Spin

1 Introduction

For the past few weeks you have been learning about the mathematical and algorithmic background of quantum computation. Over the course of the next couple of lectures, we'll discuss the physics of making measurements and performing qubit operations in an actual system. In particular we consider nuclear magnetic resonance (NMR). Before we get there, though, we'll discuss a very famous experiment by Stern and Gerlach.

2 Spins as quantized magnetic moments

The quantum two level system is, in many ways, the simplest quantum system that displays interesting behavior (this is a very subjective statement!). Before diving into the physics of two-level systems, a little on the history of the canonical example: electron spin.

In 1921, Stern proposed an experiment to distinguish between Larmor's classical theory of the atom and Sommerfeld's quantum theory. Each theory predicted that the atom should have a magnetic moment (i.e., it should act like a small bar magnet). However, Larmor predicted that this magnetic moment could be oriented along any direction in space, while Sommerfeld (with help from Bohr) predicted that the orientation could only be in one of two directions (in this case, aligned or anti-aligned with a magnetic field). Stern's idea was to use the fact that magnetic moments experience a linear force when placed in a magnetic field gradient. To see this, not that the potential energy of a magnetic dipole in a magnetic field is given by:

$$U = -\vec{\mu} \cdot \vec{B}$$

Here, $\vec{\mu}$ is the vector indicating the magnitude and direction of the magnetic moment. The direction of the moment is analogous to the orientation of a bar magnet. This expression for the potential energy can also be used to derive a force that acts on the dipole (dipole is just another name for something that posesses a magnetic moment). Recall that the force is defined as the negative gradient of the potential:

$$F = -\nabla U = \nabla \left(\vec{\mu} \cdot \vec{B} \right)$$

Let's suppose that the magnetic field looks like $\vec{B} = B_0 z \hat{z}$; this field doesn't satisfy Maxwell's equations, but it makes the analysis easier. We get a force,

$$F = \nabla \left(\vec{\mu} \cdot B \right) = \nabla \left(\mu_z B z \right) = \mu B \cos(\theta) \, \hat{z}$$

The $\cos\theta$ term comes from the dot product. If, for example, the dipole is initially aligned with the field, it will experience an 'upward' force, and if it is antialigned, it will experience a 'downward' force.

Now consider a beam of dipoles passing through this field gradient. Larmor's classical theory predicts that the dipole moment could point in *any* direction, so the beam would spread out homogenously. The Bohr-Sommerfeld theory, though, predicts that the dipole moment can take only two values, aligned or anti-aligned with the field, so the beam would be *split* into two beams.

In 1922 Gerlach performed this experiment using silver atoms. (It turns out that electrons are a bad choice for this experiment because they are also affected by the Lorentz force, which is proportional to their

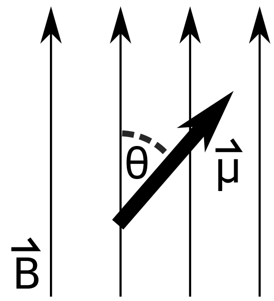

Figure 1: A magnetic moment in a magnetic field.

velocity. Any spread in the initial velocity causes a spread in the output that overwhelms the spin-gradient force.) Gerlach saw his beam split into two distinct beams, thereby demonstrating the spatial quantization of the magnetic moment and falsifying the Lorentz theory. In an interesting twist, the Sommerfeld theory was *also* incorrect, even though it predicted the correct result of this experiment. In 1925/1926, Uhlenbeck and Goudspit postulated that the electron carried its own spin magnetic moment independent of its orbital angular momentum.

In any case, the Stern-Gerlach experiment provides a toy model that we can use to learn about quantum two level systems. Let's make the language a but more precise by labeling some things:

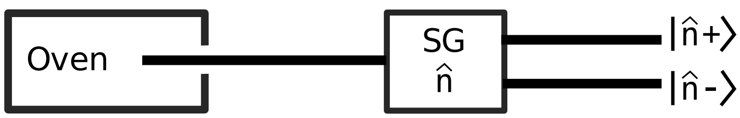

Figure 2: Schematic of Stern-Gerlach device.

In this diagram we have a cartoon picture of the Stern-Gerlach device. An oven produces a beam of particles which enters a reion with an inhomogenrous magnetic field, that gradient of which points in the $\hat{n}$ direction. The two beams that emerge we label $|\hat{n}+\rangle$ and $|\hat{n}-\rangle$. These symbols are what we use to label a quantum state, and what is written inside gives some information about this particular state. However, we should stress that what we write inside the ket is simply a label that we give to help us remember how the state behaves. We could have just as easily called the two states $|\text{Bob}\rangle$ and $|\text{Alice}\rangle$. Here, though, $|\hat{n}+\rangle$ denotes the state that is directed upwards when the field gradient is along the vector $\hat{n}$, and $|\hat{n}-\rangle$ is the state directed downwards.

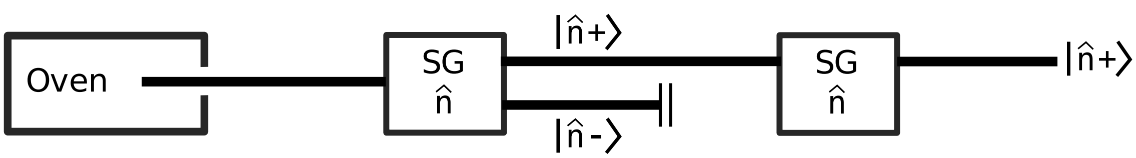

Figure 3: Cascaded Stern-Gerlach devices.

This all becomes much more interesting if we consider multiple, cascaded Stern-Gerlach devices. Let's add another identical SG device, which we'll denote as SG($\hat{n}$), after the first, but we'll discard the $|\hat{n}-\rangle$ state.

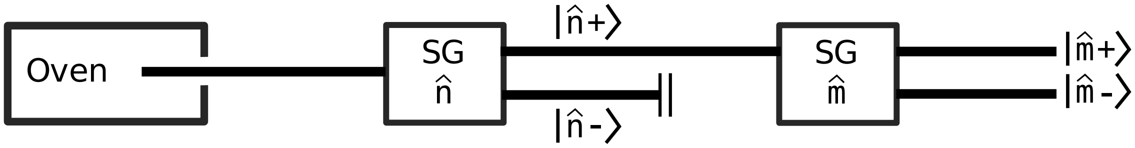

Figure 4: Cascaded Stern-Gerlach devices.

Notice that if we measure the output of the first SG($\hat{n}$) with a second SG($\hat{n}$) then we only get one beam out, the $|\hat{n}+\rangle$ state again. This shouldn't be surprising, as we have established with the first SG device that the dipole moment points along $\hat{n}$. But what would happen if we rotated the device?

Now we get two beams! The probability that $|\hat{n}+\rangle \to |\hat{m}\pm\rangle$ is found experimentally to be

$$P(|\hat{n}+\rangle \to |\hat{m}+\rangle) = \frac{1}{2}\left(1 + \hat{n}\cdot\hat{m}\right),$$

and also

$$P(|\hat{n}+\rangle \to |\hat{m}-\rangle) = \frac{1}{2}\left(1 - \hat{n}\cdot\hat{m}\right).$$

These probabilities can also be consdidered as the relatice intesitites of the two outgoing beams, $|\hat{m}\pm\rangle$, given an incoming beam $|\hat{n}+\rangle$.

Our task now is to seek a quantum mechanical description of this experiment. One way to do thisis to seach for the simplest description we can come up with, adding complexity only as we need to. Because the result of any measurement we can do is either "aligned" or "antialigned," the simplest model is the two-state system. We now consider the states $|\hat{n}\pm\rangle$ to be quantum states which are orthogonal to one another. Because we have two orthogonal states in a two-level space, these states form a basis. The states $|\hat{m}\pm\rangle$ also form a different basis. Let's pick a special basis, $|\hat{z}\pm\rangle$ and express all other states as linear combinations of these vectors. To indicate that this basis is special, let's relabel the vectors:

$$|\hat{z}+\rangle \to |0\rangle$$

$$|\hat{z}-\rangle \to |1\rangle$$

We can now represent any other state as a linear combination of the $\{|0\rangle, |1\rangle\}$ states:

$$|\hat{n}+\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,,$$

where α and β are complex numbers. Now that we know, roughly, how to express our states, we ned to figure out what the α and β are. We can start by considering the following double Stern-Gerlach experiment.

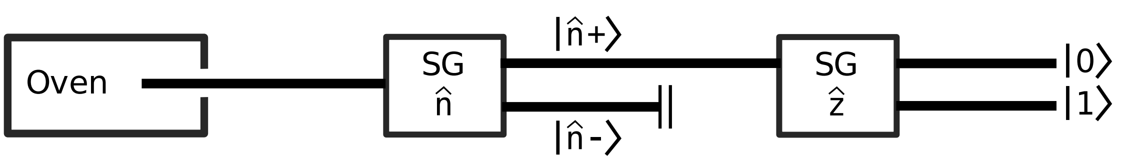

Figure 5: Cascaded Stern-Gerlach devices.

So, given the initial state $|\hat{n}+\rangle$, the probability of measuring $|0\rangle$ is

$$P_0(\hat{n}+) = |\langle 0|\hat{n}+\rangle|^2 = |\alpha\,\langle 0|0\rangle + \beta\,\langle 0|1\rangle|^2 = |\alpha|^2 = \frac{1}{2}\,(1 + \hat{n}\cdot\hat{z})$$

We have used that fact that $\langle 0|1\rangle = 0$ and $\langle 0|0\rangle = 1$. To make this a bit nicer, we are going to rewrite $\hat{n}\cdot\hat{z} = \cos\theta$, where θ is the angle between the two unit vectors $\hat{n}$ and $\hat{z}$. This angle is also equal to the spherical coordinate, θ, that corresponds to $\hat{n}$:

$$\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

Applying the trig identity, $(1+\cos\theta)/2 = \cos^2(\theta/2)$, we can say that:

$$|\alpha| = \cos(\theta/2)$$

A similar analysis shows that $|\beta| = \sin(\theta/2)$. So this simple argument has given us the magnitudes of α and β, but what about their phases? Because they are complex, they can be written as:

$$\alpha = |\alpha|\,e^{i\psi} \qquad\qquad \beta = |\beta|\,e^{i\chi}$$

giving,

$$|\hat{n}+\rangle = |\alpha|\,e^{i\psi}\,|0\rangle + |\beta|\,e^{i\chi}\,|1\rangle = \cos(\theta/2)e^{i\psi}\,|0\rangle + \sin(\theta/2)e^{i\chi}\,|1\rangle$$

However, a quantum state is only defined up to an overall phase, so we can multiply this state by $e^{-i\psi}$ to get

$$|\hat{n}+\rangle = \cos(\theta/2)\,|0\rangle + \sin(\theta/2)e^{i(\chi-\psi)}\,|1\rangle \equiv \cos(\theta/2)\,|0\rangle + \sin(\theta/2)e^{i\phi}\,|1\rangle\,,$$

where $\phi = \chi - \psi$ is the relevant phase. So what is ϕ for a given $|\hat{n}+\rangle$? Let's look at $|\hat{x}+\rangle$ and $|\hat{y}+\rangle$:

$$|\hat{x}+\rangle = \frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}e^{i\phi_x}\,|1\rangle$$

$$|\hat{y}+\rangle = \frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}e^{i\phi_y}\,|1\rangle$$

If a double SG device is set up, with $\hat{x}$ first, then $\hat{y}$, the probability of seeing $|\hat{y}+\rangle$ given that the first beam emerged in $|\hat{x}+\rangle$, is $P(x\to y) = (1+\hat{x}\cdot\hat{y})/2 = 1/2$. But this is also the overlap between the two states:

$$\begin{aligned}
P(\hat{x}\to\hat{y}) &= |\langle\hat{x}+|\hat{y}+\rangle|^2 \\
&= \left|\frac{1}{2}\,\langle 0|0\rangle + \frac{1}{2}e^{i(\phi_y-\phi_x)}\,\langle 1|1\rangle\right|^2 \\
&= \frac{1}{2} + \frac{1}{2}\cos(\phi_y-\phi_x) \\
&= \frac{1}{2}
\end{aligned}$$

So, $\cos(\phi_y - \phi_x) = 0 \to \phi_y - \phi_x = \pi/2$. This implies that we can associate the phase angle, ϕ_n, with the second spherical coordinate, ϕ, in:

$$\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

So, for any vector $\hat{n}$, $|\hat{n}+\rangle = \cos(\theta/2)\,|0\rangle + \sin(\theta/2)e^{i\phi}\,|1\rangle$, where (θ,ϕ) are the polar coordinates of the vector $\hat{n}$. But what about $|\hat{n}-\rangle$? The states $|\hat{n}+\rangle$ and $|\hat{n}-\rangle$ must be orthogonal. This gives a unique solution, up to the standard quantum mechanical phase,

$$|\hat{n}-\rangle = \sin(\theta/2)\,|0\rangle + \cos(\theta/2)e^{-i\phi}\,|1\rangle$$

But this is the same as

$$|(-\hat{n})+\rangle = \cos((\pi - \theta)/2)\,|0\rangle + \sin((\pi - \theta)/2)e^{-i\phi}\,|1\rangle = |\hat{n}-\rangle$$

So, $|\hat{n}+\rangle$ is orthogonal to $|(-\hat{n})+\rangle$! This representation we have been using (θ, ϕ as parameters for a qubit state) is known as the Bloch Sphere representation. Every point on the Bloch Sphere (a unit sphere in $\mathbb{R}^3$) corresponds to a unique state, with the orthogonal state being represented by the antipodal point on the sphere.

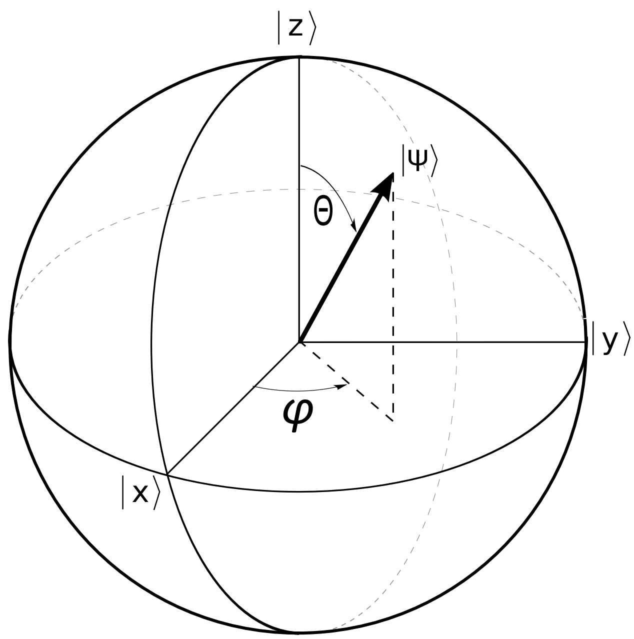

Figure 6: Bloch sphere representation of a qubit state.

We'll see a lot more of the Bloch Sphere representation over the course of the semester, so you'll have time to get accustomed to it.

Now that we've come up with a good way to represent the states in the Stern-Gerlach experiment, we need to come up with some way to mathematically describe the Stern-Gerlach devices (what we've been calling $SG(\hat{n})$. Before we can do this, however, we need to understand a little bit about how we measure quantum states.

Measurement is the assignment of a particular value to some attribute of the system under study. One of the postulates of quantum mechanics is that for every possible measurement you can do, there exists a Hermitian operator. Recall that a Hermitian operator, A, satisfies $A = A^\dagger$. The outcomes that are possible for the measurement are the eigenvalues of the operator. It might help to see an example:

Let's say we have some state: $|\hat{n}+\rangle = \cos(\theta/2)\,|0\rangle + \sin(\theta/2)e^{i\phi}\,|1\rangle$, and we want to measure its magnetic moment inthe $\hat{z}$ direction. We know from the discussion above that we should get one of two values, "aligned" or "antialigned." Because we're building a mathematical theory, let's call "aligned" = 1 and "antialigned" = -1. So now we have out eigenvalues. The eigenstates that correspond to these eigenvalues are the states $|\hat{z}+\rangle = |0\rangle$ and $|\hat{z}+\rangle = |1\rangle$. In order to represent out operators as matrices, we need to represent our states as vectors,

$$|\hat{z}+\rangle = |0\rangle \longrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad\qquad |\hat{z}-\rangle = |1\rangle \longrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

In this notation,

$$|\hat{n}+\rangle \longrightarrow \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2)e^{i\phi} \end{pmatrix} \qquad |\hat{n}-\rangle \longrightarrow \begin{pmatrix} \sin(\theta/2) \\ -\cos(\theta/2)e^{i\phi} \end{pmatrix}$$

So, our task now is to find a matrix $S(\hat{z})$, that has eigenvalues ± 1 and eigenvectors $(1, 0)^{\mathsf{T}}$ and $(0, 1)^{\mathsf{T}}$. This is easy:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Now we need to find $S(\hat{n})$, the matrix corresponding to a measurement along $\hat{n}$. The eigenvalues are still (± 1), but the eigenvalues are still ± 1, but the eigenvectors have changed. We have already written down the eigenvectors, $|\hat{n}+\rangle$ and $|\hat{n}-\rangle$, so we can use the eigenvalue decomposition of a matrix,

$$S = P\Lambda P^{-1}$$

where Λ is the diagonal matrix of eigenvalues and P is the matrix of eigenvectors. Let's use this to explicitly construct $S(\hat{x})$. The eigenvectors are

$$|\hat{n}+\rangle \longrightarrow \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \qquad |\hat{n}+\rangle \longrightarrow \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

So,

$$P_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$\Lambda_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Which gives,

$$S(\hat{x}) = P_x \Lambda_x P_x^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

This case was a little special, because $P_x = P_x^{-1}$. Doing the same for $S(\hat{y})$,

$$P_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

These three matrices we've just derived are very special. So special, in fact, that we are going to give them special names:

$$S(\hat{x}) \to \sigma_x \equiv X$$
$$S(\hat{y}) \to \sigma_y \equiv Y$$
$$S(\hat{z}) \to \sigma_z \equiv Z$$

Where $\sigma_{(x,y,z)}$ is the notation preferred by physicists and (X, Y, Z) is the notation preferred by computer scientists. These are the Pauli matrices, and you'll see them a lot this semester. A nice feature of them is that any matrix $S(\hat{n})$ can be written as

$$S(\hat{n}) = \hat{n} \cdot \vec{\sigma} = n_x\sigma_x + n_y\sigma_y + n_z\sigma_z$$

This is a nice exercise, and I recommend you try to show that you try to show that it's true.