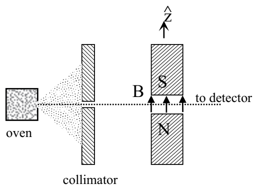
Introduction

State vectors

Stern Gerlach experiment



In the Stern Gerlach experiment

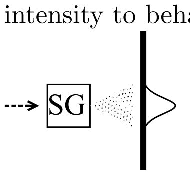
- silver atoms are heated in an oven, from which they escape through a narrow slit,
- the atoms pass through a collimator and enter an inhomogenous magnetic field, we assume the field to be uniform in the xy-plane and to vary in the z-direction,
- a detector measures the intensity of the electrons emerging from the magnetic field as a function of z.

We know that

- 46 of the 47 electrons of a silver atom form a spherically symmetric shell and the angular momentum of the electron outside the shell is zero, so the magnetic moment due to the orbital motion of the electrons is zero,
- the magnetic moment of an electron is $c\boldsymbol{S}$, where $\boldsymbol{S}$ is the spin of an electron,
- the spins of electrons cancel pairwise,
- thus the magnetic moment $\boldsymbol{\mu}$ of an silver atom is almost solely due to the spin of a single electron, i.e. $\boldsymbol{\mu} = c\boldsymbol{S}$,
- the potential energy of a magnetic moment in the magnetic field $\boldsymbol{B}$ is $-\boldsymbol{\mu}\cdot\boldsymbol{B}$, so the force acting in the z-direction on the silver atoms is

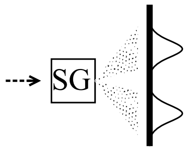
$$F_z = \mu_z \frac{\partial B_z}{\partial z}.$$

So the measurement of the intensity tells how the z-component the angular momentum of the silver atoms passing through the magnetic field is distributed. Because the atoms emerging from the oven are randomly oriented we would expect the intensity to behave as shown below.



classically

In reality the beam is observed to split into two components.



in reality

Based on the measurements one can evaluate the z-components S_z of the angular momentum of the atoms and find out that

- for the upper distribution $S_z = \hbar/2$.
- for the lower distribution $S_z = -\hbar/2$.

In quantum mechanics we say that the atoms are in the angular momentum *states* $\hbar/2$ and $-\hbar/2$.
The *state vector* is a mathematical tool used to represent the states. Atoms reaching the detector can be described, for example, by the *ket*-vectors $|S_z;\uparrow\rangle$ and $|S_z;\downarrow\rangle$. Associated with the ket-vectors there are dual *bra*-vectors $\langle S_z;\uparrow|$ and $\langle S_z;\downarrow|$. State vectors are assumed

- to be a complete description of the described system,
- to form a linear (Hilbert) space, so the associated mathematics is the theory of (infinite dimensional) linear spaces.

When the atoms leave the oven there is no reason to expect the angular momentum of each atom to be oriented along the z-axis. Since the state vectors form a linear space also the superposition

$$c_\uparrow|S_z;\uparrow\rangle + c_\downarrow|S_z;\downarrow\rangle$$

is a state vector which obviously describes an atom with angular momentum along both positive and negative z-axis.
The magnet in the Stern Gerlach experiment can be thought as an apparatus measuring the z-component of the angular momentum. We saw that after the measurement the atoms are in a definite angular momentum state, i.e. in the measurement the state

$$c_\uparrow|S_z;\uparrow\rangle + c_\downarrow|S_z;\downarrow\rangle$$

collapses either to the state $|S_z;\uparrow\rangle$ or to the state $|S_z;\downarrow\rangle$.
A generalization leads us to the measuring postulates of quantum mechanics:
Postulate 1 *Every measurable quantity is associated with a Hermitean operator whose eigenvectors form a complete basis (of a Hilbert space),*
and
Postulate 2 *In a measurement the system makes a transition to an eigenstate of the corresponding operator and the result is the eigenvalue associated with that eigenvector.*
If $\mathcal{A}$ is a measurable quantity and A the corresponding Hermitean operator then an arbitrary state $|\alpha\rangle$ can be described as the superposition

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle,$$

where the vectors $|a'\rangle$ satisfy

$$A|a'\rangle = a'|a'\rangle.$$

The measuring event $\mathcal{A}$ can be depicted symbolically as

$$|\alpha\rangle \xrightarrow{\mathcal{A}} |a'\rangle.$$

In the Stern Gerlach experiment the measurable quantity is the z-component of the spin. We denote the measuring event by $\boxed{\mathrm{SG}\hat{\boldsymbol{z}}}$ and the corresponding Hermitean operator by S_z. We get

$$\begin{array}{rcl} S_z|S_z;\uparrow\rangle & = & \frac{\hbar}{2}|S_z;\uparrow\rangle \\ S_z|S_z;\downarrow\rangle & = & -\frac{\hbar}{2}|S_z;\downarrow\rangle \\ |S_z;\alpha\rangle & = & c_\uparrow|S_z;\uparrow\rangle + c_\downarrow|S_z;\downarrow\rangle \\ |S_z;\alpha\rangle & \xrightarrow{\boxed{\mathrm{SG}\hat{\boldsymbol{z}}}} & |S_z;\uparrow\rangle \quad \text{or} \\ |S_z;\alpha\rangle & \xrightarrow{\boxed{\mathrm{SG}\hat{\boldsymbol{z}}}} & |S_z;\downarrow\rangle. \end{array}$$

Because the vectors $|a'\rangle$ in the relation

$$A|a'\rangle = a'|a'\rangle$$

are eigenvectors of an Hermitean operator they are orthognal with each other. We also suppose that they are normalized, i.e.

$$\langle a'|a''\rangle = \delta_{a'a''}.$$

Due to the completeness of the vector set they satisfy

$$\sum_{a'} |a'\rangle\langle a'| = 1,$$

where 1 stands for the identity operator. This property is called the *closure*. Using the orthonormality the coefficients in the superposition

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle$$

can be written as the scalar product

$$c_{a'} = \langle a'|\alpha\rangle.$$

An arbitrary linear operator B can in turn be written with the help of a complete basis $\{|a'\rangle\}$ as

$$B = \sum_{a',a''} |a'\rangle\langle a'|B|a''\rangle\langle a''|.$$

Abstract operators can be *represented* as matrices:

$$B \mapsto \begin{array}{c} \\ \langle a_1| \\ \langle a_2| \\ \langle a_3| \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccc} \quad|a_1\rangle\quad & \quad|a_2\rangle\quad & \quad|a_3\rangle\quad & \cdots \end{array} \\ \left(\begin{array}{cccc} \langle a_1|B|a_1\rangle & \langle a_1|B|a_2\rangle & \langle a_1|B|a_3\rangle & \cdots \\ \langle a_2|B|a_1\rangle & \langle a_2|B|a_2\rangle & \langle a_2|B|a_3\rangle & \cdots \\ \langle a_3|B|a_1\rangle & \langle a_3|B|a_2\rangle & \langle a_3|B|a_3\rangle & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array}\right) \end{array}.$$

Note The matrix representation is *not* unique, but depends on the basis. In the case of our example we get the 2×2-matrix representation

$$S_z \mapsto \frac{\hbar}{2}\left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right),$$

when we use the set $\{|S_z;\uparrow\rangle, |S_z;\downarrow\rangle\}$ as the basis. The base vectors map then to the unit vectors

$$\begin{array}{rcl} |S_z;\uparrow\rangle & \mapsto & \left(\begin{array}{c} 1 \\ 0 \end{array}\right) \\ |S_z;\downarrow\rangle & \mapsto & \left(\begin{array}{c} 0 \\ 1 \end{array}\right) \end{array}$$

of the two dimensional Euclidean space.
Although the matrix representations are not unique they are related in a rather simple way. Namely, we know that
Theorem 1 *If both of the basis $\{|a'\rangle\}$ and $\{|b'\rangle\}$ are orthonormalized and complete then there exists a unitary operator U so that*

$$|b_1\rangle = U|a_1\rangle, |b_2\rangle = U|a_2\rangle, |b_3\rangle = U|a_3\rangle, \ldots$$

If now X is the representation of an operaor A in the basis $\{|a'\rangle\}$ the representation X' in the basis $\{|b'\rangle\}$ is obtained by the similarity transformation

$$X' = T^\dagger X T,$$

where T is the representation of the base transformation operator U in the basis $\{|a'\rangle\}$. Due to the unitarity of the operator U the matrix T is a unitary matrix.
Since

- an abstract state vector, excluding an arbitrary phase factor, uniquely describes the physical system,
- the states can be written as superpositions of different base sets, and so the abstract operators can take different matrix forms,

the physics must be contained in the invariant propertices of these matrices. We know that
Theorem 2 *If T is a unitary matrix, then the matrices X and $T^\dagger XT$ have the same trace and the same eigenvalues.*
The same theorem is valid also for operators when the trace is defined as

$$\mathrm{tr}A = \sum_{a'} \langle a'|A|a'\rangle.$$

Since

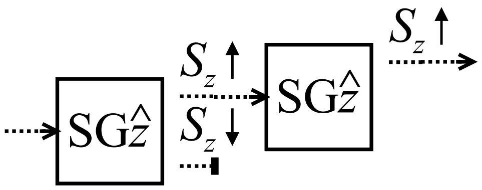
- quite obviously operators and matrices representing them have the same trace and the same eigenvalues,
- due to the postulates 1 and 2 corresponding to a measurable quantity there exists an Hermitean operator and the measuring results are eigenvalues of this operator,

the results of measurements are independent on the particular representation and, in addition, every measuring event corresponding to an operator reachable by a similarity transformation, gives the same results. Which one of the possible eigenvalues will be the result of a measurement is clarified by

Postulate 3 *If A is the Hermitean operator corresponding to the measurement $\mathcal{A}$, $\{|a'\rangle\}$ the eigenvectors of A associated with the eigenvalues $\{a'\}$, then the probability for the result a' is $|c_{a'}|^2$ when the system to be measured is in the state*

$$|\alpha\rangle = \sum_{a'} c_{a'}|a'\rangle.$$

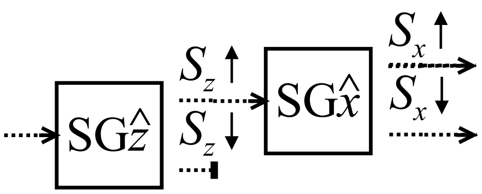
Only if the system already before the measurement is in a definite eigenstate the result can be predicted exactly. For example, in the Stern Gerlach experiment $\boxed{\mathrm{SG}\hat{\boldsymbol{z}}}$ we can block the emerging lower beam so that the spins of the remaining atoms are oriented along the positive z-axis. We say that the system is *prepared* to the state $|S_z;\uparrow\rangle$.



If we now let the polarized beam to pass through a new $\boxed{\mathrm{SG}\hat{\boldsymbol{z}}}$ experiment we see that the beam from the latter experiment does not split any more. According to the postulate this result can be predicted exactly.
We see that

- the postulate can also be interpreted so that the quantities $|c_{a'}|^2$ tell the probability for the system being in the state $|a'\rangle$,
- the physical meaning of the matrix element $\langle\alpha|A|\alpha\rangle$ is then the expectation value (average) of the measurement and
- the normalization condition $\langle\alpha|\alpha\rangle = 1$ says that the system is in one of the states $|a'\rangle$.

Instead of measuring the spin z-component of the atoms with spin polarized along the z-axis we let this polarized beam go through the $\boxed{\mathrm{SG}\hat{\boldsymbol{x}}}$ experiment. The result is exactly like in a single $\boxed{\mathrm{SG}\hat{\boldsymbol{z}}}$ experiment: the beam is again splitted into two components of equal intensity, this time, however, in the x-direction.



So, we have performed the experiment

$$|S_z;\uparrow\rangle \quad \xrightarrow{\boxed{\mathrm{SG}\hat{\boldsymbol{x}}}} \quad |S_x;\uparrow\rangle \quad \text{or}$$

$$|S_z;\uparrow\rangle \quad \xrightarrow{\boxed{\mathrm{SG}\hat{\boldsymbol{x}}}} \quad |S_x;\downarrow\rangle.$$

Again the analysis of the experiment gives $S_x = \hbar/2$ and $S_x = -\hbar/2$ as the x-components of the angular momenta. We can thus deduce that the state $|S_z;\uparrow\rangle$ is, in fact, the superposition

$$|S_z;\uparrow\rangle = c_{\uparrow\uparrow}|S_x;\uparrow\rangle + c_{\uparrow\downarrow}|S_x;\downarrow\rangle.$$

For the other component we have correspondingly

$$|S_z;\downarrow\rangle = c_{\downarrow\uparrow}|S_x;\uparrow\rangle + c_{\downarrow\downarrow}|S_x;\downarrow\rangle.$$

When the intensities are equal the coeffiecients satisfy

$$\begin{aligned}
|c_{\uparrow\uparrow}| &= |c_{\uparrow\downarrow}| &= \frac{1}{\sqrt{2}} \\
|c_{\downarrow\uparrow}| &= |c_{\downarrow\downarrow}| &= \frac{1}{\sqrt{2}}
\end{aligned}$$

according to the postulate 3. Excluding a phase factor, our postulates determine the transformation coefficients. When we also take into account the orthogonality of the state vectors $|S_z;\uparrow\rangle$ and $|S_z;\downarrow\rangle$ we can write

$$\begin{aligned}
|S_z;\uparrow\rangle &= \frac{1}{\sqrt{2}}|S_x;\uparrow\rangle + \frac{1}{\sqrt{2}}|S_x;\downarrow\rangle \\
|S_z;\downarrow\rangle &= e^{i\delta_1}\left(\frac{1}{\sqrt{2}}|S_x;\uparrow\rangle - \frac{1}{\sqrt{2}}|S_x;\downarrow\rangle\right).
\end{aligned}$$

There is nothing special in the direction $\hat{\boldsymbol{x}}$, nor for that matter, in any other direction. We could equally well let the beam pass through a $\boxed{\mathrm{SG}\hat{\boldsymbol{y}}}$ experiment, from which we could deduce the relations

$$\begin{aligned}
|S_z;\uparrow\rangle &= \frac{1}{\sqrt{2}}|S_y;\uparrow\rangle + \frac{1}{\sqrt{2}}|S_y;\downarrow\rangle \\
|S_z;\downarrow\rangle &= e^{i\delta_2}\left(\frac{1}{\sqrt{2}}|S_y;\uparrow\rangle - \frac{1}{\sqrt{2}}|S_y;\downarrow\rangle\right),
\end{aligned}$$

or we could first do the $\boxed{\mathrm{SG}\hat{\boldsymbol{x}}}$ experiment and then the $\boxed{\mathrm{SG}\hat{\boldsymbol{y}}}$ experiment which would give us

$$\begin{aligned}
|S_x;\uparrow\rangle &= \frac{e^{i\delta_3}}{\sqrt{2}}|S_y;\uparrow\rangle + \frac{e^{i\delta_4}}{\sqrt{2}}|S_y;\downarrow\rangle \\
|S_x;\downarrow\rangle &= \frac{e^{i\delta_3}}{\sqrt{2}}|S_y;\uparrow\rangle - \frac{e^{i\delta_4}}{\sqrt{2}}|S_y;\downarrow\rangle.
\end{aligned}$$

In other words

$$\begin{aligned}
|\langle S_y;\uparrow|S_x;\uparrow\rangle| &= |\langle S_y;\downarrow|S_x;\uparrow\rangle| &= \frac{1}{\sqrt{2}} \\
|\langle S_y;\uparrow|S_x;\downarrow\rangle| &= |\langle S_y;\downarrow|S_x;\downarrow\rangle| &= \frac{1}{\sqrt{2}}.
\end{aligned}$$

We can now deduce that the unknown phase factors must satisfy

$$\delta_2 - \delta_1 = \pi/2 \text{ or } -\pi/2.$$

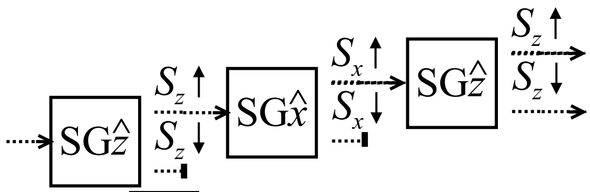
A common choice is $\delta_1 = 0$, so we get, for example,

$$\begin{aligned}
|S_z;\uparrow\rangle &= \frac{1}{\sqrt{2}}|S_x;\uparrow\rangle + \frac{1}{\sqrt{2}}|S_x;\downarrow\rangle \\
|S_z;\downarrow\rangle &= \frac{1}{\sqrt{2}}|S_x;\uparrow\rangle - \frac{1}{\sqrt{2}}|S_x;\downarrow\rangle.
\end{aligned}$$

Thinking like in classical mechanics, we would expect both the z- and x-components of the spin of the atoms in the upper beam passed through the $\boxed{\mathrm{SG}\hat{\boldsymbol{z}}}$ and $\boxed{\mathrm{SG}\hat{\boldsymbol{x}}}$ experiments to be $S_{x,z} = \hbar/2$. On the other hand, we can reverse the relations above and get

$$|S_x;\uparrow\rangle = \frac{1}{\sqrt{2}}|S_z;\uparrow\rangle + \frac{1}{\sqrt{2}}|S_z;\downarrow\rangle,$$

so the spin state parallel to the positive x-axis is actually a superposition of the spin states parallel to the positive and negative z-axis. A Stern Gerlach experiment confirms this.



After the last $\boxed{\mathrm{SG}\hat{\boldsymbol{z}}}$ measurement we see the beam splitting again into two equally intensive componenents. The experiment tells us that there are quantitities which cannot be measured simultaneously. In this case it is impossible to determine simultaneously both the z- and x-components of the spin. Measuring the one causes the atom to go to a state where both possible results of the other are present.
We know that
Theorem 3 *Commuting operators have common eigenvectors.*
When we measure the quantity associated with an operator A the system goes to an eigenstate $|a'\rangle$ of A. If now B commutes with A, i.e.

$$[A, B] = 0,$$

then $|a'\rangle$ is also an eigenstate of B. When we measure the quantity associated with the operator B while the system is already in an eigenstate of B we get as the result the corresponding eigenvalue of B. So, in this case we can measure both quantities simultaneously.
On the other hand, S_x and S_z cannot be measured simultaneously, so we can deduce that

$$[S_x, S_z] \neq 0.$$

So, in our example a single Stern Gerlach experiment gives as much information as possible (as far as only the spin is concerned), consecutive Stern Gerlach experiments cannot reveal anything new.
In general, if we are interested in quantities associated with commuting operators, the states must be characterized by eigenvalues of all these operators. In many cases quantum mechanical problems can be reduced to the tasks to find the set of all possible commuting operators (and their eigenvalues). Once this set is found the states can be classified completely using the eigenvalues of the operators.