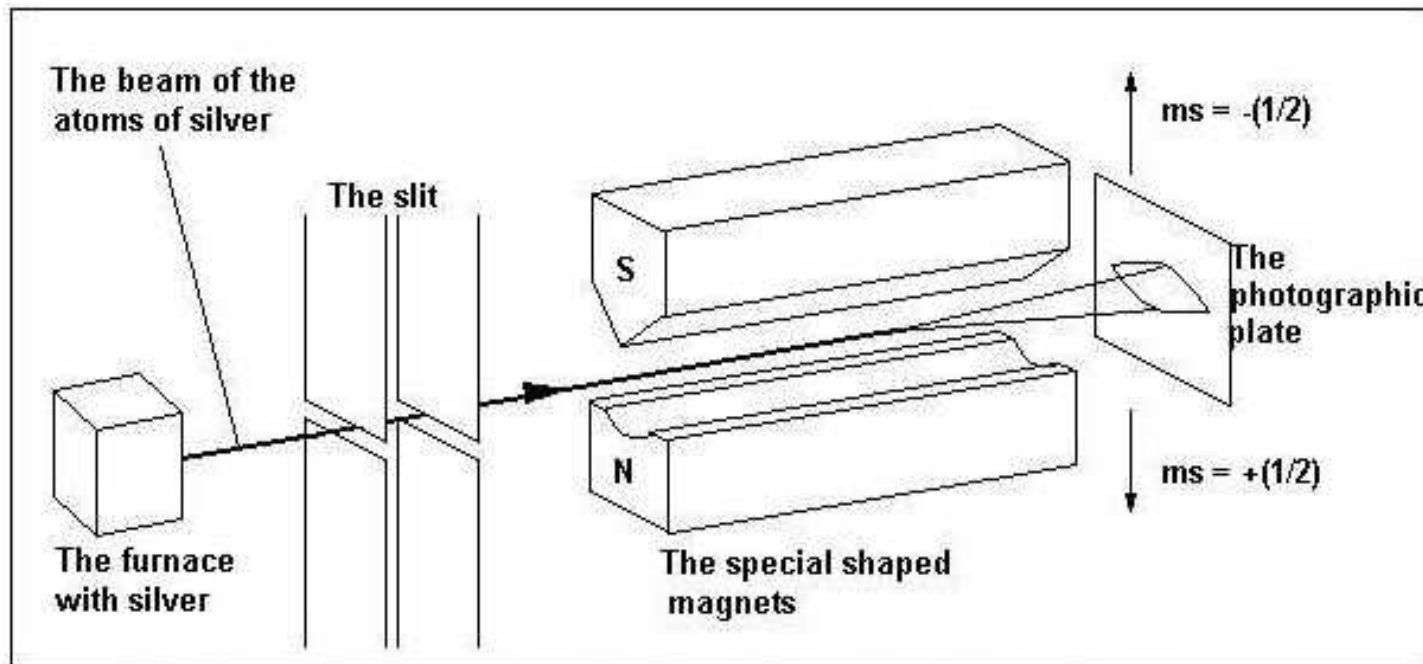


The Stern-Gerlach experiment and spin

Experiments in the early 1920s discovered a new aspect of nature, and at the same time found the simplest quantum system in existence. In the Stern-Gerlach experiment, a beam of hot atoms is passed through a *nonuniform* magnetic field. This field would interact with the magnetic dipole moment of the atom, if any, and deflect it.



The Stern-Gerlach experiment. On the photographic plate are two clear tracks.

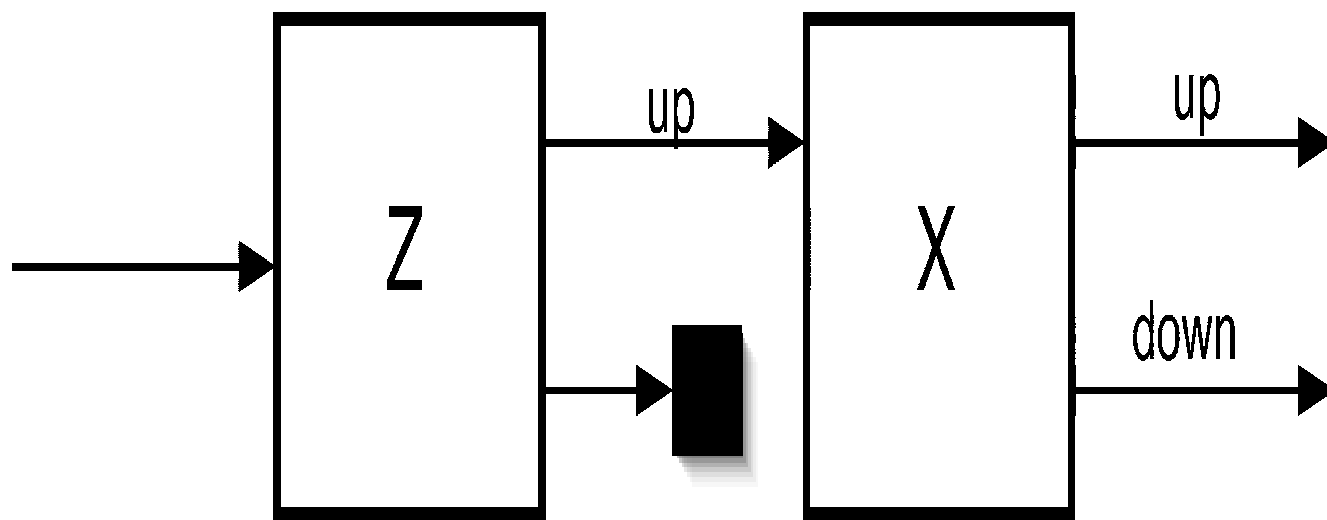
- This experiment discovered two surprising things. The atoms—specifically, the unpaired outer electron—did have a magnetic dipole moment. In effect, in addition to being charged, electrons acted like tiny bar magnets. They also, as it developed, have a tiny intrinsic amount of angular momentum, equal to $\hbar/2$. (This quantity is called *spin*, and all known elementary particles have nonzero spin.) Electrons are called *spin-1/2* particles.
- The second surprising thing was *how much* the path of the electrons was deflected. If electrons were really bar magnets, they could be oriented in any direction. The component oriented along the magnetic field gradient (say the Z direction) would determine the force on the electron, and hence how much it would be deflected.

- If electrons were like ordinary magnets with random orientations, they would show a continuous distribution of paths. The photographic plate in the Stern-Gerlach experiment would have shown a continuous distribution of impact positions.
- What was observed was quite different. The electrons were deflected either up or down by a constant amount, in roughly equal numbers. Apparently, the Z component of the electron's spin is *quantized*: it can take only one of two discrete values. We say that the spin is either *up* or *down* in the Z direction.

Sequential measurements

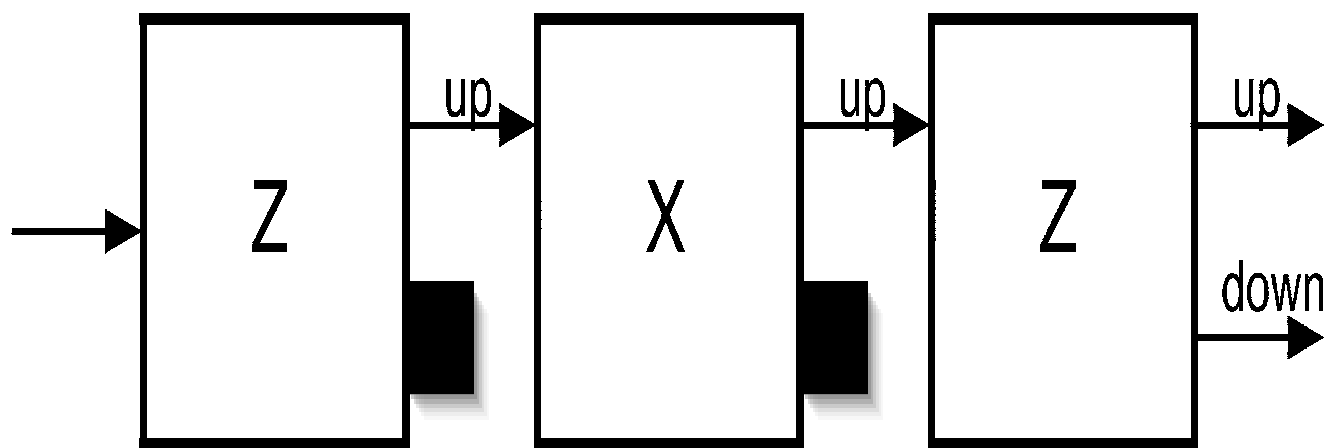
- Suppose an electron is passed through a Stern-Gerlach device, and is found to have spin up in the Z direction. If we pass it through a *second* Stern-Gerlach device, it will always be found to still have spin up in the Z direction. So this seems like an actual property of the spin, which we are measuring with the Stern-Gerlach device.
- In a similar way, we can tilt the Stern-Gerlach apparatus 90 degrees on its side and measure the component of spin in the X direction. Here again, that component of the spin is discrete: it is either up or down in the X direction. In fact, we can rotate it by any angle θ that we like, and measure the component of the spin in *any* direction; and it will always be found to have a discrete value, up or down, in that direction.

Suppose that we have determined the spin to be up in the Z direction, and we pass the spin through a second device to measure the X component of the spin. In this case, we get spin up or down in the X direction with *equal* probabilities. If we start with Z down, the same thing happens.



Suppose now that we measure Z up and then X up. What happens if we measure Z again?

In this case, we get Z up or down with equal probabilities!
Measuring X has *erased* our original measurement of Z .
Similarly, if we started with a definite state of X and
measure Z , we erase the original value of X .



By a more complicated arrangement, we can measure the
component of spin in the direction Y . If we do, we find that
measuring Y erases the value of either X or Z ; measuring
 X erases Y or Z ; and measuring Z erases X or Y .

Complementarity and randomness

- The X , Y , and Z components of the electron spin are all *complementary variables*. Knowing one of the three precludes knowing the other two. They are not all simultaneously well-defined. If a given variable is *not* well-defined for a given state of the system, when we measure it the outcome is *random*.
- Suppose that a spin is up in the Z direction. If we measure the spin component along an axis at angle θ to the Z axis, we find the spin up along that axis with probability $p_{\text{up}} = \cos^2(\theta/2)$. This is not special to the Z direction. If the spin is initially up along a given axis, measuring along an axis at angle θ has outcomes up or down with probabilities $\cos^2(\theta/2)$ and $\sin^2(\theta/2)$. We need a mathematical framework to encompass these results.

The mathematical description of spin

- Let us first choose a particular direction for spin measurements. By convention, this is usually Z .
- There are two special states: spin definitely up or definitely down in the Z direction. Let us call these **ZUP** and **ZDOWN**. How do we represent a spin up along an axis at an angle θ to the Z axis? We might try:

$$\psi(\theta) = \cos(\theta/2)\mathbf{ZUP} + \sin(\theta/2)\mathbf{ZDOWN}.$$

- The values $\cos(\theta/2)$ and $\sin(\theta/2)$ are the *amplitudes* in the up and down directions; and if we measure the Z component of spin, we get up or down with probabilities $\cos^2(\theta/2)$ and $\sin^2(\theta/2)$ (the amplitudes squared).

- If the spin is up or down in the X direction, $\theta = \pi/2$, we would write the states

$$\begin{aligned} \mathbf{XUP} &= (\mathbf{ZUP} + \mathbf{ZDOWN})/\sqrt{2}, \\ \mathbf{XDOWN} &= (\mathbf{ZUP} - \mathbf{ZDOWN})/\sqrt{2}. \end{aligned}$$

- Of course, we could equally well have written it like this:

$$\begin{aligned} \mathbf{ZUP} &= (\mathbf{XUP} + \mathbf{XDOWN})/\sqrt{2}, \\ \mathbf{ZDOWN} &= (\mathbf{XUP} - \mathbf{XDOWN})/\sqrt{2}. \end{aligned}$$

- That works fine for the X direction. But what about the Y direction? That is also at an angle $\pi/2$ to the Z axis. But \mathbf{YUP} and \mathbf{YDOWN} can't be the same as \mathbf{XUP} and \mathbf{XDOWN} !

- For the Y direction, we avoid the problem by letting the amplitudes be *complex numbers*:

$$\begin{aligned} \mathbf{YUP} &= (\mathbf{ZUP} + i\mathbf{ZDOWN})/\sqrt{2}, \\ \mathbf{YDOWN} &= (\mathbf{ZUP} - i\mathbf{ZDOWN})/\sqrt{2}. \end{aligned}$$

- The most general state, then, can be written

$$\psi = \alpha\mathbf{ZUP} + \beta\mathbf{ZDOWN},$$

where α, β are complex numbers with $|\alpha|^2 + |\beta|^2 = 1$. (This is the *normalization condition*.) We can think of \mathbf{ZUP} and \mathbf{ZDOWN} as being basis vectors for a two-dimensional complex vector space.

- In terms of this basis we can write any state as a column vector:

$$[\psi]_Z = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

We will make the assumption that **ZUP** and **ZDOWN** are *orthogonal* vectors of *unit length*. By making this assumption, the basis we have chosen is *orthonormal*.

- Of course, we needn't choose **ZUP** and **ZDOWN** to be our basis. We could choose **XUP** and **XDOWN** instead, or **YUP** and **YDOWN**. Those would *also* be orthonormal bases. In fact, the states that represent up and down along any axis will form an orthonormal basis.

- We can change from the Z basis to a different basis, say the X basis, by applying a linear transformation. We multiply $[\psi]_Z$ by the transition matrix

$$[\psi]_X = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

- In general, the matrix \hat{U} that is used to change between two orthonormal bases is a *unitary matrix*,
 $\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}$.
- The probability of spin up or down along the Z axis is given by $|\alpha|^2$ and $|\beta|^2$. What is the probability to measure spin up or down along a different axis?

It turns out that the following rule gives the right answer:

1. Find the orthonormal basis corresponding to up and down along the desired axis.
2. Transform the vector $[\psi]_Z$ to $[\psi]_S$ for the new basis S .
3. The probabilities of up and down are the absolute values squared of the two components in the new basis.
4. After the measurement, the state is up (or down) along the new axis.

For an axis at an angle θ from the Z axis (towards the X axis), the basis vectors are $\cos(\theta/2)\mathbf{ZUP} + \sin(\theta/2)\mathbf{ZDOWN}$ and $\sin(\theta/2)\mathbf{ZUP} - \cos(\theta/2)\mathbf{ZDOWN}$.

- We can write a state ψ in terms of any orthonormal basis, as a linear combination of two orthogonal vectors corresponding to spin up and down along some axis. These are called *orthogonal decompositions* of ψ .
- Every orthogonal decomposition of ψ (i.e., every orthonormal basis) corresponds to a measurement that could be made on ψ (i.e., a possible axis along which to determine the spin component). And every possible measurement of ψ corresponds to a particular choice of orthogonal decomposition. (Obviously, there is some arbitrariness in what we call “up” and “down.”)
- Only if one component vanishes for a particular decomposition does the corresponding measurement have a definite ($p = 1$) outcome. For a given state ψ this will be true only for one possible measurement.

Inner products

- If we take the inner product of two states ψ and ψ' , then $|\psi \cdot \psi'|^2$ is the probability of getting result ψ when measuring a spin, in the initial state ψ' , along the axis for which ψ is spin up. Symmetrically, it is also the probability of getting ψ' when measuring a spin, in the initial state ψ , along the axis for which ψ' is spin up.
- The inner product of two vectors is independent of the choice of orthonormal basis. For instance,

$$[\psi]_X \cdot [\psi']_X = [\psi]_Y \cdot [\psi']_Y = [\psi]_Z \cdot [\psi']_Z.$$

- However, the inner product of coordinate vectors in *different* bases is not generally meaningful.

Global phase

- Suppose we multiply the state ψ by a *pure phase* $\exp(i\phi)$,

$$\alpha \rightarrow e^{i\phi}\alpha, \quad \beta \rightarrow e^{i\phi}\beta.$$

This doesn't change the probabilities for a measurement along the Z axis. Nor does it change the probabilities for a measurement along *any other* axis.

- This means that multiplying by a pure phase has *no observable physical consequences*. We say that the *global phase* of a state is *arbitrary*.
- (If we multiplied α and β by *different* phases that *would* have observable consequences. It still wouldn't change the probabilities for a Z measurement, but it would change the probabilities for other measurements.)

Bloch Sphere Representation

- If we fix the normalization $|\alpha|^2 + |\beta|^2 = 1$ and the global phase (so, for instance, α is real), then there are only two independent parameters for the state of a spin. One useful choice of parameters is

$$\psi = \cos(\theta/2)\mathbf{ZUP} + e^{i\phi} \sin(\theta/2)\mathbf{ZDOWN},$$

where $0 \leq \theta \leq \pi$ and $-\pi \leq \phi \leq \pi$. The parameters (θ, ϕ) are the coordinates of points on the surface of a sphere. This is called the *Bloch Sphere Representation*.

- Each point on the sphere corresponds to a state; opposite states are orthogonal. **ZUP** and **ZDOWN** are the north $(\theta, \phi) = (0, 0)$ and south $(\pi, 0)$ poles; **XUP** and **XDOWN** are $(\theta, \phi) = (0, \pi/2), (\pi, \pi/2)$; and **YUP** and **YDOWN** are $(\theta, \phi) = (\pi/2, \pm\pi/2)$.

The Pauli Matrices

- Since ZUP and ZDOWN are orthogonal, we can find a Hermitian matrix $\hat{Z} = \hat{Z}^\dagger$ such that ZUP and ZDOWN are eigenvectors with eigenvalues ± 1 . We can find similar matrices for X and Y . In the standard Z basis, these matrices are:

$$\hat{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

- These are our first examples of what in quantum mechanics are called *observables*. These three matrices are called the *Pauli Spin Matrices*, and they have a number of important mathematical properties.

- The Pauli matrices have unique algebraic properties:

$$\hat{X}\hat{Y} = -\hat{Y}\hat{X} = i\hat{Z}.$$

$$\hat{Y}\hat{Z} = -\hat{Z}\hat{Y} = i\hat{X}.$$

$$\hat{Z}\hat{X} = -\hat{X}\hat{Z} = i\hat{Y}.$$

These three matrices are *anticommutative*.

- All three matrices are both *traceless* and *idempotent*:

$$\text{Tr}\{\hat{X}\} = \text{Tr}\{\hat{Y}\} = \text{Tr}\{\hat{Z}\} = 0,$$

$$\hat{X}^2 = \hat{Y}^2 = \hat{Z}^2 = \hat{I}.$$

So in addition to being *Hermitian*, they are also *Unitary*.

- Together with the identity, the Pauli matrices form a basis for all 2×2 matrices. Any 2×2 matrix \hat{O} can be written as a linear combination

$$\hat{O} = a\hat{I} + b\hat{X} + c\hat{Y} + d\hat{Z}$$

for some complex numbers a, b, c, d . From the algebraic properties of $\hat{X}, \hat{Y}, \hat{Z}$, it is easy to see that

$$a = \frac{1}{2}\text{Tr}\{\hat{O}\}, \quad b = \frac{1}{2}\text{Tr}\{\hat{X}\hat{O}\},$$

$$c = \frac{1}{2}\text{Tr}\{\hat{Y}\hat{O}\}, \quad d = \frac{1}{2}\text{Tr}\{\hat{Z}\hat{O}\}.$$

- This basis for the 2×2 matrices is often useful in quantum information theory.

Properties of the spin-1/2

Let us summarize the properties we have discovered about the spin-1/2 system, the simplest quantum system:

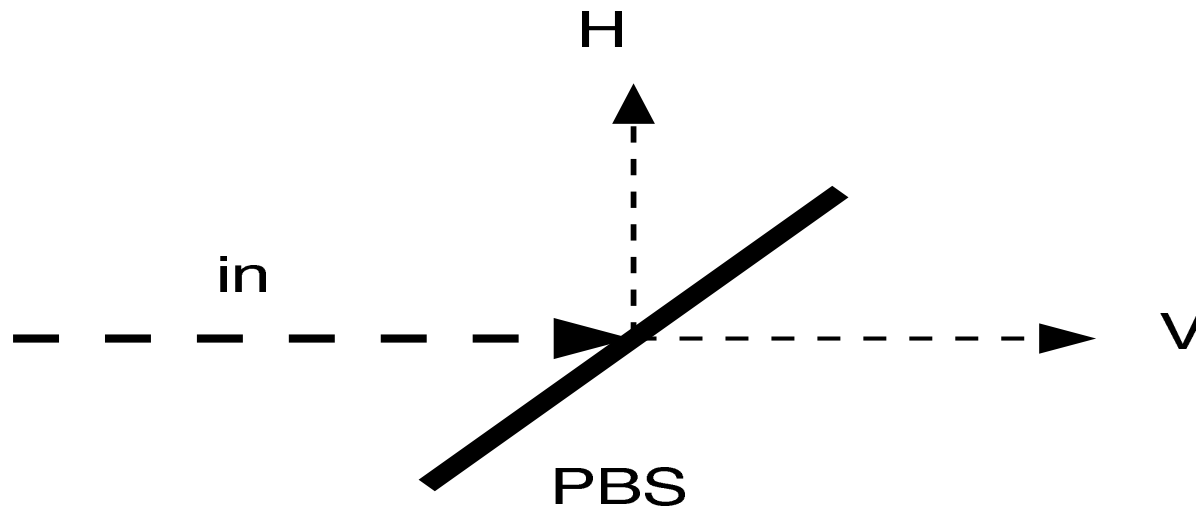
1. Every state of the spin can be represented by a (complex) vector of unit length.
2. Any state can be written as a linear combination (superposition) of orthonormal basis vectors.
3. Possible measurements of the spin correspond to choices of orthonormal bases.
4. The probabilities of different measurement outcomes are the squares of the orthogonal components (amplitudes).

And a few more:

5. We can find Hermitian operators whose eigenvalues and eigenvectors correspond to the outcomes of particular measurements.
6. Making a measurement of a particular variable *erases* (or disturbs) the values of all complementary variables.
7. Basis changes are done by multiplying by a unitary matrix.
8. The inner product of two states is unchanged by changing the basis.
9. The global phase has no physical meaning.

Photon Polarization

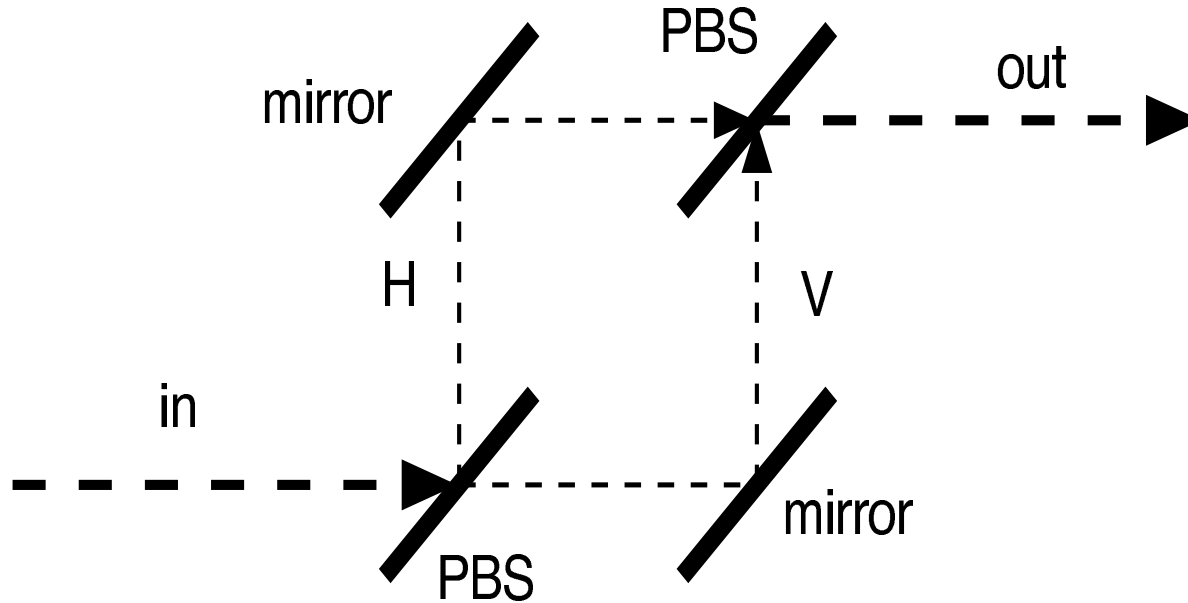
While the spin-1/2 is extremely simple, there are other physical systems that behave in the same way. The most well-known is the *photon polarization*. If light shines on a polarizing beam splitter (PBS), orthogonally polarized light (say H and V) exits from the two ports.



If a single photon arrives at a PBS, it exits from one of the two ports *with some probability*. A polarizing beam splitter for a photon acts just like a Stern-Gerlach apparatus for spin-1/2!

Interferometry

Note that having split the two components H and V, we can rejoin them to reconstruct the original state:



(In principle, we can do this with the Stern-Gerlach apparatus as well; this is called *matter interferometry*.) So we see that “measurements” are not necessarily final until the actual *read-out* process is complete. Bohr called this final step an “irreversible classical amplification.”

Quantum Bits

- The mathematical description of spin-1/2 maps directly onto photon polarization. The states ZUP and ZDOWN become the linear polarization states H and V. XUP and XDOWN become the linear polarizations at 45 degrees to H and V , and YUP and YDOWN become the *circular* polarizations $H \pm iV$. Just like Stern-Gerlach devices, we can construct PBSs to measure any polarization. The probabilities of different outcomes obey the same mathematics as the spin-1/2.
- Many other systems act the same way: the two energy levels of the hyperfine splitting, the flux in a superconducting loop, etc. In these cases, we are picking out a *two-dimensional subspace* of a larger space. All such two-dimensional systems are instances of *quantum bits*—the fundamental units of quantum information.