Chapter 2

Quantum mechanics and probability

2.1 Classical and quantum probabilities

In this section we extend the quantum description to states that include "classical" uncertainties, which add to the probabilities inherent in the quantum formalism. Such additional uncertainties may be due to lack of a full knowledge of the state vector of a system or due to the description of the system as a member of a statistical ensemble. But it may also be due to *entanglement*, when the system is the part of a larger quantum system.

2.1.1 Pure and mixed states, the density operator

The state of a quantum system, described as a wave function or an abstract vector in the state space, has a *probability interpretation*. Thus, the wave function is referred to as a *probability amplitude* and it predicts the result of a measurement performed on the system only in a statistical sense. The state vector therefore characterizes the state of the quantum system in a way that seems closer to the statistical description of a classical system than to a detailed, non-statistical description. However, in the standard interpretation, this uncertainty about the result of a measurement performed on the system is not ascribed to lack of information about the system. We refer to the quantum state described by a (single) state vector, as a *pure state* and consider this to contain *maximum available information* about the system. Thus, if we intend to acquire further information about the state of the system by performing measurements on the system, this will in general lead to a change of the state vector which we interpret as a *real modification* of the

physical state due to the action of the measuring device. It cannot be interpreted simply as a (passive) collection of additional information.[1]

In reality we will often have less information about a quantum system than the maximally possible information contained in the state vector. Let us for example consider the spin state of the silver atoms emerging from the oven of the Stern-Gerlach experiment. In principle each atom could be in a pure spin state $|\mathbf{n}\rangle$ with a quantized spin component in the $\mathbf{n}$-direction. However, the collection of atoms in the beam clearly have spins that are isotropically oriented in space and they cannot be described by a singel spin vector. Instead they can be associated with a statistical ensemble of vectors. Since the spins are isotropically distributed, this indicates that the ensemble of spin vectors $|\mathbf{n}\rangle$ should have a uniform distribution over all directions $\mathbf{n}$. However, since these vectors are not linearly independent, an isotropic ensemble is equivalent to an ensemble with only two states, spin up and spin down in *any* direction $\mathbf{n}$, with equal probability for the two states.

We therefore proceed to consider situations where a system is described not by a single state vector, but by an ensemble of state vectors, $\{|\psi\rangle_1, |\psi\rangle_2, ..., |\psi\rangle_n\}$ with a probability distribution $\{p_1, p_2, ..., p_n\}$ defined over the ensemble. We may consider this ensemble to contain both *quantum probabilities* carried by the state vectors $\{|\psi\rangle_k\}$ and *classical probabilities* carried by the distribution $\{p_k\}$. A system described by such an ensemble of states is said to be in a *mixed state*. There seems to be a clear division between the two types of probabilities, but as we shall see this is not fully correct. There are interesting examples of mixed states with no clear division between the quantum and classical probabilities.

The expectation value of a quantum observable in a state described by an ensemble of state vectors is

$$\langle A \rangle = \sum_{k=1}^{n} p_k \langle A \rangle_k = \sum_{k=1}^{n} p_k \langle \psi_k | \hat{A} | \psi_k \rangle \tag{2.1}$$

This expression motivates the introduction of the *density operator* associated with

---

[1]An interesting question concerns the interpretation of a quantum state vector $|\psi\rangle$ as being the *objective* state associated with a *single* system. An alternative understanding, as stressed by Einstein, claims that the probability interpretation implies that the state vector can only be understood as an *ensemble variable*, it describes the state of an ensemble of identically prepared systems. From Einstein's point of view this means that additional information, beyond that included in the wave function, should *in principle* be possible to acquire when we consider single systems rather than ensembles.

the mixed state,

$$\hat{\rho} = \sum_{k=n}^{n} p_k \, |\psi_k\rangle\langle\psi_k| \tag{2.2}$$

The corresponding matrix, defined by reference to an (orthogonal) basis, is called the *density matrix*,

$$\rho_{ij} = \sum_{k=n}^{n} p_k \, \langle\phi_i|\psi_k\rangle\langle\psi_k|\phi_j\rangle \tag{2.3}$$

The important point to note is that *all information about the mixed state is contained in the density operator of the state*, in the sense that the expectation value of *any* observable can be expressed in terms of $\hat{\rho}$,

$$
\begin{aligned}
\langle A\rangle &= \sum_{k=1}^{n} p_k \sum_{ij} \langle\psi_k|\phi_i\rangle\langle\phi_i|\hat{A}|\phi_j\rangle\langle\phi_j|\psi_k\rangle \\
&= \sum_i \sum_{k=1}^{n} p_k \langle\phi_j|\psi_k\rangle\langle\psi_k|\hat{A}|\phi_j\rangle \\
&= Tr(\hat{\rho}\hat{A}) \tag{2.4}
\end{aligned}
$$

The density operator satisfies certain properties, in particular,

$a)$ Hermiticity : $\hat{\rho}^\dagger = \hat{\rho} \;\Rightarrow\; p_k = p_k^*$ (real eigenvalues)

$b)$ Positivity : $\langle\chi|\hat{\rho}|\chi\rangle \geq 0$ for all $|\chi\rangle \;\Rightarrow\; p_k \geq 0$ (non − negative eigenvalues)

$c)$ Normalization : $Tr\,\hat{\rho} = 1 \;\Rightarrow\; \sum_k p_k = 1$ (sum of eigenvalues is 1) (2.5)

These conditions follows from (2.2) with the coefficients p_k interpreted as probabilities. We also note that

$$Tr\,\hat{\rho}^2 = \sum_k p_k^2 \;\Rightarrow\; 0 < Tr\,\hat{\rho}^2 \leq 1 \tag{2.6}$$

This inequality follows from the fact that for all eigenvalues $p_k \leq 1$, which means that $Tr\,\hat{\rho}^2 \leq Tr\,\hat{\rho}$.

The pure states are the special case where one of the probabilities p_k is equal to 1 and the others are 0. In this case the density operator is the projection operator on a single state,

$$\hat{\rho} = |\psi\rangle\langle\psi| \;\Rightarrow\; \hat{\rho}^2 = \hat{\rho} \tag{2.7}$$

In this case $Tr(\hat{\rho}^2) = 1$, while for all the (truly) mixed states $Tr(\hat{\rho}^2) < 1$.

A general density matrix can be written in the form

$$\hat{\rho} = \sum_k p_k |\psi_k\rangle\langle\psi_k| \tag{2.8}$$

where the states $|\psi_k\rangle$ may be identified as members of a statistical ensemble of state vectors associated with the mixed state. Note, however, that *this expansion is not unique. There are many different ensembles that give rise to the same density matrix*. This means that even if all information about the mixed state is contained in the density operator in order to specify the expectation value of any observable, there may in principle be additional information available that specifies the physical ensemble to which the system belongs.

An especially useful expansion of a density operator is the expension in terms of its eigenstates. In this case the states $|\psi_k\rangle$ are orthonomal and the eigenvalues are the probabilities p_k associated with the eigenstates. This expansion is unique unless there are eigenvalues with degeneracies.

As opposed to the pure states the mixed states are not providing the maximal possible information about the system. This is due to the classical probabilities contained in the mixed state, which to some degree makes it similar to a statistical state in a classical system. Thus, additional information may in principle be available without interacting with the system. For example, two different observers may have different degrees of information about the system and therefore associate different density matrices to the system. By exchanging information they may increase their knowledge about the system without interacting with it. If, on the other hand, two observers have maximal information about the system, which means that they both describe it by a pure state, they have to associate the same state vector with the system in order to have a consistent description.

A mixed state described by the density operator (2.8) is sometimes referred to as an *incoherent* mixture of the states $|\psi_k\rangle$. A *coherent* mixture is instead a *superposition* of the states,

$$|\psi\rangle = \sum_k c_k |\psi_k\rangle \tag{2.9}$$

which then represents a pure state. The corresponding density matrix can be written as

$$\hat{\rho} = \sum_k |c_k|^2 |\psi_k\rangle\langle\psi_k| + \sum_{k \neq l} c_k^* c_l |\psi_k\rangle\langle\psi_l| \tag{2.10}$$

Comparing this with (2.8) we note that the first sum in (2.10) can be identified with the incoherent mixture of the states (with $p_k = |c_k|^2$). The second sum involves the *interference* terms of the superposition and these are essential for the coherence effect. Consequently, if the off-diagonal matrix elements of the density matrix (2.10) are erased, the pure state is reduced to a mixed state with the same probability p_k for the states $|\psi_k\rangle$.

The time evolution of the density operator for an isolated (closed) system is determined by the Schrödinger equation. As follows from the expression (2.8) the density operator satisfies the dynamical equation

$$i\hbar\frac{\partial}{\partial t}\hat{\rho} = \left[\hat{H}, \hat{\rho}\right]$$
(2.11)

This looks similar to the Heisenberg equation of motion for an observable, but one should note that Eq.(2.11) *is valid in the Schrödinger picture*. In the Heisenberg picture, the density operator, like the state vector is time independent.

One should note the close similarity between Eq.(2.11) and Liouville's equation for the classical probability density $\rho(q, p)$ in phase space

$$\frac{\partial}{\partial t}\hat{\rho} = \{\rho, H\}_{PB}$$
(2.12)

where $\{,\}_{PB}$ is the Poisson bracket.

2.1.2   Entropy

In the same way as one associates *entropy* with statistical states in a classical system, one associates entropy with mixed state as a measure of the lack of (optimal) information associated with the state. The *von Neuman entropy* is defined as

$$S = -Tr(\hat{\rho}\log\hat{\rho})$$
(2.13)

Rewritten in terms of its eigenvalues of $\hat{\rho}$ it has the form

$$S = -\sum_k p_k \log p_k$$
(2.14)

which shows that it is closely related to the entropy defined in statistical mechanics and in information theory.

The pure states are states with zero entropy. For mixed states the entropy measures "how far away" the state is from being pure. The entropy increases

when the probabilities get distributed over many states. In particular we note that for a finitedimensional Hilbert space, a maximal entropy state exist where all states are equally probable. The corresponding density operator is

$$\hat{\rho}_{max} = \frac{1}{n} \sum_k |k\rangle\langle k| \tag{2.15}$$

where n is the dimension of the Hilbert space and $\{|k\rangle\}$ is an orthonormal set of basis vectors. Thus, the density operator is proportional to a projection operator that projects on the full Hilbert space. The corresponding maximal value of the entropy is

$$S_{max} = \log n \tag{2.16}$$

A *thermal state* is a special case of a mixed state, with a (statistical) *Boltzman distribution* over the energy levels. It is described by a *temperature dependent* density operator of the form

$$\hat{\rho} = N e^{-\beta \hat{H}} \tag{2.17}$$

where $\beta = 1/(k_B T)$ and N is a normalization factor. It is given by

$$N^{-1} = Tr e^{-\beta \hat{H}} = \sum_k e^{-\beta E_k} \tag{2.18}$$

in order to give $\hat{\rho}$ the correct normalization (2.5) consistent with the probability interpretation. In the above expression E_k are the energy eigenvalues of the system. The close relation between the normalization factor and the *partition function* in (classical) statistical mechanics is apparent.

The *quantum statistical mechanics* is based on the definition of density matrices associated with statistical ensembles. Thus, the density matrix (2.17) is associated with the *canonical ensemble*. Furthermore the *thermodynamic entropy* is, in the quantum statistical mechanics, identical to the von Neuman entropy (2.13) apart from a factor k_B, the *Boltzmann constant*. We will in the following make some futher study of the entropy, but focus mainly on its *information* content rather than thermodynamic relevance. Whereas the thermodynamic entropy is most relevant for systems with a large number of degrees of freedom, the von Neuman entropy (2.13) is also highly relevant for small systems in the context of quantum information.

2.1.3 Mixed states for a two-level system

For the two-level system we can give an explicit (geometrical) representation of the density matrices of mixed (and pure) states.

Since the matrices are hermitian, with trace 1, they may be written as

$$\rho = \frac{1}{2}(\mathbf{1} + \mathbf{r} \cdot \boldsymbol{\sigma}) \tag{2.19}$$

where $\mathbf{r}$ is a three-component vector. The condition for this to be a pure state is

$$\begin{aligned} \rho^2 &= \rho \\ \Rightarrow \quad \frac{1}{4}(\mathbf{1} + \mathbf{r}^2 + 2\mathbf{r} \cdot \boldsymbol{\sigma}) &= \frac{1}{2}(\mathbf{1} + \mathbf{r} \cdot \boldsymbol{\sigma}) \\ \Rightarrow \quad \mathbf{r}^2 &= 1 \end{aligned} \tag{2.20}$$

Thus, a pure state can be written as

$$\rho = \frac{1}{2}(\mathbf{1} + \mathbf{n} \cdot \boldsymbol{\sigma}) \tag{2.21}$$

with $\mathbf{n}$ as a unit vector. This represents a projection operator that projects on the one-dimensional space spanned by the *spin up* eigenvector of $\mathbf{n} \cdot \boldsymbol{\sigma}$. We note that this is consistent we the discussion of section (**??**), where it was shown that *any state* in the two-dimensional space would be the spin up eigenstate of the operator $\mathbf{n} \cdot \boldsymbol{\sigma}$ for some unit vector $\mathbf{n}$.

For a general, mixed state we can write the density matrix as

$$\begin{aligned} \rho &= \frac{1}{2}(\mathbf{1} + r\mathbf{n} \cdot \boldsymbol{\sigma}) \\ &= \frac{1}{2}(1+r)\frac{\mathbf{1} + \mathbf{n} \cdot \boldsymbol{\sigma}}{2} + \frac{1}{2}(1-r)\frac{\mathbf{1} - \mathbf{n} \cdot \boldsymbol{\sigma}}{2} \end{aligned} \tag{2.22}$$

with $\mathbf{r} = r\mathbf{n}$. This means that the eigenvectors of ρ, also in this case are the eigenvectors of $\mathbf{n} \cdot \boldsymbol{\sigma}$, but the eigenvalue p_+ of the spin up state and the eigenvalue p_- of the spin down state are given by

$$p_{\pm} = \frac{1}{2}(1 \pm r) \tag{2.23}$$

The interpretation of $p_{\pm}$ as probabilities means that they have to be positive (and less than one). This is satisfied if $r \leq 1$. This result shows that all the physical
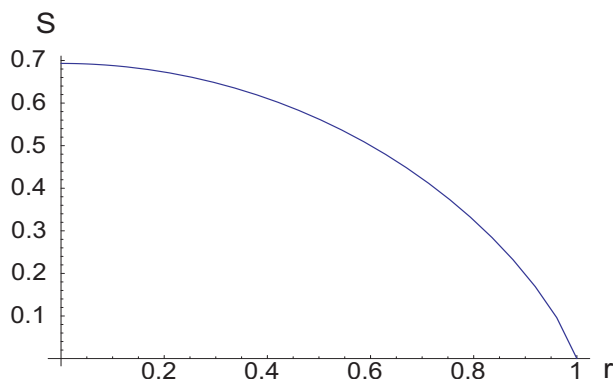
Figure 2.1: The entropy of mixed states for a two level system. The entropy S is shown as a function of the radial variable r for the *Bloch sphere representation* of the states.

states, both pure and mixed can be represented by points in a three dimensional sphere of radius $r = 1$. Points on the surface ($r = 1$) correspond to the pure state. As r decreases the states gets less pure and for $r = 0$ we find the state of maximal entropi. The entropi, defined by (2.14), is a monotonic function of r, with maximum ($\log 2$) for $r = 0$ and minimum (0) for $r = 1$. The explicit expression for the entropi is

$$S = -(\frac{1}{2}(1+r)\log\frac{1}{2}(1+r) + \frac{1}{2}(1-r)\log\frac{1}{2}(1-r)) \qquad (2.24)$$

The *sphere of states* for the two-level system is called the *Bloch sphere*. More generally the set of density matrices of a quantum system can be shown to form a *convex set* (see Problem (2.4.3)).

2.2   Entanglement

In this section we consider *composite systems*, which can be considered as consisting of two or more *subsystems*. These subsystems will in general be *correlated* due to interactions between the two parts. A quantum system may have correlations that in some sense are stronger than what is possible in a classical system. We refer to this effect as quantum *entanglement*. This is considered as one of the clearest marks of the difference between classical and quantum physics.

2.2.1 Composite systems

To this point we have mainly considered an *isolated quantum system*. Only the variables of the system then enters in the quantum description, in the form of of state vectors and observables, while the dynamical variables of other systems are irrelevant. As long as the system stays isolated all interactions with other systems are negligible and the time evolution is described by the Schrödinger equation.

With maximal information about the system we describe it by a pure state (a state vector), with less than maximal information it is described by a mixed state (a density matrix). If the system is in a pure state, it will continue to be in a pure state as long as it stays isolated. For a mixed state, the degree of "non-purity" measured by the entropy will stay constant as long as it is isolated. This follows from the fact that the time evolution is unitary and the eigenvalues of the density operator therefore do not change with time.

Clearly an isolated system is an abstraction since interactions with other systems (generally referred to as the *environment*) can never be totally absent. However, in some cases this idealization works perfectly well to a high dergee of acuracy. If interactions with the environments cannot be neglected, also variables associated with the environment have to be taken into account. If these act randomly on the system they tend to introduce *decoherence* in the system. The state develops in the direction of being less pure, *i.e.*, the entropy increases. However, a systematic manipulation with the system may change the state in the direction of being more pure. Thus, a partial or full measurement performed on the system will be of this kind.

Even if a system A cannot be considered as isolated, sometimes it will be a part of a *larger* isolated system. We will consider this sitation and assume that the total system can be descibed in terms of a set of variables for system A and a set of variables for the rest of the system, which we denote B. We assume that these two sets of variables can be regarded as independent (they are associated with independent degrees of freedom for the two subsystems A and B), and that the interactions between A and B do not destroy this relative independence. We further assume that the totality of dynamical variables for A and B provides a complete set of variables for the full system. We refer to this as *a composite system*.

The interactions between the two parts of a composite system will introduce *correlations* between the two parts. Thus, the evolution of the two subsystems will be correlated and the expectation values of variables from system A and B will in general also be correlated. Correlations is a typical feature of interacting systems

both at the classical and quantum level. We shall focus here on correlations that are special for a quantum mechanical system.

In mathematical terms we describe the Hilbert space $\mathcal{H}$ of a composite system as a *tensor product* of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ associated with the two subsytems,

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \tag{2.25}$$

A basis $|i\rangle$ of $\mathcal{H}_A$ and a basis $|\alpha\rangle$ of $\mathcal{H}_B$ then defines a tensor product basis for $\mathcal{H}$,

$$|i\alpha\rangle = |i\rangle \otimes |\alpha\rangle \tag{2.26}$$

A general vector in $\mathcal{H}$ is of the form

$$|\chi\rangle = \sum_{i\alpha} c_{i\alpha} |i\rangle \otimes |\alpha\rangle \tag{2.27}$$

and in general it is *not* a tensor product of a vector from $\mathcal{H}_A$ with a vector from $\mathcal{H}_B$.

The observables of system A and the observables of system B define *commuting observables* of the full system. This follows directly from how they act on the basis vectors (2.26). With $\hat{A}$ as an observable for system A and $\hat{B}$ as an observable for system B, we have

$$\hat{A}|i\alpha\rangle = \sum_j A_{ji}|j\alpha\rangle , \quad \hat{B}|i\alpha\rangle = \sum_\beta B_{\beta\alpha}|i\beta\rangle \tag{2.28}$$

Clearly

$$\hat{A}\hat{B}|i\alpha\rangle = \sum_{j\beta} A_{ji} B_{\beta\alpha} |j\beta\rangle = \hat{B}\hat{A}|i\alpha\rangle \tag{2.29}$$

The expectation values of observables acting only on subsystem A are determined by the *reduced density matrix* of system A. This is obtained by taking the *partial trace* with respect to the coordinates of system B,

$$\hat{\rho}^A = Tr_B(\hat{\rho}) , \quad \rho_{ij}^A = \sum_\alpha \rho_{i\alpha\,j\alpha} \tag{2.30}$$

In the same way we define the reduced density matrix for system B

$$\hat{\rho}^B = Tr_A(\hat{\rho}) , \quad \rho_{\alpha\beta}^B = \sum_i \rho_{i\alpha\,i\beta} \tag{2.31}$$

In general the reduced density matrices will correspond to mixed states for the subsystems A and B, even if the full system is in a pure state.

For a composite system there exist certain relations between the entropy $S(\rho)$ of the full system and the entropies of the subsystems, $S(\rho^A)$ and $S(\rho^B)$. Thus, for a *bipartite* system (a system consisting of two parts), like the one considered above, the following inequalities are satisfied,

$$S(\rho) \leq S_A + S_B , \quad S(\rho) \geq |S_A - S_B| \tag{2.32}$$

We note in particular the first inequality. This may seem somewhat surprising, since it indicates that there are states of a composite system that are more pure than the states of its subsystems. A corresponding situation for a classical system would be rather paradoxical since it would mean that the total system could be in a state that was more ordered than the states of its subsystems. Another way to interpret the inequality is that full information about the states of the subsystems A and B will in general not be sufficient to give full information about the state of the total system A+B. When there are *correlations* between the two subsystems, these are not seen in the description of A and B separately.

2.2.2 Correlations and entanglement

A product state of the form

$$\hat{\rho} = \hat{\rho}^A \otimes \hat{\rho}^B \tag{2.33}$$

is a state with *no correlations* between the subsystems A and B. For any observable $\hat{A}$ acting on A and $\hat{B}$ acting on B the expectation value of the product operator $\hat{A}\hat{B}$, is then simply the product of the expectation values

$$\langle AB \rangle = \langle A \rangle_A \langle B \rangle_B \tag{2.34}$$

Written in terms of the density operators,

$$Tr(\hat{\rho}\hat{A}\hat{B}) = Tr_A(\hat{\rho}^A \hat{A})Tr_B(\hat{\rho}^B \hat{B}) . \tag{2.35}$$

Conversely, if the product relation (2.34) is correct for *all* observables $\hat{A}$ acting on A and *all* observables $\hat{B}$ acting on B, the two subsystems are uncorrelated and the density operator can be written in the product form (2.34).

We next consider a (mixed) state of the form

$$\hat{\rho} = \sum_k p_k \, \hat{\rho}_k^A \otimes \hat{\rho}_k^B . \tag{2.36}$$

With $\{p_k\}$ as a probability distribution, this density matrix can be viewed as describing a *statistical ensemble* of product states, *i.e.*, of states which, when regarded separately, do not contain correlations between the two subsytems. However, the density matrix of the ensemble *does* contain (statistical) correlations between the two subsystems. Thus, the expectation value of a product of observables for A and B is

$$
\begin{aligned}
\langle AB \rangle &= \sum_k p_k \langle A \rangle_{kA} \langle B \rangle_{kB} \\
&= \sum_k p_k Tr_A(\hat{\rho}_k^A \hat{A}) Tr_B(\hat{\rho}_k^B \hat{B})
\end{aligned} \tag{2.37}
$$

and this is generally different from the product of expectation values,

$$
\begin{aligned}
\langle A \rangle_A \langle B \rangle_B &= \sum_k p_k \langle A \rangle_{kA} \sum_l p_l \langle B \rangle_{lB} \\
&= \sum_{kl} p_k p_l \, Tr_A(\hat{\rho}_k^A \hat{A}) Tr_B(\hat{\rho}_l^B \hat{B}) \, .
\end{aligned} \tag{2.38}
$$

The correlations contained in a density operator of the form (2.36) we may refer to as *classical* correlations, since these are of the same form that we find in a classical, statistical description of a composite system. However, in a quantum system there may be correlations that cannot be written in the form (2.36) and which are therefore of genuinely quantum mechanical nature. Composite systems with this kind of correlations are said to be *entangled*, and in some sense the correlations that are found in entangled states are stronger than can be achieved due to classical statistical correlations. In recent years there has been much effort devoted to give a concise quantiative meaning to entanglement in composite systems. However, for bipartite systems in mixed states and for multipartite systems in general, it is not obvious how to quantify deviations from classical correlations. But for a bipartite system in a pure state the definition of entanglement is unambiguous. We focus on this case.

We therfore now consider a general *pure* state of the total system A+B. The entropy of such a state clearly vanishes. The state has the general form

$$
|\chi\rangle = \sum_{i\alpha} c_{i\alpha} |i\rangle_A \otimes |\alpha\rangle_B \tag{2.39}
$$

but can also be written in the diagonal form

$$
|\chi\rangle = \sum_n d_n |n\rangle_A \otimes |n\rangle_B \tag{2.40}
$$

where $\{|n\rangle_A\}$ is a set of orthonormal states for the A system and $\{|n\rangle_B\}$ is a set of orthonormal states for the B system. This form of the state vector as a simple sum over product vectors rather than a double sum is referred to as the *Schmidt decomposition* of $|\chi\rangle$. It can be shown to be generally valid for a composite two-partite system.

The form of the state vector $|\chi\rangle$ is similar to that of density operator (2.36) for the correlated state. In the same way as for the density operator the system is uncorrelated if the sum includes only one term, *i.e.*, if the state has a product form. However, the correlations implied by the general form of the state vector (2.40) are different from those of the (classically) correlated state (2.36). From the Schmidt decomposition follows that the density matrix of the full system corresponding to the pure state $|\chi\rangle$ is

$$\hat{\rho} = \sum_{nm} d_n d_m^* \, |n\rangle_A \langle m|_A \otimes |n\rangle_B \langle m|_B \tag{2.41}$$

This is not of the form (2.36). The state (2.40) has *correlations between state vectors* of the two subsystems rather than correlations only between density operators.

From (2.41) also follows that the reduced density operators of the two subsystems have the form

$$\hat{\rho}^A = \sum_n |d_n|^2 |n\rangle_A \langle n|_A \, , \quad \hat{\rho}^B = \sum_n |d_n|^2 |n\rangle_B \langle n|_B \tag{2.42}$$

The expressions show that all the eigenvalues of the two reduced density matrices are the same. This implies that the von Neuman entropy of the two subsystems is the same,

$$S_A = S_B = - \sum_n |d_n|^2 \log |d_n|^2 \tag{2.43}$$

In this case, with the full system A+B in a pure state, the entropy of one of the subsystems (A or B) is taken as a quantitative measure of the entanglement between the subsystems.

The Schmidt decomposition

We show here that a general state of the composite system can be written in the form (2.40). The starting point is the general expression (2.39). We introduce a unitary transformation U for the basis of system A,

$$|i\rangle_A = \sum_m U_{mi} \, |m\rangle_A \tag{2.44}$$

and rewrite the state vector as

$$
\begin{aligned}
|\chi\rangle &= \sum_{i\alpha}\sum_{m} c_{i\alpha} U_{mi} |m\rangle_A \otimes |\alpha\rangle_B \\
&= \sum_{m} |m\rangle_A \otimes \left(\sum_{i\alpha} c_{i\alpha} U_{mi} |\alpha\rangle_B\right) \\
&\equiv \sum_{m} |m\rangle_A \otimes |\tilde{m}\rangle_B
\end{aligned}
\tag{2.45}
$$

If the unitary transformation can be chosen to make the set of states $\{|\tilde{m}\rangle\}$ and *orthogonal* set of states, then (2.40) follows. The scalar product between two of the states is

$$
\begin{aligned}
\langle \tilde{n}|\tilde{m}\rangle_B &= \sum_{i\alpha}\sum_{j\beta} c_{j\beta}^* U_{nj}^* c_{i\alpha} U_{mi} \langle \beta|\alpha\rangle_B \\
&= \sum_{ij}\sum_{\alpha} c_{j\alpha}^* U_{nj}^* c_{i\alpha} U_{mi} \\
&= (UCC^\dagger U^\dagger)_{mn}
\end{aligned}
\tag{2.46}
$$

In the last expression the coefficients $c_{i\alpha}$ have been treated as the matrix elements of the matrix C. We note that the matrix $M = CC^\dagger$ is a hermitian, positive definite operator and the unitary matrix U can therefore be chosen to diagonalize M. With the (non-negative) eigenvalues written as $|d_n|^2$ we have

$$
\langle \tilde{n}|\tilde{m}\rangle_B = |d_n|^2 \delta_{mn}
\tag{2.47}
$$

and if the normalized vectors $|m\rangle$ are introduced by

$$
|\tilde{m}\rangle = d_m |m\rangle
\tag{2.48}
$$

then the Schmidt form (2.40) of the vector $|\chi\rangle$ is reproduced with the basis sets of both systems A and B as orthonormal vectors.

2.2.3   Entanglement in a two-spin system

To exemplify this we consider the simplest possible composite system, a system that consists of two two-level subsystems (for example two spin-half systems). We shall consider several sets of vectors with different degrees of entanglement.

An orthogonal basis of product states is given by the four vectors

$$
\begin{aligned}
|\uparrow\uparrow\rangle &= |\uparrow\rangle_A \otimes |\uparrow\rangle_B \\
|\uparrow\downarrow\rangle &= |\uparrow\rangle_A \otimes |\downarrow\rangle_B \\
|\downarrow\uparrow\rangle &= |\downarrow\rangle_A \otimes |\uparrow\rangle_B \\
|\downarrow\downarrow\rangle &= |\downarrow\rangle_A \otimes |\downarrow\rangle_B
\end{aligned}
\tag{2.49}
$$

where $\{|\uparrow\rangle, |\downarrow\rangle\}$ is an orthonormal basis set for the spin-half system.

Another basis is given by those with well-defined total spin. As is well-known by the rules of addition of angular momentum, two spin half systems will have states of total spin 0 or 1. The spin 0 state is the antisymmetric (spin singlet) state

$$
|0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)
\tag{2.50}
$$

while spin 1 is described by the symmetric (spin triplet) states

$$
|1,\uparrow\rangle = |\uparrow\uparrow\rangle \ , \quad |1,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \ , \quad |1,\downarrow\rangle = |\downarrow\downarrow\rangle
\tag{2.51}
$$

Clearly the two states $|1,\uparrow\rangle = |\uparrow\uparrow\rangle$ and $|1,\downarrow\rangle = |\downarrow\downarrow\rangle$ are product states with no correlation between the two spin systems. However, the two remaining states $|0\rangle$ and $|1,0\rangle$ are entangled. These states may be included as two of the states of a third basis of orthonormal states, called the *Bell states*. They are defined by

$$
\begin{aligned}
|a,\pm\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle \pm |\downarrow\uparrow\rangle) \\
|c,\pm\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\uparrow\rangle \pm |\downarrow\downarrow\rangle)
\end{aligned}
\tag{2.52}
$$

and are all states of *maximal entanglement* between the two subsystems. We note that the two spins are strictly *anticorrelated* for the first two states ($|a,\pm\rangle$) and strictly *correlated* for the two other states ($|c,\pm\rangle$).

Let us focus on the spin singlet state $|a,-\rangle$. It is a pure state (of system A+B) with density matrix

$$
\begin{aligned}
\hat{\rho}(a,-) &= |a,-\rangle\langle a,-| \\
&= \frac{1}{2}(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow| - |\uparrow\downarrow\rangle\langle\downarrow\uparrow| - |\downarrow\uparrow\rangle\langle\uparrow\downarrow|)
\end{aligned}
\tag{2.53}
$$

The corresponding reduced density matrices of both systems A and B are of the same form

$$\hat{\rho}^A = \hat{\rho}^B = \frac{1}{2}(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|) \tag{2.54}$$

The symmetric state $\hat{\rho}(a, +)$ has the same reduced density matrices as $\hat{\rho}(a, -)$. In fact this is true for *all* the four Bell states. Thus, the information contained in the density matrices of the subsystems do not distinguish between these four states of the total system.

The loss of information when we consider the reduced density matrices is demonstrated explicitly in taking the partial trace of (2.53) with respect to subsystem A or subsystem B. The two last terms in (2.53) simply do not contribute. If we leave out the two last terms of (2.53) we have the following density matrix, which also have the same reduced density matrices,

$$\hat{\rho}(a) = \frac{1}{2}(|\uparrow\downarrow\rangle\langle\uparrow\downarrow| + |\downarrow\uparrow\rangle\langle\downarrow\uparrow|) \tag{2.55}$$

This is still strictly anticorrelated in the spin of the two particles, but the correlation is now *classical* in the sense that the full density matrix is written in the form (2.36). Thus, the terms that are important for the quantum entanglement between the two subsystems are the ones that are left out when we take the partial trace. These terms are the off-diagonal *interference terms* of the full density matrix, and this means that we can view the entanglement as a special type of interference effect associated with the composite system.

When the reduced density operators are written as 2×2 matrices they have the form

$$\rho^A = \rho^B = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.56}$$

Thus they are proportional to the 2×2 unit operator and therefore correspond to states with maximal entropy $S_A = S_B = \log 2$. With the entropy of the subsystems taken as a measure of entanglement this means that the spin singlet state is a state of *maximal entanglement*. This is true for all the Bell states, as already has been mentioned.

Finally, note that the reduced density matrices (3.8) are *rotationally invariant*. This is however not the case for the "classical" density matrix (2.55) of the full system. It describes a state where the spins of the two particles *along one particular direction* is strictly anticorrelated. It is interesting to note that we cannot define a density matrix of this form which predicts strict anticorrelation for the spin in

any direction. Also note that neither the pure state $|a, +\rangle$ is rotationally invariant. In fact the state $|a, -\rangle$ is the only Bell state that is rotationally inavriant, since it corresponds to *spin 0*. This makes the spin singlet state particularly interesting, since it describes a situation where the spins are strictly anticorrelated *along any direction in space*.

2.3 Quantum states and physical reality

In 1935 Einstein, Podolsky and Rosen published a paper where they focussed on a central question in quantum mechanics. This question has to do with the relation between the the *formalism of quantum mechanics*, with its probability interpretation, and the underlying *physical reality* that quantum mechanics describes. They pointed out, by way of a simple thought experiment, that one of the implications of quantum physics is that it blurs the distinction between the *objective reality of nature* and the *subjective description* used by the physicist. Their thought experiment has been referred to as the EPR paradox, and it has challenged physicists in their understanding of the relation between physics and the reality of natural phenomena up to this day. Einstein strongly suggested that the paradox showed that quantum mechanics is an *incomplete* theory of nature. Later, in 1964 John Bell showed that the conflict between quantum mechanics and intuitive notions of reality goes deeper. Quantum theory allows types of correlations that cannot be found in classical theories that obey the basic assumptions of *locality* and *reality*.

In this section we examine the EPR paradox in a form introduced by David Bohm and proceed to examine how the limitations of classical theory is broken in the form of Bell inequalities. As first emphasized by Erwin Schrödinger the basic element of quantum physics involved in these considerations is that of *entanglement*.

2.3.1 EPR-paradox

We consider a thought experiment, where two spin half particles (particle A and particle B) are produced in a singlet state (with total spin zero). The two particles move apart, but since they are considered to be well separated from any disturbance, they keep their correlation so that the spin state is left unchanged. The full wave function of the two-particle system can be viewed as the product of a spin state and a position state, but for our purpose only the spin state is of interest. Concerning the position state it is sufficient to know that after some time the

particles are separated by a large distance.

At a point in time when the two particles are well separated, a spin measurement is performed on particle A. This will modify the spin state of this particle, but since particle B is far away, *the measurement cannot affect particle B in any real sense*. However, and this is the paradox, *the change in the spin state caused by the measurement will also influence the possible outcomes of spin measurements performed on particle B.*

The experiment is outlined in Fig.(2.3.1). From the time when the particles are emitted and until the time when the spin experiment is performed the two particles are in the entangled state

$$|a, -\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{2.57}$$

In this state the spin of the two particles are strictly *anticorrelated*. This means that if particle A is measured to be in a spin up state, particle B is necessarily in a spin down state. But each particle, when viewed separately, is with equal probability found with spin up and spin down.

The reference to *spin up* ($\uparrow$) and *spin down* ($\downarrow$) in the state vector (2.57) seems to indicate that we have given preference to some particular direction in space. However, the singlet state is *rotationally invariant*. Therefor the state is left unchanged if we redefine this direction in space. Thus, we do not have to specify whether the z-axis, the x-axis or any other direction has been chosen, they all give rise to the same (spin 0) state.

We now consider, in this hypothetical experiment, that the spin of particle is measured along the z-axis. If the result is *spin up*, we know with certainty that the spin of particle B in the z-direction is *spin down*. Likewise, if the result of measurement on A is *spin down* the spin of particle B in the z-direction is *spin up*. In both cases, a measurement of the z spin component of particle A will with necessity project particle B into a state with well defined (quantized) z-component of the spin. Since no real change can have been introduced in the state of particle B (it is far away) it seems natural to conclude that the spin component along the z-axis must have had a sharp (although unknown value) also *before* the measurement was actually performed on particle A.

This conclusion is on the other hand in conflict with the standard interpretation of quantum mechanics. If the the z-component of the spin of particle A has a sharp value before the measurement, it will have a sharp value even if the measurement along the z-axis is not performed at all, and even if we choose to
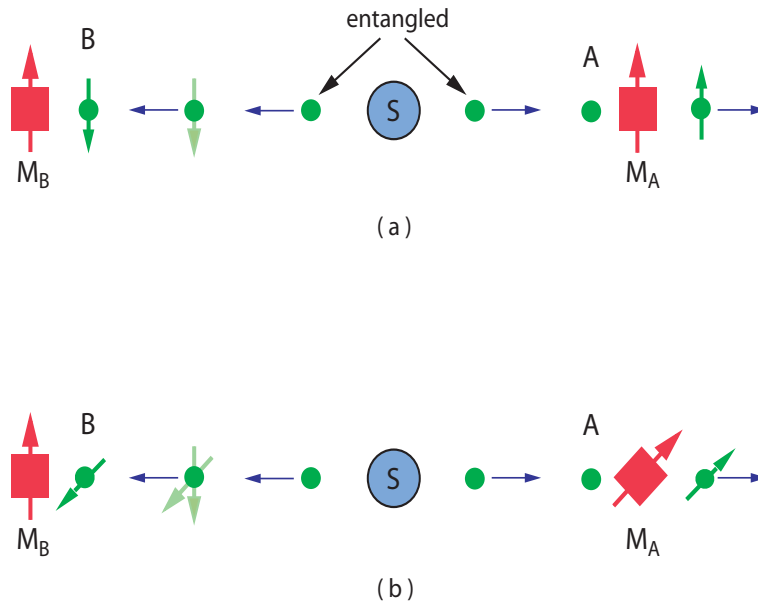
Figure 2.2: The set-up of an EPR experiment. Two entangled particles in a singlet spin state is sendt in opposite directions from a source S. The particles are shown at three different times. The spin direction of each particle is, in this initial state, totally undetermined. In Fig (a) the z-component of the spin of particle A is measured by a measuring device M_A at a time t_{M_A}. Since the spins of the two particles are strictly anticorrelated this measurement will project not only particle A, but also particle B into a state of quantized spin in the z-direction. If one assumes (like EPR) that no real change has taken place at the position of particle B, one may be tempted to conclude that particle B was in reality in this spin state also before the measurent. In Fig (b) the measurement by M_A is changed to a rotated direction. Repeating the argument would lead to the conclusion that particle B had a well-defined (although unknown) spin component in this direction before the measurement. But *before* the measurement one has a freedom of choice in what direction to perform the measurement, either in the direction of (a) or the direction of (b). This seems to indicate that both spin components have to be well defined before the measurement. But the two components of the spin are *incompatible* observables, hence the paradox.

measure the spin along the x-axis instead. But now the argument can be repeated for the x-component of the spin. This will in the same way lead to the conclusion that the x-component of the spin of particle has a sharp value before the measurement. Consequently, both the z-component and the x-component of the spin must have a sharp value before the measurement is performed on particle A. But this is not in accordance with the standard interpretation of quantum mechanics. The z-component and the x-component of the spin operator are *incompatible observables*, since they do not commute as operators. Therefore they cannot simultanuously be assigned sharp values.

The conclusion Einstein, Podolsky and Rosen drew from the thought experiment is that both observables will in reality have sharp values, but since quantum mechanics is not able to predict these values with certainty (only with probabilities) quantum theory cannot be a complete theory of nature. From their point of view there seem to be some missing variables in the theory. If such variables are introduced, we refer to the theory as a "hidden variable" theory.

Note the two elements that enter into the above argument:
Locality — Since particle B is far away from where the measurement takes place we conclude that no real change can take place concerning the state of particle B.
Reality — When the measurement on particle A makes it possible to predict *with certainty* the outcome of a measurement performed on particle B, we draw the conclusion that this represents a property of B which is there even without actually performing a measurement on B.

From later studies we know that the idea of Einstein, Podolsky and Rosen that quantum theory is an incomplete theory does not really resolve the EPR paradox. Hidden variable theories can in principle be introduced, but to be consistent with the predictions of quantum mechanics, they cannot satisfy the conditions of locality and reality used by EPR. Thus, there is a real *conflict* between quantum mechanics and basic principles of classical theories of nature.

From the discussion above it is clear that *entanglement* between the two particles is the source of the (apparent) problem. It seams natural to draw the conclusion that if a measurement is performed on one of the partners of an entangled system there will be (immediately) a change in the state also of the other partner. But if we phrase the effect in this way, one should be aware of the fact that the effect of entanglement is always hidden in correlations between measurements performed on the two parts. The measurement performed on particle A does not lead to any local change at particle B that can be seen without consulting the result of the measurement performed on A. Thus, no change that can be interpreted as a signal sent from A to B has taken place. This means that the immediate

change that is affecting the total system A+B due to the measurement on A does not introduce results that are in conflict with causality.

2.3.2 Bell's inequality

Correlation functions

The EPR paradox shows that quantum mechanics is radically different from classical statistical theories when entanglement between systems is involved. In the original presentation of Einstein, Podolsky and Rosen it was left open as a possibility that a more complete description of nature could replace quantum mechanics. However, by studying correlations between spin systems John Bell concluded in 1964 that the predictions of quantum mechanics is in direct conflict with what can be deduced from a "more complete" theory with hidden variables, when this satisfy what is known as *Einstein locality*. This was expressed in terms of a certain inequality (Bell inequality) that has to be satisfied by classical (local, realistic) theories. Quantum theory does not satisfy this inequality.

To examine this difference between classical and quantum correlations we reconsider the EPR experiment, but focus on correlations between measurements of spin components performed on both particles A and B. Let us assume that the particles move in the x-direction. The spin component in a rotated direction in the y, z-plane will have the form

$$\hat{S}_\theta = \cos\theta \hat{S}_z + \sin\theta \hat{S}_x \tag{2.58}$$

We consider now the following correlation function between spin measurements on the two particles,

$$E(\theta_A, \theta_B) \equiv \frac{4}{\hbar^2} \langle S_{\theta_A} S_{\theta_B} \rangle \tag{2.59}$$

Thus, we are interested in correlations between spin directions that are rotated by *arbitrary* angles θ_A and θ_B in the y, z-plane. The factor $\hbar^2/4$ is included to make $E(\theta_A, \theta_B)$ a dimensionless function, normalized to take values in the interval $(-1, +1)$.

For the spin singlet state it is easy to calculate this correlation function. It is

$$E(\theta_A, \theta_B) = -\cos(\theta_A - \theta_B) \tag{2.60}$$

With the two spin measurements oriented in the same direction, $\theta_A = \theta_B \equiv \theta$, we have

$$E(\theta, \theta) = -1 \tag{2.61}$$

which shows the anticorrelation between the two spin vectors.

Let us now leave the predictions of quantum mechanics and look at the limitations of a *local, realistic hidden variable theory*. We write the measured spin values as $S_A = \frac{\hbar}{2}\sigma_A$ and $S_B = \frac{\hbar}{2}\sigma_B$ so that σ_A and σ_B are normalized to unity in absolete value. We make the following assumptions

1. *Hidden variables.* The measured spin values S_A of particle A and S_B of particle B are determined (before the measurement) by one or more unknown varlables, denoted λ. Thus $\sigma_A = \sigma_A(\lambda)$ and $\sigma_B = \sigma_B(\lambda)$.

2. *Einstein locality*. The spin value S_A will depend on the orientation θ_A, but not on the orientation θ_B, when the two measuring devices are separated by a large distance. Similarly S_B will be independent of θ_A. We write this as

$$\sigma_A = \sigma_A(\theta_A, \lambda) \quad \sigma_B = \sigma_B(\theta_B, \lambda)$$

3. *Spin quantization*. We assume the possible outcome of measurements are consistent with spin quantization,

$$\sigma_A = \pm 1 \quad \sigma_B = \pm 1$$

4. *Anticorrelation*. We assume that the measured values are consistent with the known anticorrelation of the singlet state

$$\sigma_A(\theta, \lambda) = -\sigma_B(\theta, \lambda) \equiv \sigma(\theta, \lambda)$$

Since the variable λ is unknown we assume that particles are emitted from the source S with some probability distribution $\rho(\lambda)$ over λ. Thus, the expectation value (2.59) can be written as

$$
\begin{aligned}
E(\theta_A, \theta_B) &= \int d\lambda \, \rho(\lambda) \, \sigma_A(\theta_A, \lambda) \, \sigma_B(\theta_B, \lambda) \\
&= -\int d\lambda \, \rho(\lambda) \, \sigma(\theta_A, \lambda) \, \sigma(\theta_B, \lambda) \quad (2.62)
\end{aligned}
$$

We now consider the difference between correlation function E for two different set of angles

$$
\begin{aligned}
E(\theta_1, \theta_2) - E(\theta_1, \theta_3) &= -\int d\lambda \, \rho(\lambda)(\sigma(\theta_1, \lambda)\sigma(\theta_2, \lambda) - \sigma(\theta_1, \lambda)\sigma(\theta_3, \lambda)) \\
&= \int d\lambda \, \rho(\lambda)\sigma(\theta_1, \lambda)\sigma(\theta_2, \lambda) \, (\sigma(\theta_2, \lambda)\sigma(\theta_3, \lambda) - 1)
\end{aligned}
$$

$$(2.63)$$

In the last expression we have used $\sigma(\theta, \lambda)^2 = 1$, which follows from assumption 3. From this we deduce the following inequality

$$
\begin{aligned}
|E(\theta_1, \theta_2) - E(\theta_1, \theta_3)| &\leq \int d\lambda \, \rho(\lambda) \, (1 - \sigma(\theta_2, \lambda)\sigma(\theta_3, \lambda)) \\
&= 1 + E(\theta_2, \theta_3)
\end{aligned} \tag{2.64}
$$

which is one of the forms of the *Bell inequality*.

Let us make the further assumption that the correlation function is only dependent on the relative angle (rotational invariance),

$$
E(\theta_1, \theta_2) = E(\theta_1 - \theta_2) \tag{2.65}
$$

Let us also assume that $E(\theta)$ increases monotonically with θ in the interval $(0, \pi)$. (Both these assumptions are true for the quantum mechanical correlation function (2.60).) We introduce the probability function

$$
P(\theta) = \frac{1}{2}(E(\theta) + 1) \tag{2.66}
$$

This function gives the probability for measuring either *spin up* or *spin down* on *both spin measurements* (see the discussion below). It satisfies the simplified inequality

$$
P(2\theta) \leq 2P(\theta) \,, \quad 0 \leq \theta \leq \frac{\pi}{2} \tag{2.67}
$$

Due to strict anticorrelation between the spin measurements for $\theta = 0$, which means strict correlation for $\theta = \pi$, the function satisfies the boundary conditions

$$
P(0) = 0 \,, \quad P(\pi) = 1 \tag{2.68}
$$

The inequality (2.67) further implies that $P(\theta)$ is a *concave function* with these boundary values. This is illustrated in Fig.(2.3.2), where $P(\theta) = \theta/\pi$ is a limiting function which satisfies (2.67) *with equality* rather than inequality.

We now turn to the predictions of quantum theory. As follows from (2.60) the function $P(\theta)$ is given by

$$
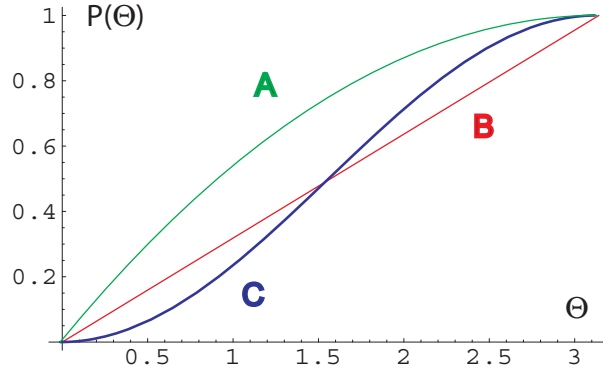P(\theta) = \sin^2 \frac{\theta}{2} \tag{2.69}
$$

Figure 2.3: The probability $P(\theta)$ for measuring two *spin up*s or two *spin down*s as a function of the relative angle θ between the spin directions. Curve A satisfies the Bell inequality, Curve B is the limiting case where the Bell inequality is replaced by equality and Curve C is the probability predicted by quantum mechanics. Curve C is not consistent with the Bell inequality, but is confirmed by experiment.

This function does not satisfy the Bell inequality. This is clear from taking a special value $\theta = \pi/3$. We have

$$
\begin{aligned}
P(2\pi/3) & = \sin^2 \frac{\pi}{3} = \frac{3}{4} \\
2P(\pi/3) & = 2\sin^2 \frac{\pi}{6} = \frac{1}{2}
\end{aligned}
\tag{2.70}
$$

Thus, $P(2\pi/3) > 2P(\pi/3)$ in clear contradiction to Bell's inequality. This breaking of the inequality is seen also in Fig.(2.3.2).

Measured frequencies

To get a better understanding of the physical content of the Bell inequality we will discuss its meaning in the context of spin measurements performed on the two particles, where a series of results *spin up* or *spin down* is registered for each particle. We are particularly interested in correlations between the results of pairs of entangled particles.

    Let us therefore consider a spin measurement experiment where the orientations of the two measurement devices are set to fixed directions. We assume N pairs of entangled particles are used for the experiment. The results are registered

for each pair, and we denote by n_+ the number of times where the two particles are registered both with *spin up* or both with *spin down*. The number of times where one of the particles are registered with *spin up* and the other with *spin down* we denote by n_-. Clearly $n_+ + n_- = N$.

The functions $E(\theta)$ and $P(\theta)$ are for large N approximated by the frequencies

$$
\begin{aligned}
E(\theta)_{exp} &= \frac{n_+ - n_-}{N} \\
P(\theta)_{exp} &= \frac{1}{2}(E(\theta)_{exp} + 1) = \frac{n_+}{N}
\end{aligned}
\tag{2.71}
$$

Thus $P(\theta)$ represents the probability for having the same result (both *spin up* or both *spin down*) for an entangled pair, as already mentioned.

Let us first focus at the list of results called *Series I* in Fig.(2.3.2). We consider this list as the outcome of a series of measurements for a setup where the directions of the two measuring devices are aligned ($\theta_A = \theta_B = 0$). Each pair of entangled particles is separated in a *particle A* whose spin is measured in M_A and a *particle B* whose spin is measured in M_B.

The list indicates, as we should expect, that the results of measurements at M_A, when considered separately, correspond to *spin up* and *spin down* with equal probability. The same is the case for the meaurements at M_B. The list also reveals the strict *anticorrelation* between measurements on each entangled pair. We write the result as

$$
P(I)_{exp} = \frac{n_+(I)}{N} = 0
\tag{2.72}
$$

Instead of performing a new series of measurements we next make some theoretical considerations based on the assumption of *locality*. We ask the question: *What would have happened if in the experiment performed the spin detector M_B had been rotated to another direction $\theta_B = \pi/3$.* Locality is now interpreted as meaning that if M_B were rotated, that could in no way have influenced the results at M_A. The series of measurements S_A would therefore have been unchanged. The results S_B would however have to change since the results for $\theta \neq 0$ are not strictly anticorrelated. Thus, some of the results would be different compared to those listed in Series I. Quantum mechanics in this case predicts $P(\pi/3) = 1/4$. In the list of Series II we have indicated a possible outcome consistent with this, where

$$
P(II)_{exp} = \frac{n_+(II)}{N} = \frac{3}{12}
\tag{2.73}
$$

Figure 2.4: Four series with results of spin measurement on pairs of entangled particles (S_A and S_B). *Spin up* is represented by + and *spin down* by - . *Series I* is considered as the result of a real expeiment with alignment of the directions of the measuring devices M_A and M_B. A strict anticorrelation between the result is observed for each pair. *Series II* gives the hypothetical results that would have been obtained in the same experiment if the direction of M_B were tilted relative to M_A. About $1/4$ of the spins S_B would be flipped. These are indicated by red circles. This would lead to four pairs with correlated spins, indicated by green lines. *Series III* is a similar hypothetical situation with M_A rotated. Finally *Series IV* gives the results if both M_A and M_B were rotated. In one of the pairs both spins are flipped relative to *Series I*. This reduces the number of correlated spins relative to the sum of correlated pairs for *Series II* and *Series III*.

Since we have a symmetric situation between M_A and M_B we clearly could have used the same argument if M_A were rotated instead of M_B. This situation is shown in Series III, where the series of measurements S_B is left unchanged, but some of the results of S_A are changed. In this case we have chosen $\theta_A = -\pi/3$. We also in this list have

$$P(III)_{exp} = \frac{n_+(III)}{N} = \frac{3}{12} \tag{2.74}$$

consistent with the predictions of quantum mechanics.

Finally, we combine these two results to a two-step argument for what would have happened if both measurement devices had been rotated, to the directions $\theta_A = -\pi/3$ and $\theta_B = \pi/3$. We can start with either Series II or Series III. Rotation of the second measuring device would then lead to a change of about $1/4$ of the results in the series that was not changed in the first step. The total number of correlated pairs would now be

$$n(IV) = n(II) + n(III) - \Delta \tag{2.75}$$

In this expression $n(II)$ is the number of flips in the result of S_B and $n(III)$ the number of flips in the results of S_A. Such flips would change the result from *anticorrelated* to *correlated*. However that would be true only if *only one* of the spin results of an entangled pair were flipped. If both spins were flipped we would be back to the situation with *anticorrelations*. Thus the number of correlated pairs would be equal to the sum of the number of flips in each series minus the number where two flips is applied to the same pair. The last number is represented by Δ in the formula (2.75). In the result of Series IV this is represented by the result

$$\begin{aligned} P(IV)_{exp} &= \frac{n_+(IV)}{N} = \frac{4}{12} \\ &< P(II)_{exp} + P(III)_{exp} = \frac{6}{12} \end{aligned} \tag{2.76}$$

The subtraction of Δ in (2.75) explains the Bell inequality, which here takes the form

$$P(2\pi/3) \leq 2P(\pi/3) \tag{2.77}$$

The results extracted from the series of spin measurements listed in Fig.(2.3.2) gives (2.76) in accordance with the Bell inequality (2.77).

The arguments given above, which reproduce Bell's inequality, may seem to implement the assumption of *locality* in a very straight forward way: *Changes in the set up of the measurements at M_A cannot influence the result of measurements performed at M_B* when these devices are separated by a large distance. The predictions of quantum mechanics, on the other hand, are not consistent with the results of these hypothetical experiments, and *real experiments* have confirmed quantum mechanics rather than the Bell inequality.

So is there anything about the arguments given for the hypothetical experiments that indicates that they may be wrong? We note at least one disturbing fact. Only one of the four series of results in Fig.(2.3.2) can be associated with a *real* experiment. The others must be based on assumptions of what would have happened if *the same series of measururements* would have been performed under somewhat different conditions. This is essential for the result. If the four series of measurements should instead correspond to *four different (real) experiments* the situation would be different. We then would have no reason to believe that the series of results S_A would be preserved when going from I to II or S_B when going from I to II. Each series would in this case rather give a new (random) distribution of results for particle A (or particle B).

In any case the conclusion seems inevitable, that quantum mechanics is in conflict with *Einstein locality*, *i.e.*, with the basic assumptions that lead to the Bell inequality. Does this mean that *some kind of influence* is transmitted from A to B when measuring on particle A, even when the two particles are very far apart?

2.4   Problems

(To be completed.)

2.4.1   Density matrix and spin orientation

a) Write up the general expression for the (2×2) density matrix ρ of a spin half system. (It should satisfy the conditions (2.5)). Show that it can be interpreted as representing a statistical ensemble of *spin up* states and *spin down* states in some direction $\mathbf{n}$. Find the probabilities p_+ for *spin up* and p_- for *spin up* as well as the unit vector $\mathbf{n}$ expressed in terms of the matrix elements of ρ. Is the vector $\mathbf{n}$ uniquely determined? Discuss the special case where the vector is completely undetermined.

b) Assume the spin-half system is in a *pure state*. Find the density matrix (or state vector) expressed in terms of the expectation values $\langle S_z \rangle$ and $\langle S_x \rangle$. (Is information about $\langle S_y \rangle$ also needed to specify the state?)

c) Assume the spin-half system is in a *mixed state*. Consider the same question as under b).

2.4.2 Spin 1 system

Find the number of parameters that determine the general density matrix of a spin 1 system. How many parameters are needed if the system is in a pure state? Assume the expectation values $\langle S_x \rangle$, $\langle S_y \rangle$ and $\langle S_z \rangle$ are known. What additional information is needed in the two cases in order to specify the state?

2.4.3 Convexity

Density matrices do not satisfy the superposition principle like the state vectors. They do however satisfy a *convexity criteria* in the following form. If $\hat{\rho}_1$ and $\hat{\rho}_2$ are two density matrices, also the following linear combination is a density matrix,

$$\hat{\rho} = \alpha \hat{\rho}_1 + (1 - \alpha)\hat{\rho}_2 \tag{2.78}$$

with α as a real number in the interval $0 \le \alpha \le 1$. Show this. Can a *pure* state be written as such a combination of two other density matrices?

2.4.4 Schmidt decomposition

A composite system consists of two two-level systems. The Hilbert space of the composite system is spanned by the four vectors

$$|00\rangle = |0\rangle \otimes |0\rangle \,, \ |01\rangle = |0\rangle \otimes |1\rangle \,, \ |10\rangle = |1\rangle \otimes |0\rangle \,, \ |11\rangle = |1\rangle \otimes |1\rangle \tag{2.79}$$

Find the Schmidt decomposition of the following three vectors

$$
\begin{aligned}
|a\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \\
|b\rangle &= \frac{1}{\sqrt{3}}(|00\rangle + |01\rangle + |10\rangle) \\
|c\rangle &= \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)
\end{aligned}
$$

$$\tag{2.80}$$