Cambridge University Press
978-0-521-11369-4 - Quantum Engineering: Theory and Design of Quantum Coherent Structures
A. M. Zagoskin
Excerpt
More information

# 1

# Quantum mechanics for quantum engineers

> Shall I refuse my dinner because I do not fully understand the process of digestion? No, not if I am satisfied with the result.
>
> O. Heaviside, *Electromagnetic Theory*, vol. 2, 1899

## 1.1 Basic notions of quantum mechanics

### *1.1.1 Quantum axioms*

Let us start with a brief recapitulation of quantum mechanics on the "how to" level. According to the standard lore, the instantaneous state of any quantum system (that is, *everything* that can be known about it at a given moment of time) is given by its wave function (state vector)[1] – a complex-valued vector in some abstract Hilbert space; the nature of this space is determined by the system. All the *observables* (i.e., physical quantities defined for the system and determined by its state – e.g., the position or momentum of a free particle, the energy of an oscillator) are described by Hermitian operators defined in the same Hilbert space. All three elements – the Hilbert space, the state vector, and the set of observables – are necessary to describe the outcome of any experiment one could perform with the system. Since humans cannot directly observe the behaviour of quantum objects, these outcomes are also called *measurements*, being the result of using some classical apparatus in order to translate the state of a quantum system into the state of the apparatus, which can then be read out by the experimentalist. The classical (i.e., non-quantum) nature of the apparatus is essential, as we shall see in a moment.

In addition, we need to know how the state of the system changes in time, and how it determines the measured values of the observables. All of the above can be presented as four textbook "axioms of quantum theory":

1) The state of a quantum system at time $t$ is described by a normalized[2] vector $|\psi(t)\rangle$ belonging to the Hilbert space $\mathcal{H}$, which is specific for the system in question.

---

[1] A very important generalization of the wave function, the density matrix (statistical operator), will be discussed in Section 1.2.

[2] This is not strictly necessary, but makes the explanations shorter without much loss of generality.

2) [Schrödinger equation] The state evolves with time according to $\imath\hbar|\dot{\psi(t)}\rangle = H|\psi(t)\rangle$, where $H$, the Hamiltonian, is a Hermitian operator associated with the energy of the system.
3) [Collapse of the wave function] The measured value of an observable is always one of the eigenvalues of its operator $A$, $a_j$; whatever the state of the system was *before* the measurement, immediately *after* it the state vector of the system is the corresponding normalized eigenvector of $A$, $|a_j\rangle$ ($A|a_j\rangle \equiv a_j|a_j\rangle$).
4) [Born's rule] The probability of measuring a particular eigenvalue, $a_j$, of the observable $A$ (collapsing the wave function $|\psi(t)\rangle$ into a given eigenvector $|a_j\rangle$) is given by the square modulus of the former's projection on the latter, $p_j(t) = \left|\langle a_j|\psi(t)\rangle\right|^2$.

The most striking features of these axioms are the special roles played by time and energy (Hamiltonian) of all the other physical quantities associated with the quantum system, and the jarring difference between the linear, reversible, unitary quantum evolution determined by the Schrödinger equation, and the nonlinear, irreversible, non-unitary measurement process. Indeed, the equation

$$\imath\hbar|\dot{\psi(t)}\rangle = H|\psi(t)\rangle \tag{1.1}$$

has the solution (in the sufficiently general case of a time-independent Hamiltonian)

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle;\ U(t) = \mathrm{e}^{-\frac{\imath}{\hbar}Ht}, \tag{1.2}$$

and for Hermitian $H$ the *evolution operator* is unitary, $U(t)^\dagger U(t) = U(t)U(t)^\dagger = I$ (where $I$ is the identity operator in the Hilbert space of the system). The latter property ensures that the normalization of the state vector is preserved. On the other hand, after the measurement at moment $t$ we find the system in a state $|a_j\rangle$, which is not related to the state $|\psi(t)\rangle$ before the measurement by any reversible transformation ($\psi(t)\rangle$ instantaneously collapses to $|a_j\rangle$). The after-collapse state $|a_j\rangle$ could be obtained from *any* state $|\psi(t)\rangle$ as long as $|a_j\rangle$ and $|\psi(t)\rangle$ are not mutually orthogonal. Taken together, axioms (3) and (4) are often called the *projection postulate*: the measurement of the observable $A$ *projects* the state vector $|\psi(t)\rangle$ on the eigenvectors of $A$; the square modulus of the projection gives the probability with which the system is likely to be found in the corresponding eigenstate of $A$. The eigenvectors of a Hermitian operator $A$ can be chosen to form a complete orthonormal basis of the Hilbert space; therefore the sum of squared projections of the state vector $|\psi\rangle$ on the vectors of this basis is, by Parseval's equality, the square of the norm of $|\psi\rangle$,

$$||\psi(t)\rangle||^2 = \langle\psi(t)|\psi(t)\rangle = \sum_j |\langle a_j|\psi(t)\rangle|^2 = \sum_j p_j. \tag{1.3}$$

Once we have chosen the state vector such that it is normalized to unity, Eq. (1.3) ensures that the probabilities of finding different outcomes of the measurement of the observable $A$ sum to one, as they should. The unitarity of the evolution given by Eq. (1.2) ensures that the probabilities "don't leak".

### *1.1.2 Quantum–classical boundary: the Schrödinger's cat paradox*

This picture does lead to serious questions. First of all, time is *very* different from spatial coordinates: it is not an observable, but a parameter, which governs the evolution of the state vector between the instantaneous "collapses". Fortunately, we do not have to deal with it, since our problems are strictly non relativistic. More pertinent is the question of the nature of "measurement" and "collapse" and their presumed instantaneity.

Measurement is understood as an interaction of the quantum system with a macroscopic object ("apparatus") such that the final state of the latter is determined by the value of the observable $a_j$, and different states of the apparatus are distinguishable and immutable (i.e., can be observed without perturbation). The terminology originates from the early days of quantum mechanics. Of course, there is no need for somebody to actually set up the apparatus; any appropriate macroscopic system will do, and now the tradition forces us to talk about, e.g., the liquid "measuring" or "observing" the quantum state of a particle travelling through it.

The classical states of the apparatus correlated to the different outcomes of the measurement are called "pointer states". (The corresponding eigenstates of the observable that is being measured we will also call pointer states, where it does not lead to confusion.) In the Copenhagen interpretation, it is the apparatus that predetermines what observables can be measured (so called "complementarity with respect to the means of observation"). In principle, for any classical variable we should be able to design an apparatus which would measure its quantum counterpart, an observable.

One problem here is that there exists no well-defined boundary between the "measured", or "observed" microscopic system, and the macroscopic "apparatus" or "observer". The Copenhagen interpretation of quantum mechanics simply posits the quantum behaviour for the one, and classical for the other, which is somewhat circular. What is worse, it does not allow a description of the system and the apparatus within a single formalism, denying us any quantitative description of the *process* of measurement. This was not too troubling when dealing with (quantum) electrons going through double slits in a (classical) screen, since it was obvious which was which, and the time of the electron's interaction with the detector can be neglected compared to all other relevant timescales. It becomes of crucial importance when the quantum systems we deal with contain huge numbers of elementary particles, almost comparable with the number of particles in our "apparatus", or are being "measured" continually. How do we describe such a situation in a consistent way?

A natural thing to do is to directly extend the quantum description (as the more fundamental one) to macroscopic systems. A difficulty of this approach lies in the fact that only very special states can be measured by a macroscopic apparatus, while at the quantum level they seem to be no different from all the rest of them. The situation is highlighted by the famous "Schrödinger's cat" paradox. Let us put together a deadly contraption: a tank of poisonous gas with an electrically controlled valve, connected to a Geiger counter. Put it in a sealed container with a live cat and a radioactive atom (e.g., $^{210}$Po, which decays into the stable lead-206, $^{206}$Pb, with the half-life $T \approx 138$ days), and wait (Fig. 1.1).

If we had to deal just with the radioactive atom, the description would be simple. Its wave function can be written as[3]

$$|\psi(t)\rangle = 2^{-t/2T}\left|{}^{210}\mathrm{Po}\right\rangle + \sqrt{1-2^{-t/T}}\left|{}^{206}\mathrm{Pb} + \alpha\ \text{particle}\right\rangle. \tag{1.4}$$

If at any given time we measure (observe) this atom, with the probability $1-2^{-t/T}$ we'll find lead-206. But putting the atom in the box with a cat makes things look bizarre, once we try to describe the cat (and the rest of the contraption) by a wave

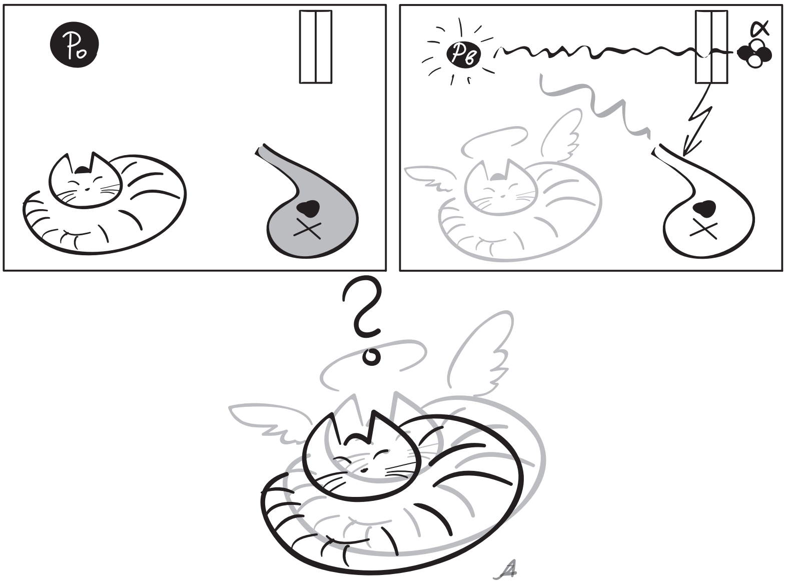

Fig. 1.1. Schrödinger's cat paradox: Should we describe macroscopic systems by wave functions, and if not, why not?

---

[3] This expression is only valid for not too short and not too long times; see § 5.5.6.

function. Now the state of the system before the observation will be

$$|\psi(t)\rangle = 2^{-t/2T}\left|{}^{210}\mathrm{Po} + \text{live cat}\right\rangle + \sqrt{1-2^{-t/T}}\left|{}^{206}\mathrm{Pb} + \alpha \text{ particle} + \text{dead cat}\right\rangle. \tag{1.5}$$

We don't doubt that after opening the box one will find either a live or a dead cat, with the probabilities $2^{-t/T}$ and $1-2^{-t/T}$ respectively. This intuitively clear outcome is in no way trivial. On the quantum side, we can always introduce an observable, which is a linear combination of other observables with real coefficients – there is nothing in the formalism to prohibit it. For example, we can build an observable, which will have as one of its eigenstates the superposition given by Eq. (1.5). Nevertheless, we evidently cannot build a classical apparatus, which would measure such a "zombie" state of the poor animal. The set of admissible classical states thus must impose some preferred set of bases on the Hilbert state of our quantum system.

Is the description (1.5) of a hybrid micro/macro system justified at all? And, what does precipitate the transition from the "live–dead" superposition to the "either–or" classical picture? The proposed answers run the full philosophical gamut from the many-worlds interpretation to the key role of the observer's conscience in the collapse. Instead of plunging into these fascinating ontological and epistemological depths, we will take a pragmatic approach and see how far it will take us.

We will not make any special difference between the "measurement" or "observation" and the interaction of our quantum system with its macroscopic surroundings.

Fig. 1.2. The canonical view of the quantum–classical transition.

Cambridge University Press
978-0-521-11369-4 - Quantum Engineering: Theory and Design of Quantum Coherent Structures
A. M. Zagoskin
Excerpt
More information

Indeed, explaining the "collapse" through the action of conscience would be exorcizing Asmodeus with Beelzebub's help, since we currently understand the latter, if anything, less than the former.

We will assume that on the fundamental level every system is governed (and described) by the laws of quantum mechanics, of which classical mechanics is a limiting case – in full agreement with Bohr's *correspondence principle*. The difference between "quantum" and "classical" systems must naturally emerge from the formalism.[4]

## 1.2 Density matrix formalism

### 1.2.1 Justification and properties

The notion of the state vector is inadequate for our purposes, but it can be generalized to the *density matrix*, or *statistical operator*, first introduced independently by Landau and von Neumann in 1927. To make the motivation clear, let us start from a quantum system, which consists of two subsystems, A and B, which we can measure independently. Suppose, for example, that we have two particles with spin $\hbar/2$, and have an apparatus, which can measure the $z$-component of spin of either particle. Let the system be in state

$$|\Psi\rangle_{AB} = \sum_{j,k=\uparrow,\downarrow} C_{jk}|j_A k_B\rangle, \tag{1.6}$$

which is obviously the most general form of such two-particle wave functions. Now measure the spin of particle B. The corresponding observable is the operator $(\hbar/2)\sigma_z$, such that $(\hbar/2)\sigma_z|\uparrow\rangle = \hbar/2|\uparrow\rangle$, and $(\hbar/2)\sigma_z|\downarrow\rangle = -\hbar/2|\downarrow\rangle$. The result of the measurement can then be written as

$$|\Psi\rangle_{AB} \rightarrow \begin{cases} \text{spin B} = \hbar/2, |\uparrow_A\rangle, \text{probability}|C_{\uparrow\uparrow}|^2; \\ \text{spin B} = \hbar/2, |\downarrow_A\rangle, \text{probability}|C_{\downarrow\uparrow}|^2; \\ \text{spin B} = -\hbar/2, |\uparrow_A\rangle, \text{probability}|C_{\uparrow\downarrow}|^2; \\ \text{spin B} = -\hbar/2, |\downarrow_A\rangle, \text{probability}|C_{\downarrow\downarrow}|^2. \end{cases} \tag{1.7}$$

If now we measure the spin of particle A, its average value will be

$$\left\langle \frac{\hbar}{2}\sigma_z \right\rangle = \frac{\hbar}{2}\left(|C_{\uparrow\uparrow}|^2 + |C_{\uparrow\downarrow}|^2 - |C_{\downarrow\uparrow}|^2 - |C_{\downarrow\downarrow}|^2\right). \tag{1.8}$$

The normalization of the initial state vector, $|||\Psi\rangle_{AB}||^2 = \sum_{ij}|C_{ij}|^2 = 1$, ensures that the probabilities of the outcomes of (1.7) add up to unity, so everything is fine,

---

[4] We are encouraged by the emergence of time-irreversible Boltzmann equations from time-reversible Newtonian dynamics of point-like particles.

Cambridge University Press
978-0-521-11369-4 - Quantum Engineering: Theory and Design of Quantum Coherent Structures
A. M. Zagoskin
Excerpt
More information

except that Eq. (1.7) is a very awkward way of dealing with even such a small system. Just imagine we had three particles!

The situation improves once the state vector is replaced with the density matrix. Let us choose a set of normalized, but not necessarily mutually orthogonal states in the Hilbert space, $\{|\Psi_j\rangle\}$, and write an operator

$$\rho = \sum_j p_j |\Psi_j\rangle\langle\Psi_j|, \tag{1.9}$$

where $p_j \geq 0$, and $\sum_j p_j = 1$. This obviously Hermitian operator can be interpreted as describing a *statistical ensemble* of quantum systems, which can be in a state $|\Psi_j\rangle$ with probability $p_j$. This is exactly the situation described by Eq. (1.7).

The convenience of using this operator becomes clear when calculating average values of the observables. According to rule (3) of quantum mechanics, the measured value of an observable $A$ is always one of the eigenvalues of $A$, and according to rule (4), the probability of measuring a particular $a_k$ in state $|\Psi_j\rangle$ is $|\langle a_k|\Psi_j\rangle|^2$. Therefore

$$\begin{aligned}\langle A\rangle_{\Psi_j} &= \sum_k |\langle a_k|\Psi_j\rangle|^2 a_k = \sum_k \langle\Psi_j|a_k\rangle a_k \langle a_k|\Psi_j\rangle \\ &= \sum_k \langle\Psi_j|a_k|a_k\rangle\langle a_k|\Psi_j\rangle = \langle\Psi_j|A\left(\sum_k |a_k\rangle\langle a_k|\right)|\Psi_j\rangle = \langle\Psi_j|A|\Psi_j\rangle,\end{aligned} \tag{1.10}$$

a standard formula. Since now the system is in state $|\Psi_j\rangle$ only with some probability $p_j$, we must average over these too, with the result

$$\langle A\rangle = \sum_j p_j \langle A\rangle_{\Psi_j} = \sum_j p_j \langle\Psi_j|A|\Psi_j\rangle \equiv \mathrm{tr}(\rho A). \tag{1.11}$$

The system, described by a density matrix, is said to be in a *mixed state*, instead of a *pure state* (when a single state vector suffices). One can write a density matrix for a pure state: it will include only one component, $\rho_{\mathrm{pure}} = |\Psi\rangle\langle\Psi|$.

As a Hermitian operator, the density matrix has an orthonormal set of eigenstates, $|\rho_j\rangle$, with corresponding eigenvalues, $\rho_j$. It can be, therefore, written in the *spectral representation*,

$$\rho = \sum_j \rho_j |\rho_j\rangle\langle\rho_j|. \tag{1.12}$$

Since $\rho_j$ can be interpreted as the probability of finding the system in state $|\rho_j\rangle$, all $\rho_j \geq 0$; in other words, the density matrix is *positive semidefinite*.

The density matrix can be written in any basis in the Hilbert space, and will usually have neither the explicit form (1.9) nor (1.12). We will therefore list here its invariant properties.

1) $\mathtt{tr}\,\rho = 1$ (follows directly from the definition (1.9)).
2) $\rho^2 = \rho$ if and only if the state is pure. Indeed, if $\rho = |\Psi\rangle\langle\Psi|$, then $\rho^2 = |\Psi\rangle\langle\Psi||\Psi\rangle\langle\Psi| = |\Psi\rangle\langle\Psi| = \rho$. Conversely, if $\rho^2 = \rho$, then (from (1.12)) $\rho_j^2 = \rho_j$, and the eigenvalues of the density matrix are either zeros, or ones. Since its trace is one, then only one eigenvalue can be equal to one, i.e., there is only one term in the spectral decomposition, and the state is indeed pure.
3) $\mathtt{tr}(\rho^2) \leq 1$, and there is equality if and only if the state is pure (therefore the quantity $\varsigma = \mathtt{tr}(\rho^2)$ is sometimes called *purity*). Starting from (1.9) and recalling that $|\langle\Psi_j|\Psi_k\rangle| \leq 1$,

$$\begin{aligned}\mathtt{tr}(\rho^2) &= \mathtt{tr}\left(\sum_{jk} p_j p_k |\Psi_j\rangle\langle\Psi_j|\Psi_k\rangle\langle\Psi_k|\right)\\ &= \sum_j p_j \sum_k p_k \left|\langle\Psi_j|\Psi_k\rangle\right|^2 \leq \sum_j p_j = 1.\end{aligned}$$

On the other hand, if $\mathtt{tr}(\rho^2) = 1$, then $\sum_j p_j \left(\sum_k p_k \left|\langle\Psi_j|\Psi_k\rangle\right|^2\right) = 1$. Since the non negative $p_j$'s add up to one, there is equality only if $p_j = \delta_{jq}$ for some $q$, that is, if the state is pure.

These properties provide us with a reliable way to check any approximate calculations of a density matrix, as well as a criterion of whether a given system is in a pure or a mixed state.

### *1.2.2 Averages, probabilities and coherences*

Consider now, as before, a quantum-mechanical system in the pure state (1.6). Its density matrix is

$$\rho_{AB} = |\Psi\rangle_{AB}\langle\Psi|_{AB} = \sum_{j,k,l,m=\uparrow,\downarrow} C_{jk}C^*_{lm}|j_A k_B\rangle\langle l_A m_B|.$$

Let's take a partial trace of this operator over the states of particle B:

$$\begin{aligned}\rho_A \equiv \mathtt{tr}_B\,\rho_{AB} &= \sum_{q=\uparrow,\downarrow} \langle q_B|\Psi\rangle_{AB}\left(\langle\Psi|_{AB}\right)|q_B\rangle\\ &= \sum_{q,j,k,l,m=\uparrow,\downarrow} C_{jk}C^*_{lm}\delta_{qk}|j_A\rangle\langle l_A|\delta_{qm} = \sum_{q,j,l=\uparrow,\downarrow} C_{jq}C^*_{lq}|j_A\rangle\langle l_A|. \qquad (1.13)\end{aligned}$$

Cambridge University Press
978-0-521-11369-4 - Quantum Engineering: Theory and Design of Quantum Coherent Structures
A. M. Zagoskin
Excerpt
More information

To find the observed value of the $z$-component of spin of particle A, we now take the trace of the spin operator with *reduced density matrix* $\rho_A$:

$$\left\langle \frac{\hbar}{2}\sigma_z^A \right\rangle = \frac{\hbar}{2}\mathtt{tr}\left(\rho_A \sigma_z^A\right) = \frac{\hbar}{2}\mathtt{tr} \sum_{q,j,l=\uparrow,\downarrow} C_{jq}C_{lq}^{*}\sigma_z|j_A\rangle\langle l_A|$$
$$= \frac{\hbar}{2}\mathtt{tr} \sum_{q,j=\uparrow,\downarrow} |C_{jq}|^2 \langle j_A|\sigma_z|j_A\rangle. \tag{1.14}$$

This equation automatically yields the same result as Equations (1.7) and (1.8), and in a compact form. We see that as long as we are interested only in the results of measurement on one subsystem, we can use the reduced density matrix by taking the trace of the full density matrix over all the irrelevant degrees of freedom. This is an important procedure, which is crucial for describing open quantum systems, i.e., systems for which one can neither neglect their interaction with the environment nor take the latter into account explicitly.

Let us look at Eq. (1.13) in more detail. This expression contains more information than just the average value of the operator $\sigma_z$. The diagonal terms provide the probabilities of finding the system with spin up/down. What about the off-diagonal ones?

Let us write a density matrix in an orthonormal basis of eigenstates of some observable $A$, and we chose $A$ such that its eigenstates are macroscopically observable pointer states – e.g., charges, positions or momenta: $\rho = \sum_{ij} \rho_{ij}|a_i\rangle\langle a_j|$. Its diagonal elements give the probabilities of finding the system in appropriate eigenstates, and if we are only performing the measurements of $A$ this is all we care about or can extract from the experiment. The off-diagonal terms could have never been there. Nevertheless, they are crucially important.

We have seen that a diagonal density matrix describes a pure state if and only if its only element is unity (invariant property number 2). If this is the case, we will always measure the same eigenvalue $a_j$ of the operator $A$ and leave the system in the same state $|a_j\rangle$.[5] (This is called *quantum non demolition (QND) measurement*, and we will discuss the implications of such experiments in § 5.5.2.) If not, the density matrix must describe a mixed state, so the absence of the off-diagonal matrix elements is a tell-tale sign. These elements describe *quantum coherence*. If they are zero, mathematically there is no basis in which the density matrix contains a single diagonal element; physically, the system is in a mixed state; from the point of view of the observation, there is no observable of which the system is in an eigenstate, and its measurement will yield some eigenstate at random, with a given

---

[5] We need not bother here with the cases of degenerate eigenstates or the continuous spectrum of the operator $A$.

Cambridge University Press
978-0-521-11369-4 - Quantum Engineering: Theory and Design of Quantum Coherent Structures
A. M. Zagoskin
Excerpt
More information

probability. This is why the evolution of the off-diagonal elements of the density matrix *(“coherences”)* is of special interest: their disappearance – “decoherence”, due to whatever processes that cause it – means the reduction of the state of the system to a mixture, with the loss of specifically quantum correlations. The rate of this reduction can, therefore, be taken as the *decoherence rate*, and its inverse as the *decoherence time*.

### *1.2.3 Entanglement*

We have already seen that when several distinct systems (e.g., a cat and a $^{210}$Po atom) are described by a single quantum state, this can lead to quite counterintuitive conclusions. Suppose that two quantum systems (e.g., spin-1/2 particles A and B) are in a pure quantum state $|\Psi\rangle_{AB}$. Then, generally speaking, separately, neither of them can be in a definite quantum state. Consider, for example, the expectation value of $\sigma_z$ for particle A, Eq. (1.14):

$$\left\langle \frac{\hbar}{2}\sigma_z^A \right\rangle = \frac{\hbar}{2}\mathtt{tr}\left(\rho_A \sigma_z^A\right).$$

The reduced density matrix $\rho_A = \mathtt{tr}_B[|\Psi\rangle_{AB}\langle\Psi|_{AB}]$ can only have the form $|\psi\rangle_A\langle\psi|_A$, corresponding to particle A being in a pure quantum state $|\psi\rangle_A$, if the quantum state of the whole system is factorized:

$$|\Psi\rangle_{AB} = |\psi\rangle_A \otimes |\psi\rangle_B. \tag{1.15}$$

Otherwise, particle A by itself is in a mixed state, even though the whole system consisting of A and B is in a pure state, and there is no interaction between A and B. The non factorized states of multipartite systems are called *entangled*, which expresses this specifically quantum correlation between its components. Essentially, it means that the properties of a quantum system in general can never be reduced to the sum of the properties of its constituents.

For a pure state $|\Psi\rangle_{ABC\ldots}$ of a system, comprising any number of subsystems, the test for entanglement is straightforward: if all reduced density matrices correspond to pure states, the state $|\Psi\rangle_{ABC\ldots}$ can be written as $|\psi\rangle_A \otimes |\psi\rangle_B \otimes |\psi\rangle_C \ldots$, and there is no entanglement. But entanglement is not limited to pure states: certain quantum correlation can exist even if the system itself is described by a density matrix. For a given density matrix one can find different numerical measures of entanglement (all of which are zero for a factorized state; see, e.g., Horodecki et al., 2009).