Quantum Entanglements

Rainer F. Hauser
rainer.hauser@gmail.com

October 4, 2019

Abstract

Quantum entanglement is an important concept in modern physics. Leonard Susskind from the Stanford University held nine lectures in 2006 covering this topic embedded in a brilliant introduction to basic quantum physics. His lectures were available on YouTube at the time this transcript has been assembled and may still be available as "Quantum Entanglements (Part 1)" today. They need minimum knowledge of mathematics and use Hilbert spaces with finite dimensions only such that linear operators can be represented by matrices, but some knowledge of complex numbers and vector spaces is required. The notation is not always the same as in the videos to emphasize that abstract concepts are important and not specific notations.

1 Physical Systems

1.1 Systems in Classical Physics

Information in classical physics is represented in bits. A *bit* is usually represented either by 0 or by 1, but could be seen as anything with two distinct states. Flipping a coin, for example, leads to the two possible states head or tail. Everything in physics can be represented by bits when approximated. To digitize a field, physical space has to be digitized as a lattice. Values can be represented by bits to any precision needed. Otherwise, digital computers could not be used to simulate physical processes.

One way of describing a system in classical physics is by representing the states in bits and the laws of updating as transitions from one state to another in discrete steps. One can draw states as dots and transitions as arrows leading from one dot to another. In computer science, such systems are called finite-state machines. Physics is deterministic, past and future are unique, and therefore there are never two or more arrows from the same node or to the same node. In other words, there is no loss of information. One can go forward in time and backward in time in a completely deterministic way.

Classical physics is based on Boolean logic and set theory. Vectors and matrices can be used to update configurations. In classical physics, this can be represented in bits as a sequence of zeros and ones. In quantum physics, this is more complicated.

1.2 Systems in Quantum Mechanics

The concept corresponding to the bit of Boolean logic is the *qubit* in the logic of quantum physics. An example for representing a qubit is using the possible states of the spin of an electron represented as vectors. The vectors in quantum physics are not the vectors in the three-dimensional space of classical physics or the four-dimensional space-time in the theory of relativity.

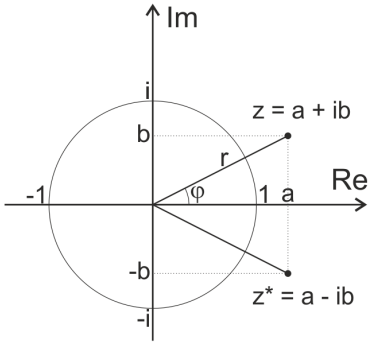
States in quantum physics are represented by vectors in an abstract vector space. Experiments require two different steps. The system is prepared in a certain state in the first step and is measured in the second step. To prepare the spin of an electron, the electron can be placed into a magnetic field. The electron may radiate a photon. In classical physics, the radiation depends on the angle, but in quantum

physics, either nothing happens or one photon is radiated. If the experiment is repeated, the energy of the photon is always the same.

The vectors with which the states of the spin of an electron can be written are $|0\rangle$ and $|1\rangle$, but could also be called $|u\rangle$ for the spin pointing up and $|d\rangle$ for the spin pointing down, similar to the bit in classical physics. The general state for the spin of an electron can be written in the form $a_u |u\rangle + a_d |d\rangle$ with two complex numbers a_u and a_d. This is called a *superposition* of the two possible states $|u\rangle$ and $|d\rangle$.

2 Principles of Quantum Mechanics

2.1 Hilbert spaces



The abstract vector space over the complex numbers $\mathbb{C}$ used for the representation of the states of a physical system in quantum mechanics is called a *Hilbert space*. Vectors are so-called ket-vectors $|\alpha\rangle$, and $z|\alpha\rangle$ for $z \in \mathbb{C}$ is also a vector. Complex numbers can be written as $z = a + ib$ with their *complex conjugate* $z^* = a - ib$ for $a, b \in \mathbb{R}$. The product $z^*z = zz^* = (a+ib)(a-ib) = a^2+b^2$ is a real number because $i^2 = -1$. Complex numbers can also be represented as $z = re^{i\varphi}$, where $r = \sqrt{a^2+b^2}$ is the length of $z = a + ib$ in a coordinate system for which the real part a is drawn in the x-direction usually called Re and the imaginary part b in the y-direction usually called Im, and where φ is the angle. The so-called *phase* $e^{i\varphi}$ can be written as $e^{i\varphi} = \cos(\varphi) + i\sin(\varphi)$ and is a point on the unit circle. The complex conjugate of $z = re^{i\varphi}$ is $z^* = re^{-i\varphi}$

Every Hilbert space has a dual space with the so-called bra-vectors $\langle\beta|$. A ket-vector $|\alpha\rangle$ and a bra-vector $\langle\beta|$ can be multiplied using the *inner product* $\langle\beta|\alpha\rangle$. The Hilbert space for the spin of an electron, for example, is two-dimensional, and

$$|\alpha\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \langle\alpha| = (a_1^*, a_2^*) \quad |\beta\rangle = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad \langle\beta| = (b_1^*, b_2^*) \quad \langle\beta|\alpha\rangle = (b_1^*, b_2^*) \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = b_1^* a_1 + b_2^* a_2$$

with $a_1, a_2, b_1, b_2 \in \mathbb{C}$ show examples for bra- and ket-vectors as well as the inner product in this Hilbert space. For higher dimensions, the vectors just have more components. Vectors representing states of a physical system are *normalized* meaning $\langle\alpha|\alpha\rangle = 1$.

As a vector space, a Hilbert space has a basis and a dimension. A basis $|\psi_n\rangle$ is called *orthonormal* if $\langle\psi_n|\psi_m\rangle = \delta_{nm}$ where δ_{mn} is the Kronecker delta which is 0 for $n \neq m$ and 1 for $n = m$. Two states $|\alpha\rangle$ and $|\beta\rangle$ correspond to the same state if they are equal except for a phase, $|\beta\rangle = e^{i\vartheta}|\alpha\rangle$. For example, the two vectors $|\alpha\rangle$ and $-|\alpha\rangle = (-1)|\alpha\rangle = e^{i\pi}|\alpha\rangle$ represent the same state.

2.2 Linear Operators

Every linear operator acting on vectors of a Hilbert space with a finite dimension can be represented as a *matrix* $\mathbf{M}$. In a two-dimensional Hilbert space, the equation $\mathbf{M}|\alpha\rangle = |\beta\rangle$ is

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \qquad |\alpha\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \qquad |\beta\rangle = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} m_{11}a_1 + m_{12}a_2 \\ m_{21}a_1 + m_{22}a_2 \end{pmatrix}$$

in components where $m_{ij} \in \mathbb{C}$. In general, a $n \times n$ matrix contains the n^2 elements m_{ij}. Multiplication of a vector with a number c can be represented as the multiplication with a matrix whose diagonal elements m_{ii} are c, and all the others are 0.

There are several operations one can apply to a matrix $\mathbf{M}$ with elements m_{ij}. The transposed matrix $\mathbf{M}^T$ has the elements $m_{ij}^T = m_{ji}$ and is the image of the original matrix mirrored about the diagonal with the elements m_{ii}. Another operation is building the complex conjugate $\mathbf{M}^*$ with the elements m_{ij}^*. The Hermite conjugate $\mathbf{M}^\dagger$ is the transposed complex conjugate $(\mathbf{M}^T)^*$ of the original matrix and has the

elements $m^\dagger_{ij} = m^*_{ji}$. Correspondingly, there are matrices with specific properties. A *symmetric* matrix satisfy $\mathbf{M}^T = \mathbf{M}$ respectively $m_{ij} = m_{ji}$, and an *antisymmetric* matrix satisfies $\mathbf{M}^T = -\mathbf{M}$ respectively $m_{ij} = -m_{ji}$. A matrix with the property $\mathbf{M}^\dagger = \mathbf{M}$ respectively $m_{ij} = m^*_{ji}$ is called *Hermitian*.

Hermitian operators or matrices play an important role in quantum physics because they represent so-called *observables* and thus measurable values. The diagonal elements fulfill $m_{ii} = m^*_{ii}$ and must therefore be real numbers. A 2×2 Hermitian matrix has the general form

$$\mathbf{M} = \begin{pmatrix} a & b+ic \\ b-ic & d \end{pmatrix}$$

where the four values a, b, c, d are real numbers..

Given a vector $|\alpha\rangle$ and a matrix $\mathbf{M}$, the matrix can act on the vector as $\mathbf{M}|\alpha\rangle = |\alpha'\rangle$ and produces the vector $|\alpha'\rangle$. The same process in the dual space is written as $\langle\alpha|\,\mathbf{M}^\dagger = \langle\alpha'|$. Because the inner product of two vectors $|\alpha\rangle$ and $|\beta\rangle$ satisfies $\langle\alpha|\beta\rangle = \langle\beta|\alpha\rangle^*$, the equation $\langle\beta|\mathbf{M}|\alpha\rangle = \langle\alpha|\mathbf{M}^\dagger|\beta\rangle^*$ holds. This equation becomes $\langle\beta|\mathbf{M}|\alpha\rangle = \langle\alpha|\mathbf{M}|\beta\rangle^*$ for a Hermitian matrix.

The number $\langle\alpha|\mathbf{M}|\alpha\rangle$ is a real number if $\mathbf{M}$ is Hermitian, because $\langle\alpha|\mathbf{M}|\alpha\rangle = \langle\alpha|\mathbf{M}|\alpha\rangle^*$. This value is called the *expectation value* or *average* of the observable $\mathbf{M}$ in state $|\alpha\rangle$. The expectation value can be written as

$$\langle\alpha|\mathbf{M}|\alpha\rangle = (a_1^*, a_2^*)\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = m_{11}a_1^*a_1 + m_{12}a_1^*a_2 + m_{21}a_2^*a_1 + m_{22}a_2^*a_2$$

in components. The result is a real number.

2.3 Probabilities

Quantum mechanics is a mechanism to calculate probabilities. If the normalized state $|\alpha\rangle$ is the superposition $|\alpha\rangle = a_1|\psi_1\rangle + a_2|\psi_2\rangle$, where the $|\psi_n\rangle$ build an orthonormal basis, then $P_1 = a_1^*a_1$ is the probability for the system prepared in the state $|\alpha\rangle$ to be found in state $|\psi_1\rangle$ when measured, and similarly for a_2. The probability to find the system either in state $|\psi_1\rangle$ or $|\psi_2\rangle$ is $P_1 + P_2 = a_1^*a_1 + a_2^*a_2 = \langle\alpha|\alpha\rangle = 1$, because the state $|\alpha\rangle$ is normalized, and therefore the system must assume one of the two states.

In general, if the underlying Hilbert space has dimension d, and the orthonormal basis chosen is $|\psi_n\rangle$ for $n \in \{1, 2, ..., d\}$, any state represented as the normalized vector $|\alpha\rangle$ satisfies

$$|\alpha\rangle = \sum_{n=1}^{d} a_n|\psi_n\rangle \qquad\qquad P_n = a_n^*a_n \geq 0 \qquad\qquad \sum_{n=1}^{d} P_n = 1$$

where the coefficients $a_n = \langle\psi_n|\alpha\rangle$ are called *probability amplitudes*. The values P_n with $0 \leq P_n \leq 1$ are the corresponding probabilities. They do not change when the phase of the state is changed.

2.4 Eigenvectors and Eigenvalues

If there is a value $e \in \mathbb{C}$ and a vector $|\eta\rangle$ such that $\mathbf{M}|\eta\rangle = e|\eta\rangle$ for a matrix $\mathbf{M}$, then e and $|\eta\rangle$ are called an *eigenvalue* and an *eigenvector* of $\mathbf{M}$, respectively. In the following all matrices are assumed to be Hermitian, because only they correspond to the interesting objects representing observables.

If the Hermitian matrix $\mathbf{M}$ has two eigenvectors $|\eta_1\rangle$ and $|\eta_2\rangle$ belonging to the different eigenvalues e_1 and e_2, respectively, then they are orthogonal because from $\langle\eta_2|\mathbf{M}|\eta_1\rangle = e_1\langle\eta_2|\eta_1\rangle$ and $\langle\eta_1|\mathbf{M}|\eta_2\rangle = e_2\langle\eta_1|\eta_2\rangle$ follows $\langle\eta_2|\mathbf{M}|\eta_1\rangle^* = e_2^*\langle\eta_2|\eta_1\rangle^*$ and therefore $\langle\eta_2|\mathbf{M}|\eta_1\rangle = e_2\langle\eta_2|\eta_1\rangle = e_1\langle\eta_2|\eta_1\rangle$. Thus, either $e_1 = e_2$ or $\langle\eta_2|\eta_1\rangle = 0$. If two states are orthogonal, a measurement can distinguish the two states.

The matrix $\mathbf{M}$ corresponding to an observable always allows to build a basis of orthonormal eigenvectors $|\eta_n\rangle$. All the eigenvalues e_n are real, because the matrix $\mathbf{M}$ is diagonal in this basis, and the diagonal elements of a Hermitian matrix are real. If a system is in state $|\eta_n\rangle$ with $\mathbf{M}|\eta_n\rangle = e_n|\eta_n\rangle$, the system is said to be in eigenstate $|\eta_n\rangle$, and a measurement of the corresponding observable gives the value e_n with probability 1. More general, eigenvalues e_n are possible outcomes of a measurement of the observable. If a system is prepared in state $|\alpha\rangle$, then the probability to measure e_n is $\langle\alpha|\eta_n\rangle\langle\alpha|\eta_n\rangle^* = |\langle\alpha|\eta_n\rangle|^2$.

3 The Spin of Electrons

3.1 A Single Spin

The *spin* of an electron or positron can assume two states. Either it eventually emits a photon or it does not when prepared in a magnetic field. Traditionally, the orthonormal basis $\{|u\rangle, |d\rangle\}$

$$|u\rangle = \begin{pmatrix}1\\0\end{pmatrix} \qquad |d\rangle = \begin{pmatrix}0\\1\end{pmatrix} \qquad |r\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix} \qquad |l\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{-1}{\sqrt{2}}\end{pmatrix} \qquad |i\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{i}{\sqrt{2}}\end{pmatrix} \qquad |o\rangle = \begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{-i}{\sqrt{2}}\end{pmatrix} \tag{3.1}$$

is used where $|u\rangle$ is pointing up parallel to the z-axis and $|d\rangle$ is pointing down antiparallel to the z-axis. The other states are pointing right or left in the x-direction respectively in or out in the y-direction, and are expressed in the basis $|u\rangle$ and $|d\rangle$. All these six vectors are normalized. A general state for the spin is $|\alpha\rangle = a_1|u\rangle + a_2|d\rangle$ with $\langle\alpha|\alpha\rangle = a_1^* a_1 + a_2^* a_2 = 1$.

The matrices representing the observable spin in the three directions x, y and z are called *Pauli matrices* denoted by σ_x, σ_y, σ_z or σ_1, σ_2, σ_3, respectively. They are defined as

$$\sigma_3 = \sigma_z = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \qquad \sigma_1 = \sigma_x = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix} \qquad \sigma_2 = \sigma_y = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix} \qquad \sigma_n^2 = \mathbf{I} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} \tag{3.2}$$

and result in the identity matrix $\mathbf{I}$ when squared. The symbol $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ is sometimes used as a kind of three-dimensional vector in ordinary space. These matrices have all the two eigenvalues ± 1, and the eigenvectors are $|u\rangle$ and $|d\rangle$ for σ_z, $|r\rangle$ and $|l\rangle$ for σ_x, and $|i\rangle$ and $|o\rangle$ for σ_y. When measuring the spin of an electron, only the value -1 for a photon to be emitted or $+1$ for no electron being emitted can occur. Note that a Hermitian matrix $\mathbf{M}$ with $\mathbf{M}^2 = \mathbf{I}$ can in general only have eigenvalues ± 1.

3.2 Probabilities for a Spin

Because $\sigma_z|u\rangle = |u\rangle$, the probability to measure 1 for the spin of an electron prepared in state $|u\rangle$ is 1. When the system is prepared in one direction, a measurement in a direction orthogonal to it, is ± 1 with equal probability $\frac{1}{2}$. As an example, if the spin is measured in direction $|r\rangle$ when the system has been prepared in direction $|i\rangle$ or vice versa, the probability is $|\langle r|i\rangle|^2 = \langle r|i\rangle^* \langle r|i\rangle = (\frac{1}{2} - \frac{i}{2})(\frac{1}{2} + \frac{i}{2}) = \frac{1}{2}$.

If the spin of an electron has been prepared along the z-axis and is supposed to be measured in the direction $\vec{n} = (n_1, n_2, n_3)$ with $|\vec{n}| = 1$ in ordinary space,

$$\sigma_{\vec{n}} = \vec{\sigma}\cdot\vec{n} = \sigma_1 n_1 + \sigma_2 n_2 + \sigma_3 n_3 = \begin{pmatrix}n_3 & n_1 - in_2\\n_1 + in_2 & -n_3\end{pmatrix} = \begin{pmatrix}n_3 & n_-\\n_+ & -n_3\end{pmatrix}$$

is the operator for measuring the spin in the direction $\vec{n}$. Because $\sigma_n^2 = \mathbf{I}$ and $\sigma_n\sigma_m = -\sigma_m\sigma_n$, the square of $\sigma_{\vec{n}}$ is $\sigma_{\vec{n}}^2 = n_1^2\sigma_1^2 + n_2^2\sigma_2^2 + n_3^2\sigma_3^2 + n_1n_2(\sigma_1\sigma_2 + \sigma_2\sigma_1) + n_1n_3(\sigma_1\sigma_3 + \sigma_3\sigma_1) + n_2n_3(\sigma_2\sigma_3 + \sigma_3\sigma_2) = \mathbf{I}$ and has therefore only the two possible eigenvalues ± 1 such that the measurable spin along any axis $\vec{n}$ is $+1$ or -1. The values $n_+ = n_1 + in_2$ and $n_- = n_1 - in_2$ satisfy $n_+n_1 = n_1^2 + n_2^2 = (n_1^2 + n_2^2 + n_3^2) - n_3^2 = 1 - n_3^2$ and $n_+^* = n_-$ as will turn out to be useful below.

If the spin is prepared in direction $\vec{n}$ with $|\vec{n}| = 1$ and is measured in direction $\vec{m}$ with $|\vec{m}| = 1$, the two operators $\sigma_{\vec{n}} = \vec{\sigma}\cdot\vec{n}$ and $\sigma_{\vec{m}} = \vec{\sigma}\cdot\vec{m}$ both with eigenvalues ± 1 can be defined. Using $\sigma_{\vec{n}}|\eta_{\vec{n}}\rangle = 1\cdot|\eta_{\vec{n}}\rangle$ and $\sigma_{\vec{m}}|\eta_{\vec{m}}\rangle = 1\cdot|\eta_{\vec{m}}\rangle$ the inner product $|\langle\eta_{\vec{n}}|\eta_{\vec{m}}\rangle|^2$ of the two eigenvectors for eigenvalue $+1$ gives the probability that the spin measured in direction $\vec{m}$ is 1 if it has been prepared such that it is 1 in direction $\vec{n}$. If one does not care about normalization of a two-dimensional ket-vector, one component can be chosen as 1. The equation

$$\vec{\sigma}\cdot\vec{n}\begin{pmatrix}1\\z\end{pmatrix} = \begin{pmatrix}n_3 & n_-\\n_+ & -n_3\end{pmatrix}\begin{pmatrix}1\\z\end{pmatrix} = \begin{pmatrix}n_3 + n_-z\\n_+ - n_3z\end{pmatrix} = \begin{pmatrix}1\\z\end{pmatrix} \Rightarrow z = \frac{1 - n_3}{n_-}$$

allows to determine z and therefore $|\eta_{\vec{n}}\rangle$. After normalizing $|\eta_{\vec{n}}\rangle$ the probability can be calculated as

$$|\langle\eta_{\vec{m}}|\eta_{\vec{n}}\rangle|^2 = \left|\left(\sqrt{\frac{1+m_3}{2}}, \sqrt{\frac{1+m_3}{2}}\frac{1-m_3}{m_+}\right)\begin{pmatrix}\sqrt{\frac{1+n_3}{2}}\\\sqrt{\frac{1+n_3}{2}}\frac{1-n_3}{n_-}\end{pmatrix}\right|^2 = 1 + \frac{\vec{m}\cdot\vec{n}}{2} = \frac{1+\cos\vartheta}{2}$$

and the probability only depends on the angle ϑ between $\vec{m}$ and $\vec{n}$. In addition, the situation is symmetric with respect to $\vec{m}$ and $\vec{n}$.

3.3 Measurements, Commutators and Anticommutators

When one measures a system, one gets a result, but one leaves the system in a new state. Specifically, one gets an eigenvalue of the operator for the measured observable as the result of the measurement and leaves the system in the corresponding eigenvector as the new state. A measurement is therefore a kind of polarizer.

If one measures the spin of an electron using first σ_1 and then σ_2, the system is in an eigenstate of σ_2. Because σ_1 and σ_2 have no common eigenvectors, one cannot measure them simultaneously. If one measures however the spin of an electron using first σ_2 and then σ_1, the system is this time in an eigenstate of σ_1 and not in an eigenstate of σ_2. The fact that σ_1 and σ_2 have no common eigenvectors has the consequence, that one gets mutually exclusive results when doing the measurements in different order.

Thus, the property of two quantities to be measured simultaneously is that they have exactly the same eigenvectors. If two operators $\mathbf{A}$ and $\mathbf{B}$ have all the same eigenvectors such that $\mathbf{A}\,|\eta_n\rangle = a_n\,|\eta_n\rangle$ and $\mathbf{B}\,|\eta_n\rangle = b_n\,|\eta_n\rangle$, then $\mathbf{AB} = \mathbf{BA}$. In other words, the two operators commute and therefore the commutator $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$ is zero. The proof is easy because the eigenvectors build a basis, and the eigenvalues and the operators commute. Numbers and operators always commute.

Sometimes two operators $\mathbf{A}$ and $\mathbf{B}$ have the property $\mathbf{AB} = -\mathbf{BA}$ and therefore anticommute. The anticommutator is written as $\{\mathbf{A}, \mathbf{B}\} = \mathbf{AB} + \mathbf{BA}$. The three Pauli matrices are examples of anticommuting operators because $\sigma_n\sigma_m + \sigma_m\sigma_n = 0$ for $m \neq n$.

4 Entanglements

4.1 Systems of Two Spins

If $|\alpha\rangle = a_1\,|u\rangle + a_2\,|d\rangle$ is a normalized vector corresponding to the state of one spin, then there exists a $\vec{n}$ such that $\vec{\sigma}\cdot\vec{n}\,|\alpha\rangle = |\alpha\rangle$, or in other words, there exists a direction $\vec{n}$ in which the state $|\alpha\rangle$ is an eigenvector of $\vec{\sigma}\cdot\vec{n}$. The normalized state $a_1\,|u\rangle + a_2\,|d\rangle$ has two degrees of freedom, because a_1 and a_2 as complex numbers consist of two real numbers each, but the normalization $a_1^* a_1 + a_2^* a_2 = 1$ and the phase which does not change the state reduce these four real numbers to two.

The effect of the Pauli matrices defined in (3.2) on the states $|u\rangle$ and $|d\rangle$ from (3.1)

$$\begin{array}{lll} \sigma_1\,|u\rangle = |d\rangle & \sigma_2\,|u\rangle = i\,|d\rangle & \sigma_3\,|u\rangle = |u\rangle \\ \sigma_1\,|d\rangle = |u\rangle & \sigma_2\,|d\rangle = -i\,|u\rangle & \sigma_3\,|d\rangle = -\,|d\rangle \end{array} \tag{4.1}$$

can be used for a system of two spins with the four basis vectors $|uu\rangle$, $|ud\rangle$, $|du\rangle$, $|dd\rangle$. The operators acting on the first spin are denoted by σ_n, and the operators acting on the second spin are denoted by τ_n to distinguish them. The three operators σ_n have therefore no effect on the second spin, and the three operators τ_n have no effect on the first spin. The table

$$\begin{array}{lll} \sigma_1\,|uu\rangle = |du\rangle & \sigma_2\,|uu\rangle = i\,|du\rangle & \sigma_3\,|uu\rangle = |uu\rangle \\ \sigma_1\,|du\rangle = |uu\rangle & \sigma_2\,|du\rangle = -i\,|uu\rangle & \sigma_3\,|du\rangle = -\,|du\rangle \\ \sigma_1\,|ud\rangle = |dd\rangle & \sigma_2\,|ud\rangle = i\,|dd\rangle & \sigma_3\,|ud\rangle = |ud\rangle \\ \sigma_1\,|dd\rangle = |ud\rangle & \sigma_2\,|dd\rangle = -i\,|ud\rangle & \sigma_3\,|dd\rangle = -\,|dd\rangle \\ \tau_1\,|uu\rangle = |ud\rangle & \tau_2\,|uu\rangle = i\,|ud\rangle & \tau_3\,|uu\rangle = |uu\rangle \\ \tau_1\,|du\rangle = |dd\rangle & \tau_2\,|du\rangle = i\,|dd\rangle & \tau_3\,|du\rangle = |du\rangle \\ \tau_1\,|ud\rangle = |uu\rangle & \tau_2\,|ud\rangle = -i\,|uu\rangle & \tau_3\,|ud\rangle = -\,|ud\rangle \\ \tau_1\,|dd\rangle = |du\rangle & \tau_2\,|dd\rangle = -i\,|du\rangle & \tau_3\,|dd\rangle = -\,|dd\rangle \end{array} \tag{4.2}$$

shows the effect of the six different operators on the four basis vectors according to (4.1) and describes completely how the σ_n and τ_n act on any vector in the four-dimensional Hilbert space of two spins.

4.2 Two Spins in a Product State

Two electrons in the states $|\alpha\rangle = a_1 |u\rangle + a_2 |d\rangle$ and $|\beta\rangle = b_1 |u\rangle + b_2 |d\rangle$ can be described with a combined state as $|\gamma\rangle = |\alpha\rangle |\beta\rangle = (a_1 |u\rangle + a_2 |d\rangle)(b_1 |u\rangle + b_2 |d\rangle) = a_1 b_1 |uu\rangle + a_2 b_1 |du\rangle + a_1 b_2 |ud\rangle + a_2 b_2 |dd\rangle$. It is called a *product state*. The spin of both electrons are independent and can have different $\vec{n}$. In other words, the two spins in a product state can be polarized in different directions.

If a combined state $|\gamma\rangle$ can be split up into a product $|\alpha\rangle |\beta\rangle$, the state is called *separable*. This is not always possible, because a single spin has two degrees of freedom, and two spins in a separable state have four. Two spins, however, in state $|\gamma\rangle = c_1 |uu\rangle + c_2 |du\rangle + c_3 |ud\rangle + c_4 |dd\rangle$ have six degrees of freedom, because the four complex numbers c_n give eight real numbers reduced by two for the normalization and the phase.

The state $1/\sqrt{2}(|du\rangle + |ud\rangle)$ cannot be separated into $a_1 b_1 |uu\rangle + a_2 b_1 |du\rangle + a_1 b_2 |ud\rangle + a_2 b_2 |dd\rangle$ because either a_1 or b_1 must be zero, but then one of the two components for $|du\rangle$ or $|ud\rangle$ cannot have a non-zero coefficient. It is therefore not a product state of two single spins.

4.3 Two Entangled Spins

Any state in the form $c_1 |du\rangle + c_2 |ud\rangle$, together with many others, is not separable. The two spins are called *entangled*. If one knows that the system of two electrons is in such a state and one knows that one of the two electrons is in state $|u\rangle$, one immediately knows also that the other electron is in state $|d\rangle$. If the two spins stay in this entangled state and get separated, the fact that one knows the state of the second electron when one knows the state of the first one has been interpreted as information flowing faster than the speed of light.

This is, however, not true as an analogy in classical physics suggests. Two balls of the same size but one black and one white are packed such that one does not know the color without unpacking them. The packages are given to two different persons who travel in two different directions without unpacking the balls. They can be lightyears away from each other, but as soon as one of them opens his package and sees a white ball, he immediately knows that the color of the ball in the package of the other person is black, or vice versa. This does not mean that information has flown faster than the speed of light.

The expectation value of σ_1 given $|du\rangle + |ud\rangle$ is $(|du\rangle + |ud\rangle)\sigma_1(|du\rangle + |ud\rangle) = (|du\rangle + |ud\rangle)(|uu\rangle + |dd\rangle)$ and is therefore 0 as also for σ_2 and σ_3. This is impossible if both electrons are independent. There is no direction with $\vec{\sigma} \cdot \vec{n}$ as for a single spin, and the probability in any direction is $\frac{1}{2}$.

To achieve a state $1/\sqrt{2}(|du\rangle \pm |ud\rangle)$ one can bring the two electrons close together, and one of the two electrons will possibly emit an electron to come into this state after some time. Electrons are a kind of magnet, and two magnets tend to align in opposite direction when their magnetic fields interact. This has nothing to do with the Pauli exclusion principle yet.

The sign in $1/\sqrt{2}(|du\rangle \pm |ud\rangle)$ has a meaning. In both cases, the probability is $\frac{1}{2}$, and if one is up, the other is down. Also the expectation value is zero for all σ_n and τ_n. The result for the operators $\sigma_n + \tau_n$ is, however, different. The result for $\sigma_1 + \tau_1$ and $\sigma_2 + \tau_2$ is a non-zero value for the plus sign and zero for the minus sign. The state with the minus sign is called *singlet* state, and the state with the plus sign is called *triplet* state. Another difference is discussed below together with the reason for these names.

4.4 Projection Operators and Subspaces

Vector spaces have the property that the dimension d is the maximum number of linearly independent vectors. Given a vector space with an inner product, d linearly independent vectors can be selected orthonormally, and these d vectors build then an orthonormal basis $|\psi_n\rangle$ with $n = 1, 2, ..., d$. The property of orthonormality can be written as $\langle\psi_m|\psi_n\rangle = \delta_{mn}$ using the Kronecker delta symbol. Any vector $|\alpha\rangle$ in

this vector space can be written as a linear combination of the basis vectors $|\psi_n\rangle$ with some coefficients a_n, and it satisfies

$$|\alpha\rangle = \sum_{n=1}^{d} a_n |\psi_n\rangle \qquad \langle\psi_n|\alpha\rangle = a_n \qquad |\alpha\rangle = \sum_{n=1}^{d} |\psi_n\rangle \langle\psi_n|\alpha\rangle \qquad \sum_{n=1}^{d} |\psi_n\rangle \langle\psi_n| = \mathbf{I}$$

as can be easily checked. The last of these four equation uses the so-called dyad operator $|\alpha\rangle \langle\beta|$ and is a way to write the identity operator $\mathbf{I}$ in the form which is called *resolution of the identity*. The identity operator has, by the way, the property that all vectors are eigenvectors with eigenvalue +1.

Given an observable $\mathbf{K}$ with $\mathbf{K} = e |\eta_1\rangle$ and $\mathbf{K} = e |\eta_2\rangle$ where $|\eta_1\rangle$ and $|\eta_2\rangle$ are linearly independent, the vector $b_1 |\eta_1\rangle + b_2 |\eta_2\rangle$ is also an eigenvector of $\mathbf{K}$ with eigenvalue e. The eigenvalue e is called *degenerate*, and the concept of an eigenvector becomes the concept of an eigenspace where $|\eta_1\rangle$ and $|\eta_2\rangle$ are basis vectors. The two vectors $|uu\rangle$ and $|ud\rangle$, for example, span a subspace where σ_3 is up and the eigenvalue is +1, and they build an orthonormal basis in this subspace.

One can find an orthonormal basis $|\varphi_n\rangle$ in the linear subspace of all the eigenvectors to eigenvalue e, and this subspace has a dimension $1 \leq d' \leq d$. For any vector $|\alpha\rangle$,

$$\mathbb{P}_{\mathbf{K}=e} |\alpha\rangle = \sum_{n=1}^{d'} |\varphi_n\rangle \langle\varphi_n|\alpha\rangle = \begin{cases} |\alpha\rangle & \text{if } |\alpha\rangle \text{ in the subspace} \\ |\alpha'\rangle & \text{otherwise} \end{cases} \tag{4.3}$$

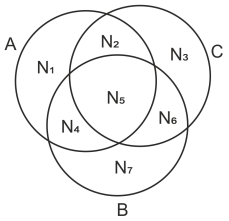
projects $|\alpha\rangle$ onto the subspace. The vector $|\alpha\rangle$ can be split into a sum of two vectors $|\alpha\rangle = |\alpha'\rangle + |\alpha''\rangle$ where $|\alpha'\rangle$ is in the subspace and $|\alpha''\rangle$ is orthogonal to the subspace. The vector $|\alpha'\rangle = \mathbb{P}_{\mathbf{K}=e} |\alpha\rangle$ is called the projection of $|\alpha\rangle$ onto the subspace, and $\mathbb{P}_{\mathbf{K}=e}$ is the corresponding projection operator. The operator $\mathbf{I}$ is the projection operator onto the whole space, and $|\psi_n\rangle \langle\psi_n|$ is the projection operator onto the one-dimensional subspace spanned by $|\psi_n\rangle$.

The probability to measure e with an experiment is given by $\langle\alpha|\mathbb{P}_{\mathbf{K}=e}|\alpha\rangle = \sum_n \langle\alpha|\varphi_n\rangle \langle\varphi_n|\alpha\rangle$ when the system is in state $|\alpha\rangle$. In other words, the expectation value of a projection operator is the probability. The projection operator, for example, for a system with two spins where the first spin is up and the eigenvalue of σ_3 is therefore +1 is $\mathbb{P}_{\sigma_3=1} = |uu\rangle \langle uu| + |ud\rangle \langle ud|$, and $\langle\alpha|uu\rangle \langle uu|\alpha\rangle + \langle\alpha|ud\rangle \langle ud|\alpha\rangle$ is the probability to measure the first spin up when the system is in state $|\alpha\rangle$.

The projection operator $\mathbb{P}_{\mathbf{K}=k} |\alpha\rangle$ gives an eigenvector to the eigenvalue k, and the projection operator $\mathbb{P}_{\mathbf{L}=l} |\beta\rangle$ gives an eigenvector to the eigenvalue l. If $[\mathbb{P}_{\mathbf{K}=k}, \mathbb{P}_{\mathbf{L}=l}] = 0$, the two observables $\mathbf{K}$ and $\mathbf{L}$ can be measured simultaneously, and $\mathbb{P}_{\mathbf{K}=k}\mathbb{P}_{\mathbf{L}=l} |\alpha\rangle = \mathbb{P}_{\mathbf{L}=l}\mathbb{P}_{\mathbf{K}=k} |\alpha\rangle$ returns a vector that is both an eigenvector of $\mathbf{K}$ to eigenvalue k and an eigenvector of $\mathbf{L}$ to eigenvalue l. The operator $\mathbb{P}_{\mathbf{K}=k}\mathbb{P}_{\mathbf{L}=l}$ projects on the intersection of the two subspaces given by $\mathbb{P}_{\mathbf{K}=k}$ and $\mathbb{P}_{\mathbf{L}=l}$. Two spins with operators σ_m and τ_n, for example, are compatible because they belong to two subsystems, and different subsystems are always compatible.

Similarly to classical probability theory, two properties connected by the word "and" result in the product of the projection operators, and two properties connected by the word "or" result in the sum of the operators. Thus, $\mathbb{P}_{\mathbf{K}=k}\mathbb{P}_{\mathbf{L}=l}$ gives the subspace of the eigenvectors for both if the two projection operators $\mathbb{P}_{\mathbf{K}=k}$ and $\mathbb{P}_{\mathbf{L}=l}$ commute.

4.5 Bell's Theorem



Bell's theorem states a rather simple inequality in classical logic or classical set theory. The depth of Bell's inequality is that there are properties A, B and C in quantum mechanics where the inequality is violated. Given are three properties A, B and C. The inequality can be expressed either in numbers or probabilities. In numbers, $N(A, \neg B)$ represents the number of objects with property A but not with property B and analog for the other terms. Bell's inequality

$$N(A, \neg B) + N(B, \neg C) \geq N(A, \neg C) \tag{4.4}$$

can be proved using a Venn diagram with all the $N_i \geq 0$ where $N(A, \neg B) = N_1 + N_2$, $N(B, \neg C) = N_4 + N_7$, and $N(A, \neg C) = N_1 + N_4$. Obviously, $N_1 + N_2 + N_4 + N_7 \geq N_1 + N_4$. To get the statement for probabilities, one has to divide all numbers N_i by the sum over all N_i.

For two spins in a singlet state, the properties are A: "Spin 1 points up along the z-axis", B: "Spin 1 points up along 45°-angle in the zx-plane", and C: "Spin 1 points up along 90°-angle in the zx-plane. Because of the entanglement in the singlet state, negated statement for spin 1 is positive statement for spin 2 such that $\neg B$: "Spin 2 points up along 45°-angle in the zx-plane", and $\neg C$: "Spin 2 points up along 90°-angle in the zx-plane". The two probabilities $P(A, \neg B)$ and $P(B, \neg C)$ are the same because the singlet state does not change when the system is rotated about an angle of 45°. Because $P(A, \neg B)$ is the probability for A and $\neg B$, it is calculated as a product $P(A, \neg B) = P(A)P(\neg B)$.

In classical physics, properties correspond to subsets of all the things, but in quantum mechanics, they correspond to subspaces of a vector space and can therefore be represented by projection operators. The two projection operators

$$\mathbb{P}_{\sigma_3=1}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} \qquad\qquad \mathbb{P}_{\sigma_3=-1}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$$

project vectors on the two eigenvectors of σ_3. The projection operator $\mathbb{P}_{\sigma_3=1}$ can be written as

$$\mathbb{P}_{\sigma_3=1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \frac{\sigma_3 + \mathbf{I}}{2} \qquad (a^*, b^*)\frac{\sigma_3 + \mathbf{I}}{2}\begin{pmatrix} a \\ b \end{pmatrix} = a^* a$$

together with the probability as an expectation value $\langle\psi|\mathbb{P}|\psi\rangle$. The same is true for σ_1 and σ_2. Thus, the probabilities using (4.2) are

$$\left.\begin{matrix} P(A, \neg B) \\ P(B, \neg C) \end{matrix}\right\} = \frac{1}{\sqrt{2}}(|du\rangle - |ud\rangle)^* \frac{\sigma_3 + \mathbf{I}}{2} \frac{\frac{\tau_1 + \tau_3}{\sqrt{2}} + \mathbf{I}}{2} \frac{1}{\sqrt{2}}(|du\rangle - |ud\rangle)$$

$$= \frac{1}{\sqrt{2}}(|du\rangle - |ud\rangle)^* \left(\frac{-1}{2\sqrt{2}} + \frac{1}{2}\right)\frac{1}{\sqrt{2}}|ud\rangle = \frac{1}{\sqrt{2}}\left(\frac{-1}{2\sqrt{2}} + \frac{1}{2}\right)\frac{1}{\sqrt{2}} \qquad \approx 0.075$$

$$P(A, \neg C) = \frac{1}{\sqrt{2}}(|du\rangle - |ud\rangle)^* \frac{\sigma_3 + \mathbf{I}}{2} \frac{\tau_1 + \mathbf{I}}{2} \frac{1}{\sqrt{2}}(|du\rangle - |ud\rangle) = \frac{1}{\sqrt{2}}\frac{1}{2}\frac{1}{\sqrt{2}} \qquad = 0.25$$

for the three statements. The probabilities $P(A, \neg B)$ and $P(B, \neg C)$ are the same and have been calculated for $P(A, \neg B)$. Thus, $P(A, \neg B) + P(B, \neg C) \approx 0.15 < 0.25 = P(A, \neg C)$. This violates Bell's inequality, and therefore the physical configuration of two entangled spins forming a singlet state cannot be modeled as a classical system with classical logic under the assumption that signals cannot travel faster than the speed of light. The conclusion, in other word, is that quantum mechanics cannot be a statistical theory governed by classical logic.

To do this experiment practically, two electrons are put together such that they interact and get into the singlet state with lowest energy. Afterwards they are taken apart such that they are separated in space. One does this with many identically prepared pairs, and this set of pairs is divided into three groups. With the three groups, the properties A, B and C are measured, and the probabilities are calculated.

4.6 Entanglement and the No-Cloning Theorem

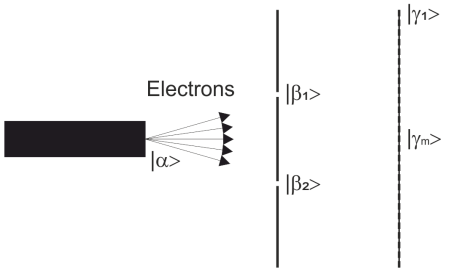
Entanglement is difficult to keep up because the two electrons can be disturbed by the environment by collisions with photons or particles of the atmosphere. Entanglement is, however, a very common phenomenon. A measurement of a property of an electron in an experiment, for example, is entanglement between apparatus and electron. If one of two entangled water molecules is put into a bathtub full of water, then the molecule outside is somehow entangled with the state of the bathtub.

A cloning apparatus looks at a state of a system and spits out two clones of the same system. It is the goal of this apparatus that one system goes in and two identical systems come out. For example, one $|u\rangle$ enters and $|u\rangle\,|u\rangle$ comes out. Similarly, one $|d\rangle$ enters and $|d\rangle\,|d\rangle$ comes out. If a spin $|r\rangle$ rotated by 90° is entered, $|r\rangle\,|r\rangle = 1/\sqrt{2}(|u\rangle + |d\rangle)1/\sqrt{2}(|u\rangle + |d\rangle)$ should come out, but this products gives also terms $|u\rangle\,|d\rangle$. Thus, cloning is inconsistent with the linearity principle of quantum mechanics. Multiplication and squaring things is not a linear operation, and the mechanism for combining two systems is the tensor product, but the evolution of state vectors is linear.

One can build machines that clone $|u\rangle$, but then it is not able to clone $|r\rangle$. If one knows the state, then one can build a machine that clones this state, but one cannot build a machine that can clone all states. This is the so-called no-cloning theorem.

4.7 The Experiment with Two Slits

In the two-slit experiment a ray of electrons is emitted towards a screen with two slits and observed on a second screen behind the first one. On the second screen, the arriving electrons build an interference pattern. If one measures through which slit the electrons went, the interference pattern disappears and becomes the bipolar distribution as one would expect classically for particles. The electrons not going through one of the two slits are just reflected and can be ignored. For some calculations only the position without the spin is used to determine the state of the electrons.



If $|\alpha\rangle$ evolves to $|\alpha'\rangle$ and $|\beta\rangle$ evolves to $|\beta'\rangle$, then $|\alpha\rangle + |\beta\rangle$ evolves to $|\alpha'\rangle + |\beta'\rangle$. In other words, time evolution is linear. The electrons get emitted in state $|\alpha\rangle$ and either go through the upper slit or through the lower slit. If the distance to the two slits is exactly the same, the electron has proceeded to state $|\beta_1\rangle + |\beta_2\rangle$ with normalization ignored. The second screen is assumed for simplicity to be a discrete set of points each corresponding to a state $|\gamma_n\rangle$. The electrons in state $|\beta_1\rangle$ evolve into a superposition $\sum a_n |\gamma_n\rangle$, and those in state $|\beta_2\rangle$ also evolve into a superposition $\sum b_n |\gamma_n\rangle$. The state $|\beta_1\rangle + |\beta_2\rangle$ therefore evolves into a state $\sum (a_n + b_n) |\gamma_n\rangle$. Calculation of the probability for the electron to arrive in state $|\gamma_n\rangle$ contains the terms $(a_n^* + b_n^*)(a_n + b_n)$ and some factors for normalization. This product is not just the sum $a_n^* a_n + b_n^* b_n$ as it would be in classical physics where the probability to go through the upper slit and the probability to go through the lower slit are added. The cross-terms $a_n^* b_n + b_n^* a_n$ in quantum mechanics is responsible for the interference.

If $|\gamma_m\rangle$ is the state of the point on the second screen in the middle such that the distance from the two slits is exactly the same, then $a_m = b_m$ and the probability is $4 \cdot |a_m|^2 = 4 \cdot |b_m|^2$. Thinking in terms of classical physics, one would expect it to be just twice as much as the probability to go through one of the two slits, but it is four times this probability. This is due to constructive interference.

If one adds something such as a little potential on the path through the lower slit such that the time evolution from $|\alpha\rangle$ becomes $|\beta_1\rangle - |\beta_2\rangle$, then also the time evolution from the lower slit to the second screen changes sign to $-\sum b_n |\gamma_n\rangle$. Now, the probabilities for the point in the middle of the second screen with the state $|\gamma_m\rangle$ becomes 0 because $b_m = -a_m$. This is an example of destructive interference.

If the upper slit is closed and electrons can only go through the lower slit, the minus sign due to the little potential has no influence because it is just a phase. Only when both slits are open, the influence from the potential has an effect, because the amplitudes add and not the probabilities.

No longer ignoring the spin, an apparatus to measure through which slit an electron goes can change its spin from down to up such that one can detect when it has gone through the upper slit. The state changes from $|\alpha d\rangle$ to $|\beta_1 u\rangle$, when it goes through the upper slit, and to $|\beta_2 d\rangle$, when it goes through the lower slit. When both slits are open, the state $|\beta_1 u\rangle + |\beta_2 d\rangle$ changes to $\sum (a_n |\gamma_n u\rangle + b_n |\gamma_n d\rangle)$. Because $|\gamma_n u\rangle$ and $|\gamma_n d\rangle$ are no longer the same state but are orthogonal, there is no longer interference, and the probability is proportional to $a_n^* a_n + b_n^* b_n$.

Making a measurement is establishing entanglement between a device and a system. The destruction of the interference is called the *collapse of the wave function*. It is especially the collapse of the interference part of the wave function corresponding to the cross-terms. If one ignores the spin, one talks about the collapse of the wave function. If one does not ignore the spin, one talks about entanglement.

This destruction of the interference is a matter of degree. If the apparatus only sometimes changes the spin, only part of the interference gets eliminated. If the state changes from $|\alpha d\rangle$ to the state $\sqrt{1-\varepsilon}\,|\beta_1 u\rangle + \sqrt{\varepsilon}\,|\beta_1 d\rangle$ and not to the state $|\beta_1 u\rangle$ when the electron goes through the upper slit, the spin is sometimes not changed. The two-slit experiment is actually easier to do with photons, because the electrons may interact with the material of the screen with the two slits.

In classical physics an experiment does not influence the system, but in quantum mechanics an entanglement is always established with a measurement, but it is not important that a person makes the measurement. The atmosphere, for example, continuously measures objects such as a tree. Therefore, entanglement is an important phenomenon for us to see the world as it is for us.

The Schrödinger cat is not in a superposition of dead and alive, it is in an entangled state with the gun, and the composite system is in a superposition of the state not fired and alive and the state fired and dead. When Schrödinger looks into the box with the cat he becomes entangled as well.

5 Entropy

5.1 Classical Entropy

Given a system with a finite number N of states, the entropy has to do with the system and the knowledge about it. The system is assumed to be known to be in one of n states. If $n = 1$ then there is no ignorance, and if $n = N$ then there is complete ignorance. Entropy is defined as $S = \log n$. Complete knowledge means $S = 0$ and complete ignorance means $S = \log N$. For a system of m spins, the possible number of states is 2^m, and $\log_2 2^m = m$ is the number of bits needed to encode all possibilities.

The above definition assumes that the probability for i among the n states is $P_i = \frac{1}{n}$ and zero otherwise, but the entropy is defined as $S = -\sum P_i \log(P_i)$ for any probability distribution. In the case of $P_i = \frac{1}{n}$, the entropy is simply $S = -n\frac{1}{n}\log\left(\frac{1}{n}\right) = \log(n)$ as used above.

The bigger the entropy, the less one knows about the system. Maximum entropy for a system with N states is $\log(N)$. The difference between the maximum entropy and the actual entropy is the amount of information available.

5.2 The Trace of a Linear Operator

The *trace* of a linear operator $\mathbf{M}$ is defined as $\mathrm{tr}(\mathbf{M}) = \sum_n \langle\psi_n|\mathbf{M}|\psi_n\rangle$ where $|\psi_n\rangle$ is a basis. Because the trace of an operator is independent of the basis, the basis can be selected such that $\mathbf{M}$ is diagonal which is always possible for a Hermitian operator and where the sum of the eigenvalues is the trace. For the trace of the product of two operators, one can easily prove $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$.

Classically, for a system with N possible states the expectation value or average of an observable $F(i)$ with probability P_i is $\overline{F} = \sum_i P_i F(i)$. In quantum mechanics, the trace is the analog where one sums over all possibilities.

5.3 Density Matrix

If a system has been prepared in a known state, the system is in a so-called *pure state*. However, if somebody has created the system in a state but the state is not known, this adds another degree of probability. In this case, one uses *density matrices* specifying a probability distribution for the different states instead of a single state. This is therefore a looser concept, and the unknown state in which the system has been prepared is called a *mixed state*.

A system in a mixed state is assumed to have been prepared in one of the states $|\psi_i\rangle$ with a probability ρ_i. One can build a diagonal matrix ρ with the elements ρ_i. All the ρ_i are real numbers between 0 and 1, and the sum over all ρ_i is 1. The last fact can be written as $\mathrm{tr}(\rho) = 1$. This Hermitian matrix is called the *density matrix*. Obviously, the values ρ_i are the eigenvalues of ρ.

If a system is in a pure state $|\alpha\rangle$, the density matrix ρ is the projection operator on that state and has one eigenvector with eigenvalue 1. This can be written as $\rho = |\alpha\rangle\langle\alpha|$. Assuming that $|\alpha\rangle$ is normalized, then $|\alpha\rangle\langle\alpha|\alpha\rangle = |\alpha\rangle$. Any state $|\beta\rangle$ orthogonal to $|\alpha\rangle$ gives $|\alpha\rangle\langle\alpha|\beta\rangle = 0$. Independent of whether the state is pure or mixed, $\overline{\mathbf{M}} = \mathrm{tr}(\mathbf{M}\rho) = \mathrm{tr}(\rho\mathbf{M})$ is the average of an observable $\mathbf{M}$. Because the trace is independent of the basis chosen, $\overline{\mathbf{M}} = \sum_i \langle\psi_i|\rho\mathbf{M}|\psi_i\rangle$ where $|\psi_i\rangle$ is just a complete basis and needs not to be the one with ρ diagonal. If $|\alpha\rangle$ is a pure state, $\overline{\mathbf{M}} = \sum_i \langle\psi_i|\alpha\rangle\langle\alpha|\mathbf{M}|\psi_i\rangle = \sum_i \langle\alpha|\mathbf{M}|\psi_i\rangle\langle\psi_i|\alpha\rangle$, and

using the resolution of the identity the result is not very surprisingly the expectation value $\overline{\mathbf{M}} = \langle\alpha|\mathbf{M}|\alpha\rangle$ in the state $|\alpha\rangle$.

If the system, on the other hand, is in a mixed state with the basis $|\psi_i\rangle$ chosen such that the density matrix ρ is diagonal, $\overline{\mathbf{M}} = \sum_i \langle\psi_i|\rho\mathbf{M}|\psi_i\rangle = \sum_{i,j} \langle\psi_i|\rho|\psi_j\rangle \langle\psi_j|\mathbf{M}|\psi_i\rangle$ using again the resolution of the identity. Because the basis has been selected such that ρ is diagonal, $\langle\psi_i|\rho|\psi_j\rangle$ is the eigenvalue ρ_i of ρ if $i = j$ and 0 otherwise. Thus, $\overline{\mathbf{M}} = \sum_i \rho_i \langle\psi_i|\mathbf{M}|\psi_i\rangle$ where ρ_i is the probability to be in state $|\psi_i\rangle$ in the basis where ρ is diagonal, but $\rho_i \langle\psi_i|\mathbf{M}|\psi_i\rangle$ is just the expectation value of $\mathbf{M}$ in state $|\psi_i\rangle$ weighted with the probability ρ_i of the state $|\psi_i\rangle$. The values ρ_i are the probabilities for the system to be in state $|\psi_i\rangle$ only in the basis where ρ is diagonal.

5.4 Entanglement Entropy

The entropy in quantum mechanics is also called *entanglement entropy* and is defined as

$$S = -\operatorname{tr}(\rho)\log\rho \tag{5.1}$$

where $\log\rho$ is the matrix with the values $\log\rho_{ij}$. If one knows the state in which the system has been created, the entropy in quantum mechanics is zero. Thus, if only one of the ρ_i in the density matrix is not zero, then there is zero entropy. If one has no knowledge in which of n possible states the system is, the density matrix is proportional to the unit matrix and each eigenvalue is $\frac{1}{n}$.

Two discrete subsystems A and B with the basis $|\alpha_a\rangle$ with $a \in \{1, ..., N_A\}$ and $|\beta_b\rangle$ with $b \in \{1, ..., N_B\}$ can have the $N_A \cdot N_B$ combined states $|\alpha_a\beta_b\rangle$ as an entangled system. If one puts, for example, the spin of an electron with two states together with an atom with three states, the combined system has six states. The most general pure state of the combined system is a superposition $\sum_{ab} f(a,b)\,|\alpha_a\beta_b\rangle$ where $f(a,b)$ is a function which takes two integers as input and returns a complex number as output. This state vector is assumed to be normalized such that $\sum_{ab} f(a,b)^* f(a,b) = 1$.

An experiment measuring observable $\mathbf{M}$ for subsystem A without involving subsystem B gives

$$\sum_{a_2 b_2 a_1 b_1} f(a_2,b_2)^* \langle\alpha_{a_2}\beta_{b_2}|\mathbf{M}|\alpha_{a_1}\beta_{b_1}\rangle\, f(a_1,b_1)$$

as the expectation value, but because $\mathbf{M}$ does not act on the subsystem B, b_1 and b_2 must be the same and can be replaced by b. Because $\sum_b f(a_2,b)^* f(a_1,b) = \rho_{a_1 a_2}$, and $\sum_{a_1 a_2} M_{a_2 a_1}\rho_{a_1 a_2}$ with $\mathbf{M}$ in components is the trace $\operatorname{tr}(\mathbf{M}\rho)$, the subsystem B appears as a density matrix in the expectation value.

If one has a composite system and looks only at one subsystem, this subsystem is described by a density matrix, and it is usually the density matrix of a mixed state. A pure state of the combined system does not mean that the subsystems are in a pure state. For the subsystems to be in a pure state, the function $f(a,b)$ must be separable such that $f(a,b) = g(a)h(b)$. Thus, $\rho_{a_1 a_2} = \sum_b g(a_1)h(b)g^*(a_2)h^*(b) = g(a_1)g^*(a_2)$, because $\sum_b h(b)h^*(b) = 1$. The expectation value for $\mathbf{M}$ is $g(a_2)M_{a_2 a_2}g(a_1)$ in components, and this is $\overline{\mathbf{M}} = \langle\alpha|\mathbf{M}|\alpha\rangle$.

5.5 Examples of Systems with Two Spins

The singlet state $1/\sqrt{2}(|ud\rangle - |du\rangle)$ has $f(uu) = f(dd) = 0$ and $f(ud) = -f(du) = 1/\sqrt{2}$. The density matrix for the subsystem of the first spin has the four elements $\rho_{uu} = f(uu)^* f(uu) + f(ud)^* f(ud) = \frac{1}{2}$, $\rho_{dd} = f(du)^* f(du) + f(dd)^* f(dd) = \frac{1}{2}$, and $\rho_{ud} = f(du)^* f(uu) + f(dd)^* f(ud) = 0 = \rho_{du}$. The density matrix is therefore

$$\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\mathbf{I}$$

and the entropy is $S = -(\frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\frac{1}{2}) = \log 2$. The density matrix for the subsystem of the second spin is the same, and the entropy is for both subsystems maximal.

The expectation value for the first spin can be calculated using the density matrix. The expectation value of the first spin along any axis $\vec{n}$ is $\overline{\sigma\cdot\vec{n}} = \operatorname{tr}(\rho\,\sigma\cdot\vec{n}) = \frac{1}{2}\,\sigma\cdot\vec{n} = 0$ because the trace of all the Pauli

matrices is zero. Thus, in the singlet state the trace of any component in any direction is zero. This means that the spin is equally likely to point in one direction or in the other direction. It is a maximally mixed state.

Another combined system with two spins is one with $f(uu) = f(dd) = f(ud) = f(du) = \frac{1}{2}$. The density matrix is therefore

$$\rho = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

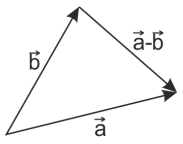
with $\mathrm{tr}(\rho) = 1$. Because the product of the eigenvalues is the determinant, and the sum of the eigenvalues is the trace, the eigenvalues ρ_1 and ρ_2 satisfy the two equations $\rho_1 \cdot \rho_2 = 0$ and $\rho_1 + \rho_2 = 1$ with the solution $\rho_1 = 1$ and $\rho_2 = 0$. This is therefore a product state. The expectation values of the Pauli matrices are $\mathrm{tr}(\sigma_2\rho) = \mathrm{tr}(\sigma_2\rho) = 0$, but $\mathrm{tr}(\sigma_1\rho) = 1$, and the entropy is zero. Changing the probabilities from $\frac{1}{2}$ a little bit would result in a small entropy.

6 The Evolution of a System with Time

6.1 Unitary Operators

Although there are only two observable states up and down for the spin of an electron and the vector space of states is discrete, the state can change continuously in time. However, two orthogonal states should stay orthogonal because orthogonal states represent those states of a system which can be distinguished by a measurement. Physics is deterministic in the sense that two distinguishable states stay two distinguishable states when evolving in time.

The first principle for evolution in time is that the operator performing the change is linear. The second principle is that the inner product of two vectors is independent of the evolution in time. The operators satisfying both principles are called *unitary*. Because the inner product is not changed by unitary operators, the triangle spanned by the three vectors $\vec{a}$, $\vec{b}$ and $\vec{a} - \vec{b}$ does not change, and the lengths as well as the angles are therefore preserved.



Evolution in time can be represented by $|\alpha(t)\rangle = \mathbf{U}(t)\,|\alpha(0)\rangle$. Thus, $\mathbf{U}(0) = \mathbf{I}$. For any two vectors, the inner product $\langle\alpha(t)|\beta(t)\rangle = \langle\alpha(0)|\mathbf{U}^\dagger(t)\mathbf{U}(t)|\beta(0)\rangle = \langle\alpha(0)|\beta(0)\rangle$. Thus, $\mathbf{U}^\dagger(t)\mathbf{U}(t) = \mathbf{I}$, and operators with this property are the unitary operators. This can be written in the form $\mathbf{U}^\dagger(t) = \mathbf{U}^{-1}(t)$.

6.2 Generalized Schrödinger Equation

For a very small time change ε, time evolution can be approximated in the form $\mathbf{U}(\varepsilon) = \mathbf{I} - \frac{i}{\hbar}\varepsilon\mathbf{H}$ for some operator $\mathbf{H}$ with the Hermitian conjugate $\mathbf{U}^\dagger(\varepsilon) = \mathbf{I} + \frac{i}{\hbar}\varepsilon\mathbf{H}^\dagger$. (The factor $\frac{i}{\hbar}$ will later turn out to be convenient.) As a unitary operator,

$$\mathbf{U}^\dagger(\varepsilon)\mathbf{U}(\varepsilon) = \mathbf{U}^{-1}(\varepsilon)\mathbf{U}(\varepsilon) = (\mathbf{I} + \frac{i}{\hbar}\varepsilon\mathbf{H}^\dagger)(\mathbf{I} - \frac{i}{\hbar}\varepsilon\mathbf{H}) = \mathbf{I}$$

or $\mathbf{I} + \frac{i}{\hbar}\varepsilon\mathbf{H}^\dagger - \frac{i}{\hbar}\varepsilon\mathbf{H} + O(\varepsilon^2) = \mathbf{I}$. Ignoring terms ε^2 and higher, it follows $\mathbf{H}^\dagger = \mathbf{H}$, and the operator $\mathbf{H}$ is therefore Hermitian and corresponds to an observable.

The operator $\mathbf{H}$ may be different for every system, but it is something general that every system has. It is called the *Hamiltonian* and corresponds to the measurable energy with the energy levels as its eigenvalues. Its most important property is that it does not change in time. Because $t = 0$ is an arbitrary point in time, the derivative

$$|\alpha(\varepsilon)\rangle - |\alpha(0)\rangle = -\frac{i}{\hbar}\varepsilon\mathbf{H}\,|\alpha(0)\rangle \qquad \frac{|\alpha(t+\varepsilon)\rangle - |\alpha(t)\rangle}{\varepsilon} = -\frac{i}{\hbar}\mathbf{H}\,|\alpha(t)\rangle \qquad \frac{\partial}{\partial t}\,|\alpha(t)\rangle = -\frac{i}{\hbar}\mathbf{H}\,|\alpha(t)\rangle \qquad (6.1)$$

can be built, and the last of these equations is the Schrödinger equation generalized for any system.

As every Hermitian operator, $\mathbf{H}$ has a complete set of eigenvectors $|\eta_n\rangle$. If $|\eta_n\rangle$ is an eigenvector, it stays an eigenvector because of $\mathbf{H}\,|\eta_n\rangle = E\,|\eta_n\rangle$. With $|\eta_n(t)\rangle = f(t)\,|\eta_n(0)\rangle$ and $\frac{\partial f(t)}{\partial t} = -i\frac{E}{\hbar}f(t)$, the

function $f(t)$ is an exponential function $f(t) = \exp\left(-i\frac{E}{\hbar}t\right)f(0)$. The evolution of an eigenvector in time is therefore the multiplication with a time-dependent phase with angular frequency ω such that $E = \hbar\omega$. Any state $|\alpha\rangle$ can be expanded with eigenvectors $|\eta_n\rangle$ of $\mathbf{H}$ as

$$|\alpha(0)\rangle = \sum_n a_n |\eta_n\rangle \qquad\qquad |\alpha(t)\rangle = \sum_n a_n e^{-i\frac{E_n}{\hbar}t} |\eta_n\rangle$$

where E_n is the eigenvalue of eigenvector $|\eta_n\rangle$. Each eigenvector of $\mathbf{H}$ evolves with it own frequency. High energy vectors evolve very fast.

6.3 Time Evolution of Expectation Values

The expectation value of an operator $\mathbf{A}$ is $\overline{\mathbf{A}} = \langle\alpha|\mathbf{A}|\alpha\rangle$. Because states change and not operators, the expectation value of the operator $\mathbf{A}$ is

$$\frac{d\overline{\mathbf{A}}}{dt} = \dot{\overline{\mathbf{A}}} = \langle\dot{\alpha}|\mathbf{A}|\alpha\rangle + \langle\alpha|\mathbf{A}|\dot{\alpha}\rangle = \langle\alpha| + \frac{i}{\hbar}\mathbf{H}\mathbf{A}|\alpha\rangle + \langle\alpha|\mathbf{A} - \frac{i}{\hbar}\mathbf{H}|\alpha\rangle = \frac{i}{\hbar}\langle\alpha|\mathbf{H}\mathbf{A} - \mathbf{A}\mathbf{H}|\alpha\rangle$$

using the Schrödinger equation (6.1). The change in time of the average of $\mathbf{A}$

$$\dot{\overline{\mathbf{A}}} = \frac{i}{\hbar}\overline{[\mathbf{H}, \mathbf{A}]} \tag{6.2}$$

is the average of the commutator of $\mathbf{A}$ with $\mathbf{H}$. The conservation law for energy follows from $\dot{\overline{\mathbf{H}}} = 0$.

The operator $\mathbf{H}$ is diagonal in the basis of its eigenvectors, and the diagonal elements are the eigenvalues E_n. Therefore the operators $\mathbf{H}$ and $\mathbf{U}(t)$ are

$$\mathbf{H} = \begin{pmatrix} E_1 & 0 & 0 & ... \\ 0 & E_2 & 0 & ... \\ 0 & 0 & ... & ... \\ ... & ... & ... & ... \end{pmatrix} \qquad \mathbf{U}(t) = \begin{pmatrix} e^{\frac{-iE_1}{\hbar}} & 0 & 0 & ... \\ 0 & e^{\frac{-iE_2}{\hbar}} & 0 & ... \\ 0 & 0 & e^{\frac{-iE_3}{\hbar}} & ... \\ ... & ... & ... & ... \end{pmatrix} = e^{\frac{-i\mathbf{H}}{\hbar}}$$

and the operator $\mathbf{U}(t)$ for the time evolution is also diagonal.

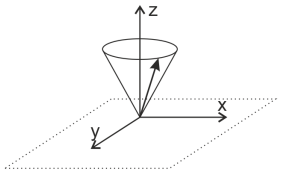
6.4 Single Spin in a Magnetic Field

A single spin with the components σ_1, σ_2, σ_3 is prepared in a magnetic field $\vec{B}$. The magnetic moment of the particle is $\vec{\mu} = \mu\vec{\sigma}/2$. The energy and therefore the Hamiltonian is

$$\mathbf{H} = \frac{\mu\vec{\sigma}}{2} \cdot \vec{B} \qquad\qquad \vec{B} = \begin{pmatrix} 0 \\ 0 \\ B \end{pmatrix} \qquad\qquad \mathbf{H} = \frac{\mu}{2}B\sigma_3$$

with $\vec{B}$ pointing in z-direction. One can now calculate the time evolution of the expectation values of σ_1, σ_2 and σ_3.

The time evolution of the expectation value of σ_3 is $\dot{\overline{\sigma_3}} = i[\mathbf{H}, \sigma_3]/\hbar = 0$ because $[\sigma_3, \sigma_3] = 0$. The component of the spin along the z-axis does not change with time, and therefore the projection of the spin to this axis and the angle between this axis and the spin does not change. Whatever the spin does, the expectation value of the spin must move on a cone of fixed angle. This is a consequence of conservation of energy.



For the σ_1 and the σ_2 components, the product with $\mathbf{H}$ respectively σ_3 is needed. The result for σ_1 is

$$\sigma_3\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i\,\sigma_2 \qquad \sigma_1\sigma_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -i\,\sigma_2$$

and therefore $[\sigma_3, \sigma_1] = 2i\sigma_2$. Thus,

$$\dot{\overline{\sigma_1}} = \frac{i\mu}{2\hbar} B[\sigma_3, \sigma_1] = -\frac{\mu}{\hbar} B\sigma_2 \qquad\qquad \dot{\overline{\sigma_2}} = +\frac{\mu}{\hbar} B\sigma_1$$

after a similar calculation for σ_2. This is a rotation in a plane and has the simple solution

$$\sigma_1 = \cos\left(\frac{\mu B}{\hbar} t\right) \qquad\qquad \sigma_2 = \sin\left(\frac{\mu B}{\hbar} t\right)$$

corresponding to the motion of the spin on the cone like a gyroscope.

6.5 Time Evolution of Two Entangled Spins

Two spins are assumed to be not in a magnetic field, but providing a magnetic field for each other. The Hamiltonian $\mathbf{H}$ is proportional to $\vec{\sigma}\vec{\tau}$ as the 4×4 matrix

$$\sigma_1\tau_1 + \sigma_2\tau_2 + \sigma_3\tau_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

in the basis $\{|uu\rangle, |ud\rangle, |du\rangle, |dd\rangle\}$ and has the eigenvectors $|uu\rangle$, $|dd\rangle$, and $|ud\rangle \pm |du\rangle$. The two states $|uu\rangle$ and $|dd\rangle$ have the same energy E_T and are therefore degenerate. The state $|ud\rangle + |du\rangle$ has also the same energy E_T. This is the reason why the triplet of states $|uu\rangle$, $|dd\rangle$ and $|ud\rangle + |du\rangle$ is called triplet. The state $|ud\rangle - |du\rangle$ has a different energy level E_S, and this is the reason why it is called singlet. The energy level of $|uu\rangle$ and $|dd\rangle$ is bigger than that of $|ud\rangle$ and $|du\rangle$ because magnets close together prefer to point in different directions. Thus, one can assume $E_T > E_S$.

The four vectors $|uu\rangle$, $|dd\rangle$, and $1/\sqrt{2}(|ud\rangle \pm |du\rangle)$ build a orthonormal basis, and every state of the two spins can be represented with them. The system can start, for example, in state $|ud\rangle$ which is

$$|ud\rangle = \frac{|ud\rangle + |du\rangle}{2} + \frac{|ud\rangle - |du\rangle}{2} = \frac{|T\rangle}{\sqrt{2}} + \frac{|S\rangle}{\sqrt{2}}$$

with $|T\rangle$ for the triplet state $1/\sqrt{2}(|ud\rangle + |du\rangle)$ and $|S\rangle$ for the singlet state $1/\sqrt{2}(|ud\rangle - |du\rangle)$. The state $|ud\rangle$ evolves with time to

$$e^{-i\frac{E_T}{\hbar}t}\frac{|T\rangle}{\sqrt{2}} + e^{-i\frac{E_S}{\hbar}t}\frac{|S\rangle}{\sqrt{2}} \qquad\qquad \frac{1}{2}\left(e^{-i\frac{E_T}{\hbar}t} + e^{-i\frac{E_S}{\hbar}t}\right)|ud\rangle + \frac{1}{2}\left(e^{-i\frac{E_T}{\hbar}t} - e^{-i\frac{E_S}{\hbar}t}\right)|du\rangle$$

specified in terms of $|T\rangle$ and $|S\rangle$ respectively converted back to the states $|ud\rangle$ and $|du\rangle$.

At time $t = 0$, the system starts in state $|ud\rangle$ and the coefficient for $|du\rangle$ is 0. After a little bit of time, the coefficient for $|ud\rangle$ get a bit smaller than 1 and consequently the coefficient for $|du\rangle$ gets a bit bigger than 0. Thus, first only a bit of $|du\rangle$ gets mixed in, but after a while, the system is completely in the state $|du\rangle$ with no $|ud\rangle$ added. Obviously, the system goes back and forth between $|ud\rangle$ and $|du\rangle$. The evolution in time can be written as

$$\frac{e^{i\frac{E_T-E_S}{2\hbar}t} + e^{-i\frac{E_T-E_S}{2\hbar}t}}{2}|ud\rangle + i\frac{e^{i\frac{E_T-E_S}{2\hbar}t} - e^{-i\frac{E_T-E_S}{2\hbar}t}}{2i}|du\rangle = \cos\left(\tfrac{E_T-E_S}{2\hbar}t\right)|ud\rangle + i\sin\left(\tfrac{E_T-E_S}{2\hbar}t\right)|du\rangle$$

and shows converted to sine and cosine that it is an oscillation between $|ud\rangle$ and $|du\rangle$. The shift of the energy to $E_T - E_S$ corresponds to a multiplication with a time-dependent phase $\exp\left(i\frac{E_T+E_S}{2\hbar}t\right)$ and does not change the probabilities. This shows that the energy levels per se are not relevant, and only differences of energies are important.