a meaningful definition of "physical reality."

This is, of course, not easy and it is probably impossible in physical terms alone.

However, certain elements of physical reality can, so they affirm, be given a precise meaning.

Indeed, if the value of a physical quantity for a physical system can be determined with certainty without in any manner whatsoever perturbing the state of the system, then this quantity has for them an element of "physical reality" in that system.

The authors then proceed to construct an example which seems to lead to the conclusion that quantum mechanics is in contradiction with a complete description of all elements of physical reality.

We reproduce this example here in a simplified form.

Let us assume that we have two systems I and Il, which at a given time can interact with each other.

We assume that the states of each system are completely described by a two-dimensional vector space.

Let $\varphi_{\pm}$ represent a complete orthonormal set of vectors in the first space and $\psi_{\pm}$ a similar set in the second space.

Let us further assume that the interaction between the two systems is such that at some time the (pure) state of the joint system is given by

$$\Phi = \frac{1}{\sqrt{2}} [(\varphi_+ \otimes \psi_+) + (\varphi_- \otimes \psi_-)]. \tag{38}$$

We now assume that the two systems can be isolated from each other, for instance by separating them spatially, so that any observation carried out on one of the component systems cannot have any physical effect on the other system.

After this separation the state is still given by Eq. (38).

If we now measure on system I whether it is in the state φ_+ or φ_-, we find that it is in $\varphi_\pm$ with probability 1/2.

The interesting point is that a measurement of $\varphi_\pm$ constitutes at the same time a measurement of $\psi_\pm$ on system II.

Indeed, according to the general theory of the measuring process, we know that whenever a measurement on system I has given the result φ_+, any future measurement on system II will give the result ψ_+.

Since the two systems are physically separated we have a means of determining the state of system I "without in any manner whatsoever perturbing the state" of that system.

According to the criterion of Einstein, Podolsky, and Rosen, the quantity with the eigenstates $\psi_\pm$ of system I must therefore have an element of physical reality.

The value of this quantity is of course not known before the measurement on system I is completed, but that does not invalidate the conclusion that it has a definite value, since one can determine it by a measurement carried out entirely on system I .

Moreover this definite value must have had the same element of reality even *before* the measurement on system I was carried out, since a measurement on system I cannot produce any physical effect whatsoever on system I and thus cannot change the reality of a physical quantity in that system.

We are thus driven to the conclusion that the system (I + Il) is in a mixture of two different states, namely, the states $\varphi_+ \otimes \psi_+$ and $\varphi_- \otimes \psi_-$ mixed, with probabilities 1/2.

But such a state is different from the state expressed by Eq. (38).

Thus the acceptance of the notion of " physical reality" has led us to a contradiction.

This paradox can be given still another form.

It is possible to carry out a simultaneous change of coordinate systems in the vector spaces referring to systems I and II respectively in such a way that the vector remains invariant.

This means we can find other orthonormal pairs $\varphi'_\pm$ and $\psi'_\pm$ such that

$$\Phi = \frac{1}{\sqrt{2}}(\varphi'_+ \otimes \psi'_+ + \varphi'_- \otimes \psi'_-). \qquad (38)'$$

The same reasoning that was applied for the form Eq. (38) can now be repeated identically for the representation Eq. (38)', with the conclusion that system I is in one of the arbitrary states $\psi'_{\pm}$.

But a system cannot be simultaneously in two different states; hence we have encountered another contradiction.

Einstein, Podolsky, and Rosen have drawn the conclusion from this paradox that quantum mechanics does not furnish a complete description of the physical reality of individual systems but merely describes the statistical properties of ensembles of systems.

This paradox was discussed by Bohr, who showed that it could not be considered a refutation of the basic principles of quantum mechanics but that it merely revealed the limits of the traditional concepts of natural philosophy.

In a rejoinder Einstein admits the logical possibility of Bohr's viewpoint, but reaffirms his belief in and preference for another point of view.

(c) Wigner's friend. In 1962 Wigner added a new element to the paradoxes already known by including consciousness for the physical systems involved.

The situation discussed by Wigner is identical with that of formula of Eq. (34), in our earlier discussion, except that Wigner endows system I (the measuring apparatus) with the facility of consciousness.

He then proceeds to introduce the ultimate observer Ω who observes and communicates with the (conscious) apparatus II.

When Ω asks II what he has observed, he will receive the answer that he has observed the state φ_+ (as the case may be) and this with probability $|\alpha_+|^2$. All this is quite satisfactory and in agreement with the theory of measurement.

However, Wigner now inquires what would happen if asked his friend However, Wigner now inquires what would happen if Ω asked his friend (system II): " What did you feel just before I asked you?"

Then the friend will answer, "I told you already I observed φ_+ (or φ_-)," as the case may be.

In other words, the question whether his friend did observe φ_+ or φ_- was already present in his consciousness before asked him.

But at that moment there was no question of any interference of the observer Ω into the natural process of evolution of the two interacting systems; thus its state at that time must have been the superposition of Eq. (34).

But this does not seem to be compatible with the information directly accessible to the conscious friend who is aware of his state before he was asked by the observer what his state is.

For if his awareness is correct, then the state of I + II *before* Ω asked his friend, was already a mixture of the two states $(\varphi_+ \otimes \psi_+)$ and $(\varphi_- \otimes \psi_-)$ and not the superposition of Eq. (34).
Wigner considers this paradox an indication of the influence of consciousness on the physicochemical conditions of living systems.

He finds such an influence entirely in accord with the general principle of action and reaction, since it is known that these physicochemical conditions have in turn a profound influence on conscious sensations.

(d) Discussion of the paradoxes. The similarity between the three paradoxes is obvious.

In all three cases one considers two interacting systems, and the paradox is produced by obtaining some information on the state of one of the systems which seems to be in contradiction with the state obtained from the principle of superposition.

The difference in the three cases refers only to the method of obtaining this information.

In case (a) one appeals to the common-sense notion that a cat is either dead or alive, and that no other state which would leave us undecided about these two alternatives can occur.

In case (b) the information about system II is obtained by looking at the other system I and using the known correlation of observations in I with those in II.

Finally in case (c) we have the "consciousness" of system II which, seemingly without outside interference, is capable of determining the state of II by introspection.

Having thus stressed the similarity, we now pay attention to the differences.

Here one sees at once that (b) stands in a class apart, since in this case only does one obtain information about the system through an outside observer which interacts with the system (I + II).

To be sure, the interaction is assumed to affect only system I and not system II, about which we thus obtain information without outside interference.

Case (b) differs from the other two in another respect.

In cases (a) and (c) one appeals to notions which are outside the confines of physics.

To “be alive" or to " be conscious" are presumably certain states of very complicated physical systems, but it is impossible to express in physical terms what these states are.

In paradox (b), on the other hand, an effort is made to reduce the problem entirely to physical terms.

For this reason it is easier to discuss this case and we shall do it first.

If science is possible then there is nothing paradoxical about the physical world, and insofar as quantum mechanics is a correct physical theory it cannot contain paradoxes.

Thus if paradoxes seem to appear, they must originate either from an inconsistent (and hence incorrect) physical theory, or they must indicate the limitation of concepts in physics which have acquired their meaning outside the domain of physics.

In case (b) we can exclude the second possibility, and so we can discuss this case entirely within existing physical theory, without first having to interpret the physical content of nonphysical concepts.

What does quantum mechanics tell us about the state of the physical systems I, II, and (I + II) after a third observer has carried out a measurement of the quantities $P_{\pm}$ on system I?

We use the notation and the theory from earlier discussions.

This theory tells us that after measurement of the quantity $P_{\pm}$ the system is in the state $W_{\mathrm{I}} \otimes W_{\mathrm{II}}$ where W_{I} and W_{II} are the reductions of the state $W = P_{\phi}$ to the subsystems I and II respectively.

This result is a direct consequence of the earlier analysis we discussed, the only difference being that the system is now (I + II) while the apparatus is the observer Ω.

The effect of the observer Ω on the system (I + II) was thus to change the state W of that system to the state $W_{\mathrm{I}} \otimes W_{\mathrm{II}}$.

This change of the state of the entire system is exactly the same as the change which would have been obtained by measuring the quantity $Q_{\pm}$ of system II.

We see now quite clearly that the attempt at restricting the observation to I is illusory.

The effect on the entire system is exactly the same, whether we observe $P_{\pm}$ in system I or $Q_{\pm}$ in system II.

To be sure, in neither case is the state of subsystem I or subsystem II modified in any manner whatsoever.

This state is before and after the measurement given by W_{I} for I and W_{II} for Il.

The paradox originates in our habit of thinking that the states of two subsystems determine uniquely the state of the composite system.

As we have shown earlier, this is usually not the case.

In the present example the two *different* states $W = P_\phi$ and $W_I \otimes W_{II}$ have the same reductions to the systems I and II and the measurement of either $P_\pm$ or $Q_\pm$ changes the state W of the combined system to the state $(W_I \otimes W_{II})$.

This shows that the application of Einstein's criterion of physical reality becomes ambiguous.

It all depends how we want to interpret the condition "in any manner whatsoever."

If we refer it only to the states of the subsvstems I or II, it is obviously fulfilled; if we refer it to the entire system (I + II), it is not.

In no case is there a contradiction of the uncertainty relation, because, as we have seen, a measurement of $P_\pm$ has exactly the same effect on the states as a measurement of $Q_\pm$.

Thus the "paradox" of Einstein, Podolsky, and Rosen does not reveal any contradiction of quantum mechanics; it merely emphasizes in a most striking way the essential nonclassical consequences of the quantum-mechanical superposition of states.

It is this very superposition principle which leads to the ambiguity in the application of Einstein's criterion of "physical reality.

Let us now turn to the more difficult discussion of paradoxes (a) and (c).

The difficulty in these cases stems from the fact that the outside observer is pushed into the background.

In case (a) he may merely be needed to verify whether the cat is dead or alive, an observation which may reasonably be assumed to have no effect whatsoever on the biological state of the cat.

In case (c) he is even entirely superfluous since consciousness becomes aware of itself by introspection.

Of course this faculty of observing the state of II without any observer is obtained here with properties which are difficult to express in physical terms, namely, "being alive" in case (a) and "being conscious" in case (c).

In either case the alternatives of the microscopic system are transferred to the crude macroscopic level and thus are no longer subject to the quantum-mechanical ambiguities associated with coherent interferences of two different states.

One might reformulate the Schrödinger cat paradox by using only the macroscopic features of system II.

Such a reformulation has been given, for instance, by Einstein, who replaced the "hellish contraption" of Schrödinger by a moving film strip which records the event of the radioactive decay in a permanent, macroscopic, and unobserved record.

In this form the paradox is formulated entirely within the confines of physics, and yet at first sight it seems to retain its paradoxical character.

The essential point here is that system II, which contains this recording device, can be made as large as one wishes.

In a subsequent observation on this system, the inevitable interaction of the outside observer with system I can therefore be made as small as one wishes, and thus (one is tempted to conclude) it can be neglected altogether.

It would, however, be incorrect to neglect it altogether. For we must not forget that the distinction between state W and state $W_{\mathrm{I}} \otimes W_{\mathrm{II}}$ becomes increasingly difficult to detect with the increase in size of the whole apparatus, and it is precisely this distinction which is under discussion here.

The theory of the preceding two sections has shown with sufficient generality that the remaining interaction between I and an outside observer Ω is in fact the essential effect which will indeed obliterate the distinction between the two states.

If the information of the recording device does have objective validity, that is, if it can be communicated, then this very property makes it impossible to distinguish between the two states W and $W_{\mathrm{I}} \otimes W_{\mathrm{II}}$

Thus the paradox of Schrödinger's cat can be resolved when it is reformulated entirely in physical terms.

Wigner's friend could be treated in the same manner with the same conclusion, but this would not meet the heart of Wigner's problem.

As long as one insists on including consciousness as a property of quantum-mechanical systems, the outsider observer Ω can be dispensed with altogether and then we have no answer to the paradox.

Must we conclude from this, as Wigner does, that quantum mechanics, as we know it now, would be inapplicable for systems with consciousness?

The answer to such a question obviously presupposes a characterization and analysis of "consciousness" in physical terms, a task which seems to transcend the present limitations of physics.