

Lecture 1

Eslam Khalaf

September 6, 2023

A standard QM course usually takes a historical route showing how a series of unexplained experimental results lead scientists in the early 20th century to abandon many of the classical notions leading to the formulation of quantum mechanics as a fully fledged theory. In this course, we are going to take a different route. Imagine we have a late 19th century physicist time-travel to us nowadays and we want to convince them of quantum mechanics using a single experiment that most embodies quantum mechanics and would most easily illustrate the structure of theory. What would this experiment be? This experiment is the Stern-Gerlach experiment performed by Stern and Gerlach in Frankfurt in 1922.

1 The Stern-Gerlach experiment

In the Stern-Gerlach experiment, silver atoms are heated in a furnace that has a hole in it. A collimator is then used to generate a beam of silver atoms. This beam is then subjected to an **inhomogenous** magnetic field that can be produced using a magnet shaped as shown in Fig. 1. We now suppose that the silver atoms has some magnetic moment $\boldsymbol{\mu}$. At this point, we are assuming we do not really know anything about quantum mechanics, spin, or the atomic number of silver. We are just assuming that there is a quantity called magnetic moment that couples to magnetic field such that an object with magnetic moment $\boldsymbol{\mu}$ has an energy $\boldsymbol{\mu} \cdot \mathbf{B}$ in a magnetic field. A particle with magnetic moment $\boldsymbol{\mu}$ placed in an inhomogenous magnetic field will feel a force given by

$$F_a = \partial_a[\boldsymbol{\mu} \cdot \mathbf{B}(\mathbf{r})] = \boldsymbol{\mu} \cdot \partial_a \mathbf{B}(\mathbf{r}) \quad (1)$$

Assuming the field points in the z -direction, $\mathbf{B} = B\hat{z}$ and it only changes along the z -direction as well, we find the force in the z -direction to be

$$F_z = \mu_z \partial_z B(z) \quad (2)$$

We then have a screen at some distance that detects the silver atoms. If we make the reasonable assumption that the magnetic moments of silver atoms all have the same magnitude $\boldsymbol{\mu} = \mu_0$ ¹ and come out of the oven with random orientation, then we expect μ_z to have all possible values between $-\mu_0$ and $+\mu_0$ leading to a continuous bundle of beams which leaves a continuous line on the screen as shown in Fig. 1 .

Rather remarkably, this is not what is seen in experiment. Instead, the beam splits into only two beams, leaving two spots on the screen as shown in Fig. 1. This means that the atoms in the beam seem to have only two possible values of their magnetic moment $\boldsymbol{\mu} = \pm\mu_0\hat{z}$. First, let us try to find some plausible classical explanation for this observation. After thinking for a bit, we can see that the only possibility is that somehow the magnetic moments of the silver atoms are pointing in the z -direction with an equal probability to be pointing up or down. To check whether this is really the case, we can simply rotate the device. If the moments are pointing in the $\pm z$ -direction, then using a device rotated by an angle θ relative to the z -axis, let's say in the $x - z$ plane, we should see two spots corresponding to the projection of the z -magnetic moment along the axis of the device. This yields $\boldsymbol{\mu} = \pm\mu_0 \cos\theta$. In particular, if we rotate the device by $\pi/2$, we expect to get no splitting at all. This corresponds to the rather obvious statement that a classical vector pointing in some direction has zero projection in a perpendicular direction.

¹This assumption is not really needed. See Problem 1 in the first problem set.

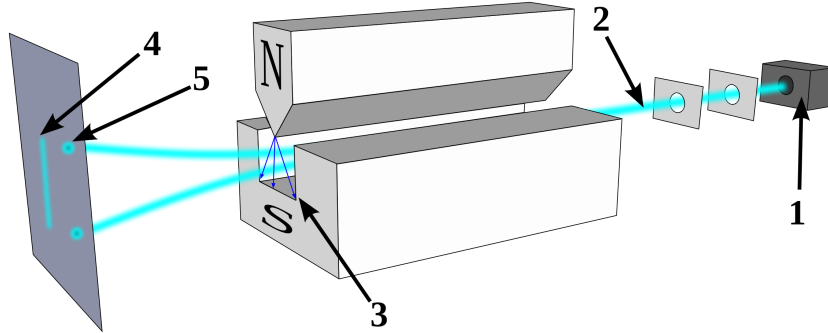


Figure 1: Stern–Gerlach experiment: Silver atoms travelling through an inhomogeneous magnetic field, and being deflected up or down depending on their spin; (1) furnace, (2) beam of silver atoms, (3) inhomogeneous magnetic field, (4) classically expected result, (5) observed result

The main surprise here is that this very reasonable expectation is false. No matter how we rotate the Stern-Gerlach device, we always find the beam split into two corresponding to magnetic moment $\pm\mu_0$ in the direction of the splitting. To make the discussion more quantitative, I will introduce the notation $SG\hat{n}$ to denote a Stern-Gerlach device which splits the beam according to the magnetic moment in the \hat{n} direction. For example, the device that splits the beam in the z direction is denoted SGz . We also introduce the notation $SG\hat{n}\pm$ to denote an $SG\hat{n}$ device followed by a filter that only lets the \pm beam through. Let us now consider the following cascaded SG setup where we first let the beam pass through $SGz+$ then let the output beam of that device go through $SG\hat{n}_\theta^{xz}$ where \hat{n}_θ is the unit vector in the xz -plane rotated by an angle θ relative to the z -axis. If $\theta = 0$, this is basically placing an SGz after an $SGz+$ and since we have filtered out the $-z$ component, we get only one beam corresponding to $\mu_z = +\mu_0$ (see top panel of Fig. 2). Once we start rotating the second SG so that $\theta \neq 0$, we immediately find that the beam is split in two corresponding to magnetic moments $\mu_{\hat{n}} = \pm\mu_0$. However, we should still expect the information about the rotation angle to be encoded somehow in the result of the experiment (otherwise, the limit $\theta \rightarrow 0$ will be discontinuous). The only other experimental variable where this info can be encoded is the intensity of the two beams. Assuming the original beam had an intensity I_0 , the \pm beams will have intensities $I_\pm(\theta)$ which satisfy

$$I_+(\theta) + I_-(\theta) = I_0 \quad (3)$$

For $\theta = 0$, we should have $I_+(0) = I_0$ and $I_-(0) = 0$. We also note that rotating the SG device by π exchanges the $+$ and $-$ beams which means that $I_\pm(\theta + \pi) = I_\mp(\theta)$. The simplest positive functions satisfying these properties are $I_+ = \cos^2(\theta/2)$ and $I_- = \sin^2(\theta/2)$ which turns out to be the right answer. This particularly means that for a device rotated by $\theta = \pi/2$, i.e. a device which splits the beam along the x -axis, we get two beams of equal intensity equal to half the original beam intensity. Thus, the beam whose magnetic moment is aligned in the $+z$ direction can be decomposed into an equal mix of beams whose magnetic moments are along $\pm x$ direction. By now, it is clear that this behavior is incompatible with any classical vector aligned with the z axis.

Let us consider one final setup where we have a succession of $SGz+$ and $SG+x$ devices followed by SGz (bottom panel of Fig. 2). Here, we find that we get an equal amount of $\mu_z = \pm\mu_0$. Thus, although

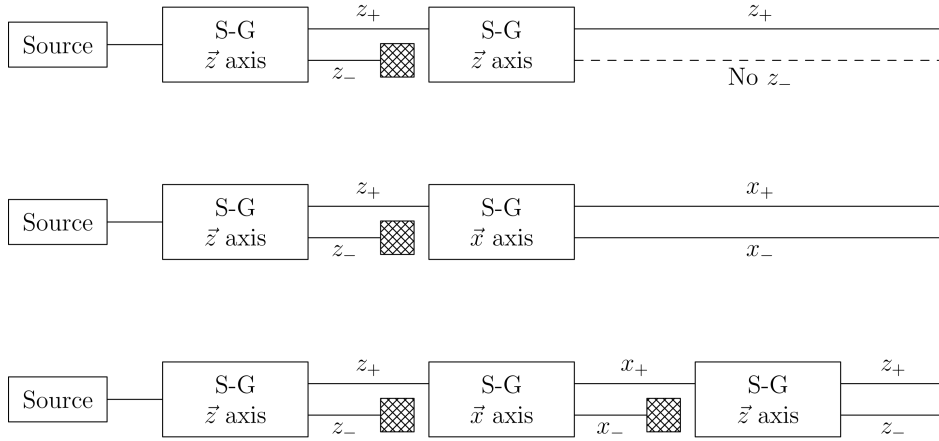


Figure 2: Different configurations for cascaded Stern-Gerlach device.

the original beam had only $+z$ component, after passing through $SGx+$, the electrons forgot they were polarized in the $+z$ direction and they again split into two beams. This means that if we place a filter that selects the $-z$ component at the last stage, we will get an output even though the input in the second stage was purely polarized in the $+z$ direction. In fact, if we remove the middle part ($SGx+$), this output will vanish. Thus, rather remarkably, we can increase the output at the end of the device by blocking some intermediate beams. This is an indication of some interference physics as we will see later (if we waves interfere destructively, blocking one of them will remove the interference and increase the signal).

Although this is a very counter-intuitive result, it has classical analogs which are realized when we have a device that decomposes a vector into parallel and perpendicular components and filters out one of them. One very cool example is how sail boats manage to sail perpendicular to or even into the wind although they only use the force generated by the wind to move. The main part of a sailboat is the sail. If the sail is parallel to the wind, it feels no force. This means that for a general angle between the sail and the wind, the sail decomposes the wind into a parallel and perpendicular components and only picks out the perpendicular one. The next important part in a sailboat is something called the keel. It's the bottommost longitudinal part that prevents the boat from moving sideways. This part picks out the component of the force that is parallel to the body of the boat. If the sail is at an angle of 45° relative to the wind direction and the keel is an angle 45° relative to the sail, we can successfully convert the wind force from pointing along the y -direction to the x -direction as shown in Fig. 3c. By using different angles, we can even manage to have a force component against the wind direction as shown in Fig. 3d. As an exercise, try to identify the ingredients of the Stern-Gerlach setup in the sailing example above.²

The sailing example above hints towards the mathematical structure needed to describe the Stern-Gerlach experiment. An SG device defines a coordinate system in some abstract space of possible quantum states with two basis vectors. The SG device decomposes any given entering beam into a parallel and a perpendicular component, which can be identified by the \pm values of the magnetic moment. An $SG\hat{n}\pm$ device acts by projecting out one component, i.e. projecting the input on the direction on one of the two basis vectors. Finally, rotating the SG apparatus amounts to rotating coordinate frame or basis vectors. However, there is an important subtlety. If we identify the $\pm z$ magnetic moments with the x and y directions in some abstract space, we see that the angle in this abstract space between two vectors is *half* the physical angle between the directions of the magnetic moment. As a consequence, physically rotating the SG device by an angle θ corresponds to a rotation by $\theta/2$ in the abstract space describing the state. With all these observations, we can write out some expressions to capture what we discussed. We denote the “quantum state” of a beam that gets deflected in the \pm direction by an $SG\hat{n}$ device by $|\pm, \hat{n}\rangle$, where the notion of quantum state

²Of course we lose some of the force or intensity of the wind by doing this, so it is still easier to sail in the wind direction.

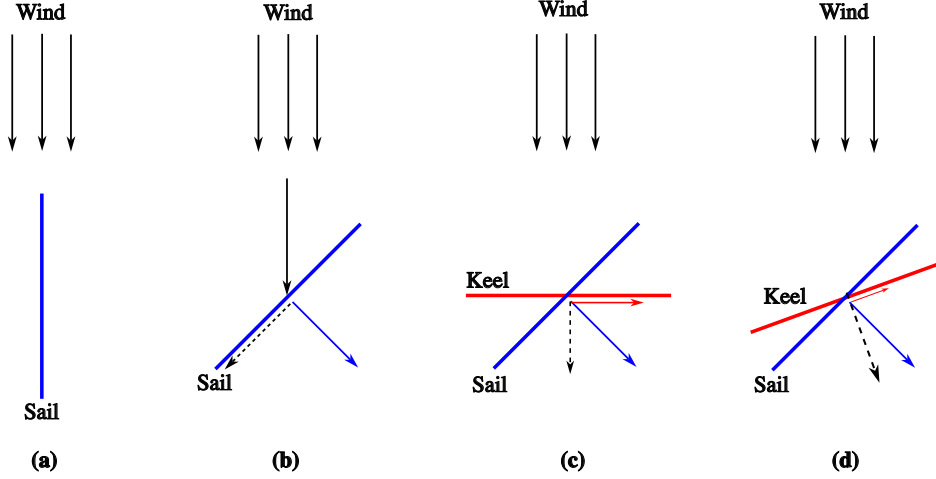


Figure 3: Illustration of how a sailboat transforms wind force: (a) when the sail is parallel to the wind, it feels no force. (b) when the sail is at an angle (45° here), it feels only the component of the force perpendicular to the sail. (c) A horizontal keel ensures the boat only picks up the x -component of force (the one parallel to the boat). (d) a different angle for the keel allowing the boat to sail against the wind.

will be defined precisely later. The results of the SG experiment can be summarized by the relation

$$\begin{aligned} |+, \hat{n}_\theta^{xz}\rangle &= \cos(\theta/2)|+, z\rangle + \sin(\theta/2)|-, z\rangle \\ |-, \hat{n}_\theta^{xz}\rangle &= -\sin(\theta/2)|+, z\rangle + \cos(\theta/2)|-, z\rangle \end{aligned} \quad (4)$$

Furthermore, we can see that the intensities $I_\pm(\theta)$ can be identified with the coefficients of these expansions. For example, the intensity of the $+$ beam coming out of a cascade of an SG_x whose input is polarized in the $+z$ direction is given by the square of the $|+, x\rangle$ coefficient of the expansion of the $|+, z\rangle$ in terms of the basis vectors $|\pm, x\rangle$.

While the theory we have right now seems to capture all aspects of the SG experiment, there is one crucial missing ingredient: how can we describe rotations in the third (y) direction? We can surely rotate our SG apparatus also in the yz plane and we expect the same results as we got for xz . However, using the same basis decomposition for yz as for xz (Eqs. 4) will mean that $|+, x\rangle$ and $|+, y\rangle$ are identical. However, clearly having a cascade of SG_x and SG_y will yield identical behavior to that of SG_z and SG_x we considered before.

To identify this last remaining ingredient, we invoke another classical analog. This is again something that our late 19th century physicist will be familiar with. From a classical perspective, light is an electromagnetic wave consisting of oscillating electric and magnetic fields (we are not discussing photons at all here). An electromagnetic wave is characterized by a quantity called polarization specifying the direction of the electric field vector. If that direction is a constant, we say the light is linear polarized. For example, a linear polarized light in the x -direction, propagating along the z -axis is described by the electric field

$$\mathbf{E} = E_0 \hat{x} \cos(kz - \omega t) \quad (5)$$

A device called polarizer can select a specific direction of linear polarization. Its simplest implementation is a series of parallel metallic rods which absorbs the electric field parallel to the rods but lets the field perpendicular to the rods through. For example, for the x -polarized light above, a polarizer in the y -direction lets it pass through whereas a polarizer in the x -direction completely blocks it. Thus, a general linear polarizer picks up the component of \mathbf{E} perpendicular to the direction of the polarizer. It is easy to see that the linear polarizer works exactly as an SG device followed by a filter which selects one of the two beams. In fact, we can realize the same results we got from the cascade of SGs with a cascade of linear polarizers. For example, an x -polarizer followed by a y -polarizer will let no light through. This acts the same way as

cascaded SG setup consisting of SG $z+$ followed by SG $z-$ which also has no output. The setup discussed in the bottom panel of Fig. 2 where an SG $x+$ is placed between SG $z+$ and SG $z-$ is the same as placing a linear polarizer at an angle of 45° between the x and y polarizers.

So far, the polarization example has reproduced what we understood about the SG setup. We have still not answered the question about including the third direction. To answer this, we recall that linear polarization is not the only possibility for an electromagnetic wave. There is also circular polarization and more generally elliptic polarization. These more general patterns of polarization arise if the two perpendicular components of an electromagnetic wave have a phase shift

$$\mathbf{E} = E_x \hat{x} \cos(kz - \omega t) + E_y \hat{y} \cos(kz - \omega t + \phi) \quad (6)$$

In general, this field will trace an ellipse in the x - y plane as the light propagates that reduces to a line if $\phi = 0$ and to a circle if $\phi = \pm\pi/2$ and $E_x = E_y$ (the two signs correspond to right and left circular polarization). Any EM wave can be decomposed into a circularly polarized wave and a linearly polarized wave which in turn can be decomposed into x and y components. Thus, the phase shift allows us to describe one more degree of freedom. Introducing a vector notation for the x and y components, we can write any EM wave as

$$\mathbf{E} = \text{Re}\{e^{i(kz-\omega t)} \mathbf{P}\}, \quad \mathbf{P} = \begin{pmatrix} E_x \\ E_y e^{i\phi} \end{pmatrix} \quad (7)$$

This means that by allowing the polarization vector to be *complex*, we can describe an arbitrary polarization state with the special cases of x and y -linearly polarized light given respectively by $\mathbf{P}_x = (1, 0)^T$ and $\mathbf{P}_y = (0, 1)^T$ whereas circularly polarized light given by $\mathbf{P}_{\text{circ}} = (1, \pm i)^T$ (Here, T denotes vector transposition that maps a row vector to a column vector). This provides us with the missing ingredient to understand the SG experiment if we identify the rotation into the y -direction by a relative phase shift between the two basis vectors and think of our abstract space as a complex rather than real vector space.

To do this explicitly, we can simply write

$$|+, y\rangle = \frac{1}{\sqrt{2}}(|+, z\rangle + i|-, z\rangle) \quad (8)$$

$$|-, y\rangle = \frac{1}{\sqrt{2}}(|+, z\rangle - i|-, z\rangle) \quad (9)$$

If we identify the intensities by the square absolute value of these complex coefficients, we see that the results of a cascaded SG experiment where SG $z+$ is followed by SG x or by SG y is identical. On the other hand, if we have SG $x+$ followed by SG y , we see the beam splits as expected.

Thus, if the only experiment our 19-th century physicist learned about is the Stern-Gerlach experiment, they will be lead to two main conclusions:

1. The magnetic moment of a silver atom and, in general, a microscopic object cannot be described by a classical vector in 3D space. Instead, the different directions should correspond to basis vectors in an abstract finite dimensional *complex* vector space.
2. We need a notion of length of the projection of a vector on another one to extract the measured intensities out of these abstract vectors i.e. we need a notion of an inner product.

In the next lecture, we will see how we can build the quantum theory formalism from the example of the Stern-Gerlach experiment. This example makes it clear that the proper mathematical structure is that of a complex vector space with an inner product associated with it. Such a space is called a Hilbert space.