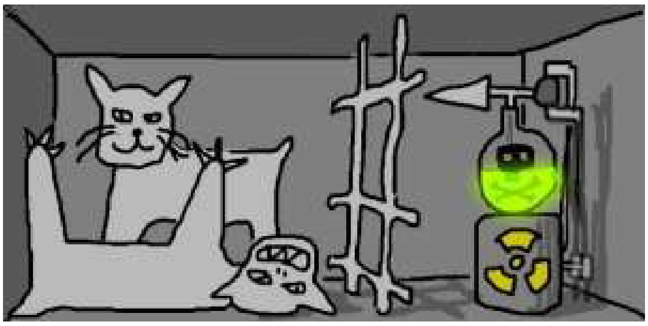
Formale Grundlagen und Messprozess in der Quantenmechanik

Feng Xu

Institut für Theoretische Physik
Universität Münster

Seminar zur Theorie der Teilchen und Felder (Heitger, Klasen, Münster),
April 2011



When I hear of Schrödinger's cat, I reach for my gun.

Stephen Hawking

Contents

1 The Mathematical Formalism of Quantum Mechanics

The main purpose of this presentation is to introduce the conceptual problem of the quantum mechanical description of the measuring process. Before going into a detailed discussion about that, let's first review the mathematical structure of quantum mechanics, its description of states, and its calculation of the probabilities of the outcomes of measurements.

1.1 Hilbert Space

As in classical physics we postulate the existence of *isolated systems* in quantum mechanics. In both theories, if a complete description of an isolated system is given at one time, a complete description for any other time is determined, as long as the system is not influenced by any other system. The means of description of the state of the system have drastically changed over the course of the history of physics. In the original Newtonian mechanics, the state of the system was described by the positions and velocities of its constituents. In the later field theories, it was described by the field strengths at all points of space. The usual quantum mechanical description of the state of a system is given by a vector (or more accurately by a *ray*) in an abstract complex Hilbert space. This vector is called the *state vector*. The state can be specified by an infinite set of complex numbers, a_1, a_2, a_3, $\ldots$, the components of the state vector in Hilbert space satisfy the condition that the sum

$$\sum |a_i|^2 < \infty. \tag{1.1}$$

This means that the *length* of the vectors in Hilbert space is finite. The set of vectors, the components of which differ by a common factor, form a *ray*. We usually choose one of these normalized vectors

$$\sum |a_i|^2 = 1 \tag{1.2}$$

to describe the state.

Schrödinger's original formulation of his wave mechanics was, of course, not in terms of the Hilbert space. Wave mechanics characterized the states of systems in terms of complex valued functions. However, if we introduce an orthonormal set of functions $u_n(x_1, x_2 \ldots)$

$$\int \cdots \int u_n(x_1, x_2 \ldots)^* u_m(x_1, x_2 \ldots) dx_1 dx_2 \cdots = \delta_{nm}, \tag{1.3}$$

we can expand the wave function ψ of any state in terms of this set. Thus,

$$\psi(x_1, x_2 \ldots) = \sum a_n u_n(x_1, x_2 \ldots); \tag{1.4}$$

and the number a_n can be considered to be the components of a vector in Hilbert space as long as

$$\int \cdots \int |\psi(x_1, x_2 \ldots)|^2 dx_1 dx_2 \cdots = \sum |a_i|^2 \tag{1.5}$$

is finite, which we assume to be the case for Schrödinger's wave functions.

The most important derivative concept in Hilbert space is that of the scalar product. The scalar product of two vectors a and b is defined as

$$(a, b) = \sum a_n^* b_n. \tag{1.6}$$

In terms of the wave functions this is

$$(\psi, \phi) = \int \cdots \int \psi(x_1, x_2 \ldots)^* \phi(x_1, x_2 \ldots) dx_1 dx_2 \cdots . \tag{1.7}$$

1.2 Linear Operators in Hilbert Space

An operator in Hilbert space transforms a vector in Hilbert space into another vector in the same space. For linear operators A we write $A\phi$ for the vector into which the operator A transforms the vector ϕ. An operator A is called linear if for any two vectors ϕ, ψ and any two numbers α, β

$$A(\alpha\phi + \beta\psi) = \alpha A\phi + \beta A\psi \tag{1.8}$$

holds.

A special kind of linear operators of basic significance are the self-adjoint operators. In order to define them, we first define the adjoint $A^\dagger$ to an operator A by

$$(\phi, A\psi) = (A^\dagger\phi, \psi). \tag{1.9}$$

And the self-adjoint operators are those which are equal to their adjoints,

$$A = A^\dagger. \tag{1.10}$$

For sake of accuracy it should be point out that the existence of the adjoint $A^\dagger$ to an arbitrary linear operator is by no means obvious.

1.3 Normal Form of Self-Adjoint Operators

Now let's consider the bounded self-adjoint operators, i.e., operators such that $(A\psi, A\psi)$ has an upper limit. The usual procedure to obtain the normal form is to look for the characteristic vectors ψ_ν of A, with

$$A\psi_\nu = \lambda_\nu \psi_\nu. \tag{1.11}$$

We can easily prove that the characteristic values λ_ν have to be real and the characteristic vectors ψ_ν, which are solutions for different λ_ν, are orthogonal. If the characteristic vectors ψ_ν form a complete orthonormal set, we say that A has only a *point spectrum*, this consisting of the λ_ν. The effect of A on an arbitrary vector ϕ

$$\phi = \sum (\psi_\nu, \phi)\psi_\nu \tag{1.12}$$

is then, because of its linear character,

$$\begin{aligned} A\phi = A\sum(\psi_\nu, \phi)\psi_\nu &= \sum(\psi_\nu, \phi)A\psi_\nu \\ &= \sum(\psi_\nu, \phi)\lambda_\nu\psi_\nu. \end{aligned} \tag{1.13}$$

The measurement theory then postulates that the measurement of A on a system in the state ϕ gives one of the values λ_ν with the probability

$$p_\nu = |(\psi_\nu, \phi)|^2. \tag{1.14}$$

The preceding remarks apply to a self-adjoint operator which has a point spectrum only. One can well say that most operators do not have such a spectrum. A self-adjoint operator may also have a continuous spectrum. In the general case the preceding equations have to be replaced by much more complicated ones. In order to illustrate this point, it is helpful to rewrite (1.12), (1.13) and (1.14) by decomposing A into projection operators P_ν. A projection operator is defined by

$$P_\nu\phi = (\psi_\nu, \phi)\psi_\nu. \tag{1.15}$$

It then follow from (1.12) that

$$\sum P_\nu = 1, \tag{1.16}$$

and from (1.13),

$$\sum \lambda_\nu P_\nu = A. \tag{1.17}$$

we can easily verify that the P_ν are self-adjoint, identical with their squares, and more generally satisfy the equation

$$P_\nu P_\mu = \delta_{\nu\mu} P_\nu. \tag{1.18}$$

The expression for the transition probability that the outcome of the measurement of A on ϕ will be λ_ν, becomes

$$\begin{aligned} p_\nu &= |(\psi_\nu, \phi)|^2 = (\psi_\nu, \phi)^*(\psi_\nu, \phi) \\ &= [(\psi_\nu, \phi)\psi_\nu, \phi] = (P_\nu \phi, \phi) \\ &= (\phi, P_\nu \phi). \end{aligned} \tag{1.19}$$

We now proceed to the more general case that A may also have a continuous spectrum. In this case one has to admit that the measurement will not yield a mathematically precise value. It is more reasonable to ask, for instance, whether the outcome is smaller than a number λ, and the probability for that can be written, in the case of a discrete spectrum, as

$$p(\lambda) = [\phi, P(\lambda)\phi], \tag{1.20}$$

with

$$P(\lambda) = \sum_{\lambda_\nu < \lambda} P_\nu. \tag{1.21}$$

It then follows from (1.20) that the probability for an outcome of the measurement between λ' and $\lambda > \lambda'$ is

$$p(\lambda) - p(\lambda') = \{\phi, [P(\lambda) - P(\lambda')]\phi\}. \tag{1.22}$$

It follows from the theory of self-adjoint operators in Hilbert space that the operators $P(\lambda)$ can also be defined if the spectrum is not exclusively discrete. In other words, the $P(\lambda)$ can be defined for every self-adjoint operator, for any real λ. We clearly have

$$P(\lambda) \to 0 \quad \text{for} \quad \lambda \to -\infty; \tag{1.23}$$

and the analogs of (1.16) and (1.18) are

$$P(\lambda) \to 1 \quad \text{for} \quad \lambda \to \infty \tag{1.24}$$

and

$$P(\lambda)P(\lambda') = P(\lambda')P(\lambda) = P(\lambda') \quad \text{for} \quad \lambda' < \lambda. \tag{1.25}$$

These equations, of course, do not determine the $P(\lambda)$, one has to write the analog of (1.17). This becomes a limit of a sum with increasing λ values, which cover the real spectrum with increasing density. If the density is $1/N$, we have the sum

$$A = \sum_{n=-\infty}^{\infty} \sum_{m=1}^{N} \left(n + \frac{m}{N}\right) \left[P\left(n + \frac{m}{N}\right) - P\left(n + \frac{m-1}{N}\right)\right], \qquad (1.26)$$

and this is valid in the limit $N = \infty$. This is, of course, a special form for A in which the distance between successive λ values is uniformly $1/N$. There are infinitely many other ways to increase the density of the λ. The mathematician writes a Stieltjes integral for (1.26) instead. The point is that (1.26), together with (1.23), (1.24) and (1.25), fully defines the λ and hence determines the mathematical expression (1.22) for the probability of the outcome of the measurement of A lying between λ' and λ.

2 The Problem of Measurement

Now before we are finally going to discuss about the measurement process in the second part, I'd like to introduce the concept of the direct product first, which is an extension of the mathematical formalism in the first part, and since it greatly facilitates the quantum mechanical description of the measurement process, we may need to have a look at it.

2.1 The Direct Product of Hilbert spaces

Starting from two Hilbert spaces, we can construct a single larger Hilbert space. If the axes of the first Hilbert space are specified by indices ν, and the second by n, then νn specifies a single axis of the Hilbert space which is the direct product of the original two Hilbert spaces. The scalar product of two vectors Ψ and Φ in the space of the direct product is

$$(\Psi, \Phi) = \sum_{\nu n} \Psi^*_{\nu n} \Phi_{\nu n}. \qquad (2.1)$$

The Hilbert space which is the direct product of two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ is usually denoted by $\mathcal{H}_1 \otimes \mathcal{H}_2$.

We also define the direct product of two vectors $\phi \in \mathcal{H}_1$ and $\psi \in \mathcal{H}_2$ by $\phi \otimes \psi$. Its νn component is

$$(\phi \otimes \psi)_{\nu n} = \phi_\nu \psi_n. \qquad (2.2)$$

The direct product is linear in its two factors,

$$(\alpha\phi + \alpha'\phi') \otimes \psi = \alpha(\phi \otimes \psi) + \alpha'(\phi' \otimes \psi). \tag{2.3}$$

The square length of $\phi \otimes \psi$ is

$$\begin{aligned}(\phi \otimes \psi, \phi \otimes \psi) &= \sum_{\nu n} |(\phi \otimes \psi)_{\nu n}|^2 = \sum_{\nu n} |\phi_\nu \psi_n|^2 \\ &= \sum_{\nu} |\phi_\nu|^2 \sum_n |\psi_n|^2, \end{aligned} \tag{2.4}$$

and is the product of the squares of the lengths of the two factors ϕ and ψ. The scalar product of two direct product vectors $\psi \otimes \phi$ and $\psi' \otimes \phi'$ in the new Hilbert space becomes

$$\begin{aligned}(\psi \otimes \phi, \psi' \otimes \phi') &= \sum_{\nu n} (\psi \otimes \phi)^*_{\nu n} (\psi' \otimes \phi')_{\nu n} = \sum_{\nu n} (\psi_\nu \phi_n)^* \psi'_\nu \phi'_n \\ &= \sum_\nu \psi^*_\nu \psi'_\nu \sum_n \phi^*_n \phi'_n = (\psi, \psi')(\phi, \phi'), \end{aligned} \tag{2.5}$$

the product of the scalar products of the two factors.

It is important to remark that the direct product of two factors ϕ and ψ is the same vector in the new Hilbert space no matter which coordinate systems are used for its definition in (2.2).

The direct product of two Hilbert spaces, and the direct product of vectors in them, is introduced in order to describe the joining of two systems into a single system. This is important if one wants to describe the interaction of two systems which were originally separated. If it is possible to describe the two systems in separate Hilbert spaces, their union can indeed be most easily characterized in the direct product of these Hilbert spaces. The probability that the two systems are in state ϕ' and ψ' when their actual state vectors are ϕ and ψ is

$$|(\psi' \otimes \phi', \psi \otimes \phi)|^2 = |(\psi', \psi)|^2 |(\phi', \phi)|^2, \tag{2.6}$$

which is the product of the transition probabilities from ϕ into ϕ' and from ψ into ψ'. Since two system are assumed to be independent, this is the expected result.

2.2 Direct Product of Operators

In order to obtain the more general expressions corresponding to (1.15), etc. for the joint system, it is useful to introduce the concept of the direct product of two operators A and B, acting in the two original Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. The action of A is described by the equation

$$(A\psi)_\nu = \sum_{\nu'} A_{\nu'\nu}\psi_{\nu'}; \tag{2.7}$$

that of B by

$$(B\phi)_n = \sum_{n'} B_{n'n}\phi_{n'}. \tag{2.8}$$

Their direct product, to be denoted by $A \otimes B$, will then transform Ψ with the components $\Psi_{\nu n}$ into $(A \otimes B)\Psi$, the components of which are

$$[(A \otimes B)\Psi]_{\nu n} = \sum_{\nu' n'} A_{\nu'\nu}B_{n'n}\Psi_{\nu' n'}. \tag{2.9}$$

This is the definition of $A \otimes B$. Clearly, if $\Psi = \phi \otimes \psi$, the action of $A \otimes B$ on this will give

$$(A \otimes B)(\psi \otimes \phi) = A\psi \otimes B\phi, \tag{2.10}$$

the direct product of the results of the actions of A and of B in their respective Hilbert spaces.

The concept of the direct product of two operators enables us to generalize the expression for the transition probability (2.6) as the similar expression (1.15) for a single system was generalized. First, if both operators A and B, to be measured on the two systems, have only a discrete spectrum, the projection operators which correspond to the characteristic values a_ν of A and b_n of B are denoted with P_ν and Q_n respectively. The probability that the outcome is a_ν for the measurement of A on ϕ is then $(\phi, P_\nu\phi)$. The probability that the outcome is b_n for the measurement of B on ψ is $(\psi, Q_n\psi)$. The probability that the two outcomes on the joint system be a_ν and b_n is then given by

$$P_{\nu n} = [\phi \otimes \psi, (P_\nu \otimes Q_n)(\phi \otimes \psi)] = (\phi, P_\nu\phi)(\psi, Q_n\psi), \tag{2.11}$$

which follows from (2.10). We conclude that $P_\nu \otimes Q_n$ is the projection operator for the outcome a_ν of A and b_n of B. Similarly, we can generalize (1.20), the probability for A giving a result smaller than λ and B giving a result smaller than l is

$$p(\lambda, l) = \{\phi \otimes \psi, [P(\lambda) \otimes Q(l)](\phi \otimes \psi)\} = [\phi, P(\lambda)\phi][\psi, Q(l)\psi]. \tag{2.12}$$

The formula for the probability that the outcomes will fall in the intervals (λ', λ) and (l', l), respectively, is obtained as

$$\begin{aligned}&[p(\lambda, l) - p(\lambda', l)] - [p(\lambda, l') - p(\lambda', l')] \\ &= \{\phi \otimes \psi, [P(\lambda) - P(\lambda')][Q(l) - Q(l')](\phi \otimes \psi)\},\end{aligned} \tag{2.13}$$

in generalization of (1.22).

2.3 The Orthodox View

As we are about to discuss the measurement process, it's useful to review the standard view of the conceptual foundations of quantum mechanics in the late 1920s, what we shall call the orthodox view.

The possible states of a system can be characterized, according to quantum mechanical theory, by state vectors. These state vectors change in two ways. As a result of the passage of time, they can change continuously, according to Schrödinger's time-dependent equation, which will be also called the quantum mechanical equation of motion. The state vector can also change discontinuously, according to the probability law, if a measurement is carried out on the system. This second type of change is often called the reduction of the state vector, which is unacceptable to many physicists.

The simplest way that we may try to reduce the two kinds of changes of the state vector to a single kind is to describe the whole process of measurement as an event in time, governed by the quantum mechanical equation of motion. Unfortunately, the situation will turn out to be not this simple.

Let's first consider a measurement on an object which is in a state, the state vector σ_k of which is a characteristic vector of the quantity to be measured. If we denote the initial state of the apparatus by a_0, then the initial state of apparatus plus object is $a_0 \otimes \sigma_k$. After the measurement, the apparatus has assumed a state which shows the outcome of the measurement; we denote its state vector by a_k. The object did not change its state as a result of the measurement, the interaction between apparatus and object transforms $a_0 \otimes \sigma_k$ into

$$a_0 \otimes \sigma_k \rightarrow a_k \otimes \sigma_k, \tag{2.14}$$

and the measurement surely yields one definite value.

It now follows from the linear nature of the quantum-mechanical equation of motion that if the initial state of the object is a linear combination of the σ_k, say $\sum \alpha_k \sigma_k$, the final state of object plus apparatus will be given by

$$a_0 \otimes \sum \alpha_k \sigma_k = \sum \alpha_k (a_0 \otimes \sigma_k) \rightarrow \sum \alpha_k (a_k \otimes \sigma_k). \tag{2.15}$$

Naturally, there is no statistical element in this result, as there cannot be. However, what the final state of (2.15) does show is that a statistical correlation between the state of the apparatus and that of the object has been established. The problem of measurement on the object is transformed into the problem of an observation on the apparatus. If we measure the state of the apparatus a by a second apparatus b, and the state a_k of a definitely put the apparatus b into the state b_k, in other words, if

$$b_0 \otimes a_k \to b_k \otimes a_k, \tag{2.16}$$

then the interaction of b_0 with the result of the measurement on the general state will give

$$b_0 \otimes \sum_k \alpha_k (a_k \otimes \sigma_k) \to \sum_k \alpha_k (b_k \otimes a_k \otimes \sigma_k). \tag{2.17}$$

Thus, a correlation between the states of all three systems (object, apparatus a and b) is established. A similar statement applies if a fourth apparatus is used to measure the state of b, and so on. However, the fundamental point remains unchanged and a full description of an observation must remain impossible, since the quantum mechanical equations of motion are deterministic and contain no statistical element, whereas the measurement does.

2.4 Critiques of the Orthodox Theory

There are attempts to modify the orthodox theory of measurement as illustrated in the picture epitomized by (2.14) and (2.15). The only attempt which will be discussed here presuppose that the result of the measurement is not a state vector, but a so-called mixture of the state vectors

$$a_\nu \otimes \sigma_\nu, \tag{2.18}$$

and *one* of the particular states will emerge from the interaction between the object and apparatus with the probability $|\alpha_\nu|^2$. If this were so, the state of the system would not be changed when we ascertain which of the state vectors (2.18) corresponds to the actual state of the system. In other words, the final observation only increases our knowledge of the system; it does not change anything. This is not true if the state vector, after the interaction between the object and apparatus, is given by (2.15), because the state represented by the vector (2.15) has properties which neither of the states (2.18) has. It may be worthwhile to illustrate this point by the example Stern-Gerlach experiment.

In the Stern-Gerlach experiment the projection of the spin of an incident beam of particles is measured. The index ν of the state of the spin has two values in this case, they correspond to the two possible orientations of the spin. The apparatus is that positional coordinate of the particle. The experiment illustrates the statistical correlation between the state of the apparatus (the position coordinate) and the state of the object (spin orientation). The ordinary use of the experiment is to obtain the spin orientation, by observing the position, i.e., the location of the beam. But now let's make a thought experiment after the measurement. If the two beams are brought together and interfere, the initial state of the spin will be produced again. However, if the two beams do not interfere, they build a mixture, as the mixture passes through a second Stern-Gelach-apparatus, both beams in the mixture will show equal probabilities to assume its initial and the opposite spin directions. Clearly we cannot reproduce the initial state of the spin through the mixture.

In general we can establish the following relation between an arbitrary state and a mixture. For an arbitrary state

$$\phi = \sum_k \alpha_k (a_k \otimes \sigma_k) = |\phi\rangle = \sum_k \alpha_k |a_k, \sigma_k\rangle, \tag{2.19}$$

its statistical operator contains the mixture and a interference term

$$\begin{aligned} \rho_\phi &= \sum_{k,l} \alpha_l^* \alpha_k |a_l, \sigma_l\rangle\langle a_k, \sigma_k| \\ &= \underbrace{\sum_k |\alpha_k|^2 |a_k, \sigma_k\rangle\langle a_k, \sigma_k|}_{\text{mixture}} + \underbrace{\sum_{k \neq l} \alpha_l^* \alpha_k |a_l, \sigma_l\rangle\langle a_k, \sigma_k|}_{\text{interference term}} . \end{aligned} \tag{2.20}$$

The interference term of a macroscopic object is very small, and can be hardly observed. Besides in reality the whole object-plus-apparatus system is not isolated. The interaction with the environment, despite being weak, destroys the interference term. Therefore the thought experiment above to bring the two beams into interference is actually not realistic. However, the thought experiment illustrates the point that the state of the system, object-plus-apparatus, shows characteristics which neither of the separated beams alone would have.

References

[1] G. Münster, *Quantentheorie*, de Gruyter (2010)

[2] E. P. Wigner, *Interpretation of Quantum Mechanics*, Princeton University Press (1983)

[3] E. P. Wigner, *American Journal of Physics, 31*, 6-15 (1963)