1 Quantization of Light and Matter

1.1 Why bother building 'quantum mechanics' as a framework?

For this section, please refer: Day 1 | Why QM

1.2 Quantum Mechanics

1.2.1 Shifting from Classical to Quantum

In classical physics, the state of a physical system is specified by a set of real-valued dynamical variables. For a single particle moving in one dimension, these are typically its position $x(t)$ and momentum $p(t)$ as functions of time. More generally, for a system with N degrees of freedom, one introduces generalized coordinates $\{q_i\}$ and their conjugate momenta $\{p_i\}$.

The Hamiltonian formulation. A particularly powerful formulation of classical mechanics is the Hamiltonian formalism. In this approach, all dynamical information about the system is encoded in a single function

$$H(q_1, \ldots, q_N, p_1, \ldots, p_N, t),$$

called the *Hamiltonian*. Physically, the Hamiltonian corresponds to the total energy of the system, expressed in terms of coordinates and momenta. [1] For example, for a single particle of mass m moving in a potential $V(x)$,

$$H(x, p) = \frac{p^2}{2m} + V(x).$$

Given the Hamiltonian, the time evolution of the system is determined by *Hamilton's equations*:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

These equations are deterministic: if the initial values of all q_i and p_i are known exactly at some time t_0, the state of the system at all later times is uniquely determined.

Observables and predictions in classical mechanics. In the classical framework, any physical quantity of interest for example position, momentum, energy, angular momentum etc. is represented by a real-valued function $A(q, p, t)$ on phase space. Measurements are assumed to reveal pre-existing values of these quantities, and probabilities enter only through ignorance about initial conditions, not as a fundamental feature of the theory.

Why this framework is insufficient. The experiments discussed in the previous section reveal phenomena that cannot be consistently described within this classical structure.

- In interference experiments (such as the double-slit experiment with electrons), classical trajectories governed by Hamilton's equations cannot account for the emergence of interference patterns when particles are sent one at a time.
- In atomic spectroscopy, the classical atomic Hamiltonian $H = p^2/2m + V(r)$ predicts a continuous range of allowed energies, whereas experiments reveal discrete spectral lines.

[1] The corresponding poisson bracket is defined as

$$\{A, B\} = \sum_{i=1}^{N} \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right).$$

- In spontaneous emission, a classically bound charge interacting with the electromagnetic field does not radiate energy unless an explicit time-dependent perturbation is introduced. Classical electrodynamics predicts radiation from accelerating charges, but it cannot account for the observed irreducible decay of an excited atomic state in vacuum.

  This signals that the classical Hamiltonian description of matter *and* fields is incomplete at a fundamental level.

Importantly, these failures are not due to an incorrect choice of Hamiltonian, but to the deeper assumption that physical observables possess simultaneously well-defined values and evolve deterministically according to classical equations of motion.

The Stern-Gerlach Experiment

We present a historic experiment to motivate the postulates of quantum mechanics.

O. Stern and W. Gerlach carried out the following experiment (Figure 1) in 1922. They produced a beam of neutral silver atoms which was passed through an inhomogeneous magnetic field (such that the magnetic field was oriented a certain way, say along $\hat{\mathbf{z}}$). A glass plate was placed at the end to observe how the beam was deflected.

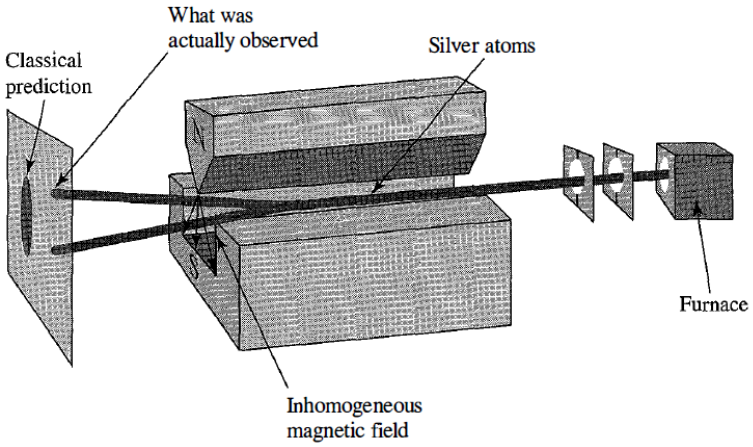

Figure 1: pictorial depiction of the Stern-Gerlach experiment, from Sakurai & Napolitano (2011)

Silver was specifically chosen because it has a single unpaired electron in its 5s orbital.[2] Recall that electrons possess a property called *spin*, which is intimately linked to its magnetic dipole moment. You would also recall that a magnetic dipole placed in an inhomogeneous magnetic field experiences a force as $F_B = \nabla(m \cdot B)$[3].

If spin were a classical quantity that could take any value, then you would expect that the beam of silver atoms would cause a single smear on the metal plate, since any given atom in the beam would have a randomly oriented magnetic dipole. (The cosine of the angle between m and B could take any value between -1 to 1, so a continuous smear is expected classically.)

[2]For a discussion on why silver was chosen for the Stern-Gerlach experiment, see Physics Stack Exchange | Why silver atoms were used in Stern-Gerlach experiment.

[3]Recall that potential energy U of a magnetic dipole in a magnetic field is $-m \cdot B$, and force is $-\nabla U$

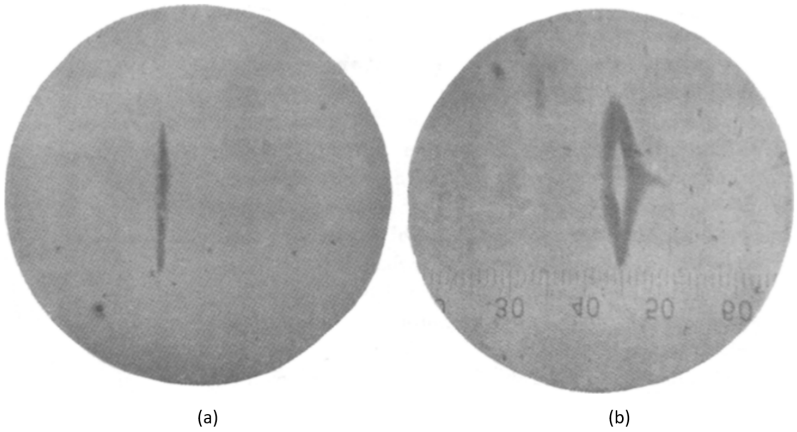

Figure 2: The result of the original Stern-Gerlach experiment, showing the deposition of silver on the glass plate. Please note that the plates are rotated by 90° in comparison to Figure 1. Here the separation occurs horizontally, not vertically.
(a): The control experiment where the beam of silver atoms is directly allowed to deposit on a glass plate without passing through an inhomogeneous magnetic field. The vertical extent of the deposition is due to the width of the slit used to produce a collimated beam.
(b): Clear separation of deposition. The separation is strongest at the centre because the magnetic field is strongest at that region. As the magnetic field weakens (when you move away from the centre along the vertical), the separation decreases. Classically, you would have expected to see the deposition everywhere, i.e., you wouldn't see any "empty" spot without deposition in the centre.

What was surprising was that two clear smears were observed. This suggests that spin (and in turn, the associated magnetic dipole moment) can only take two distinct values, causing roughly half of the silver atoms to get deflected one way (say corresponding to spin "up"), and the other half the other way (corresponding to spin "down").

The external magnetic field could have been oriented in any direction. Several interesting observations can be made when the Stern-Gerlach experiment is carried out sequentially.

To aid our discussion, let us say that the silver atoms that get deflected upwards when passing through the magnetic field along $\hat{\mathbf{z}}$ be said to have a spin s_z+ and the silver atoms deflecting downwards are labelled to have s_z- (We refer to this orientation of the external magnetic field as the SG$\hat{\mathbf{z}}$ apparatus from now on).

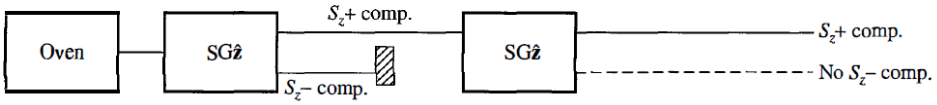

Figure 3: The sequential Stern-Gerlach Experiment, repeating SG$\hat{z}$ on the s_z+ beam. Image taken from Sakurai & Napolitano (2011)

In Figure 3, the upward deflected beam is passed into yet another Stern-Gerlach set-up oriented in the $\hat{\mathbf{z}}$ direction. The result seems intuitive–the s_z- component was filtered out, and is not present after the second set-up as well.[4]

[4]Here you see only one smear, but it is deflected upwards, in the same position where the smear corresponding to the "up" spin was in the non-sequential experiment.

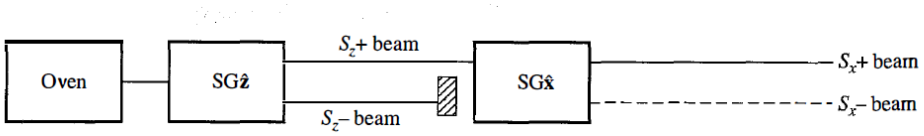

Figure 4: The sequential Stern-Gerlach Experiment, using the SG$\hat{x}$ on the s_z+ beam. Image taken from Sakurai & Napolitano (2011)

In Figure 4, we make a slight modification, where we rotate the second SG set-up such that the magnetic field is oriented in the $\hat{\mathbf{x}}$ direction (We refer to this as the SG$\hat{x}$ apparatus). Interestingly, we see two smears, now deflected horizontally. We might be tempted to say that the s_z+ spin actually consists of s_x+ an s_x- which is separated by SG$\hat{x}$

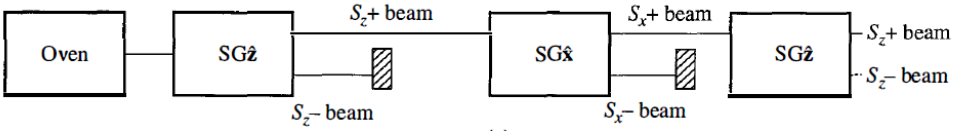

Figure 5: The sequential Stern-Gerlach Experiment, first using SG$\hat{x}$ on the s_z+ beam then SG$\hat{x}$ on the s_x+ beam. Image taken from Sakurai & Napolitano (2011)

Now a counter-intuitive observation occurs once more in Figure 5. Here, after the set-up in (b), the s_x+ beam (say the beam that got reflected to the right) is passed into an SG$\hat{z}$ set-up, and once again you see two smears! This is surprising, because intuitively it would be that since we 'filtered' out all the s_z- right at the start, we wouldn't see any at the end. Our classical intuitions fail us, and we resort to a quantum mechanical description to make sense of the results of this experiment.

1.2.2 Postulates of Quantum Mechanics

Before stating the postulates, we briefly introduce the mathematical language in which they are expressed.

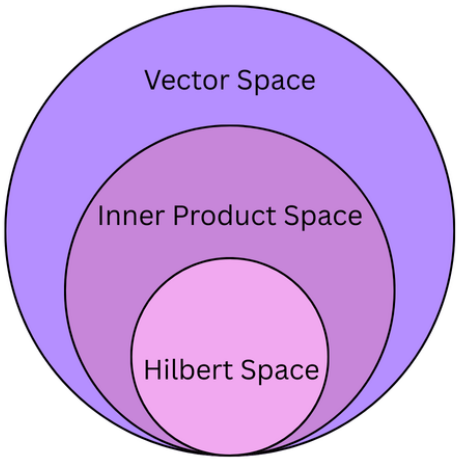

Figure 6: Image depicting the hierarchy of vector space, inner product space and the Hilbert space

Mathematical interlude: vector spaces and Hilbert spaces. A **vector space** $\mathcal{V}$ over a field $\mathbb{F}$ is a set equipped with two operations: vector addition and scalar multiplication, satisfying the standard linearity axioms. In quantum mechanics, vectors represent physical states.

An **inner product space** is a vector space equipped with an inner product

$$\langle \phi | \psi \rangle,$$

[5] which assigns a complex number to each ordered pair of vectors and allows one to define lengths, angles, and orthogonality.

A **Hilbert space** $\mathcal{H}$ is a complete inner product space, over the field of complex numbers $\mathbb{C}$. The inner product structure for the Hilbert space is essential for defining probabilities.

Quantum states are represented by vectors in a Hilbert space $\mathcal{H}$.

Postulate 1: Quantum State
Key concepts: Hilbert space, state vector, wavefunction.

The state of a quantum system is completely specified by a vector

$$|\psi\rangle \in \mathcal{H},$$

called the *quantum state*.[6]

Crucially, the quantum state does not encode a single definite value of position, momentum, or any other observable. Instead, it encodes all possible measurement outcomes and their probabilities for every observable. This is a fundamental departure from classical mechanics, where the state specifies definite values of all dynamical variables simultaneously.

Postulate 2: Observables or Operators
Key concepts: linear operator, Hermitian operator, eigenvalue, eigenstate.

Every physical observable A is represented by a *Hermitian* linear operator $\hat{A}$ acting on the Hilbert space $\mathcal{H}$.

An operator $\hat{A}$ is said to be *Hermitian* if

$$\langle \phi | \hat{A} \psi \rangle = \langle \hat{A} \phi | \psi \rangle \quad \text{for all } |\phi\rangle, |\psi\rangle \in \mathcal{H},$$

or equivalently, if $\hat{A}^\dagger = \hat{A}$, where $\dagger$ denotes the conjugate transpose. Hermiticity guarantees that measurement outcomes are real.

If a state $|a\rangle$ satisfies

$$\hat{A}|a\rangle = a|a\rangle,$$

then a is called an *eigenvalue* of $\hat{A}$ and $|a\rangle$ the corresponding *eigenstate*. Examples of observables include position operator, momentum operator and the Spin operator, which we will use in the coming sections on the SG experiment in terms of the mathematical framework being outlined currently. [7]

Postulate 3: Measurement and the Born Rule ??
Key concepts: measurement, measurement outcome, measurement operator, probability, identity.

Measuring a physical quantity in quantum mechanics corresponds mathematically to measuring an operator $\hat{A}$. The possible outcomes of such a measurement are the eigenvalues a_n of $\hat{A}$.

[5] The $|.\rangle$ symbol (the "ket") is like a column vector and the $\langle .|$ symbol "bra" is like a row vector, provided you are working in a finite dimensional space. What goes inside is for you to keep track of variables, and could be anything, e.g., $|a\rangle, |1\rangle, |\alpha\rangle$, etc. Putting them together, as in the the inner product, e.g.,$\langle \phi | \psi \rangle$ gives you a creatively named 'bra-ket' ("bracket" :))

[6] In the position representation, the state vector is represented by a complex-valued function

$$\psi(x) = \langle x | \psi \rangle,$$

called the *wavefunction*. We return to the concept of a wavefunction in a bit once we start dealing with infinite-dimensional vector spaces.

[7] In the position representation, eigenstates are represented by *eigenfunctions*.
Examples of observables include:

$$x \to \hat{x}, \qquad p \to \hat{p} = -i\hbar \frac{d}{dx}, \qquad E \to \hat{H},$$

where $\hat{H}$ is the Hamiltonian operator.

Since physical observables are represented by Hermitian operators, their eigenstates form an orthonormal set and can be chosen to span the state space of the system.[8]

The probability of obtaining the outcome a_n is given by the *Born rule*:

$$P(a_n) = |\langle a_n|\psi\rangle|^2.$$

In terms of projection operators $\hat{P}_n = |a_n\rangle\langle a_n|$, the same probability can be written as

$$P(a_n) = \langle\psi|\hat{P}_n|\psi\rangle\,,$$

This probabilistic rule is intrinsic to quantum mechanics and does not arise from experimental imperfections. The best example is the Stern-Gerlach experiment, which demonstrates the quantization of angular momentum.

We would also note that the probability of the state being itself is 1, so $\langle\phi|\phi\rangle = 1$ for any valid state vector. This is the **normalization** condition on the state vector. Essentially all states in QM are normalized, or all vectors in this vector space have norm 1.

We can make the notion of a quantum measurement mathematically precise using measurement operators. A quantum measurement is not just "reading off a value"; it is a physical interaction that both produces an outcome and disturbs the state. As mentioned earlier, the number of measurement outcomes is equal to the number of eigenstates of the operator. Each possible outcome m is associated with a measurement operator $\{M_m\}$, which acts on the state space of the system. If the system is initially in the state $|\psi\rangle$, the probability that outcome m occurs $p(m)$:

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$$

This form ensures that probabilities are non-negative and in accordance with the Born rule.

Since the m outcomes are all the possible outcomes exhaustively, $\sum_m p(m) = 1$, which implies:

$$\begin{aligned}
&\sum_m p(m) = 1 \\
\implies &\sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle = 1 \\
\implies &\langle\psi|\left(\sum_m M_m^\dagger M_m\right)|\psi\rangle = 1 \\
\implies &\sum_m M_m^\dagger M_m = \mathbb{1}
\end{aligned}$$

where $\mathbb{1}$ is the identity over the state space of the system of interest. For a more detailed discussion on measurements, the reader is referred to Sec. 2.2.3 in Nielsen & Chuang (2011)

[8] For orientation, the diagram summarizes the hierarchy and mutual relationships between several common classes of linear operators on a Hilbert space, including Hermitian, unitary, positive operators, and projectors.

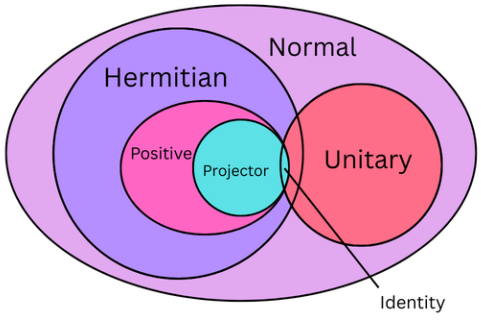

Figure: Operators on a Hilbert space

The Uncertainty Principle

The Uncertainty Principle follows directly from the above mathematical framework.[9] Suppose A and B are two Hermitian operators and $|\psi\rangle$ is a quantum state. Suppose $\langle\psi|AB|\psi\rangle = x + iy$ where x and y are real.

Taking the Hermitian conjugate:

$$\begin{aligned}\langle\psi|AB|\psi\rangle^\dagger &= (|\psi\rangle^\dagger)(B^\dagger A^\dagger)(\langle\psi|^\dagger)\\ &= \langle\psi|BA|\psi\rangle\\ &\quad\text{(Since A and B are Hermitian, } A^\dagger = A \text{ and } B^\dagger = B)\\ &= x - iy\end{aligned}$$

i.e.,

$$\langle\psi|BA|\psi\rangle = x - iy$$

Note that:[10]

$$\langle\psi|[A,B]|\psi\rangle = 2iy, \tag{1}$$

$$\langle\psi|\{A,B\}|\psi\rangle = 2x \tag{2}$$

$$\implies |\langle\psi|[A,B]|\psi\rangle|^2 + |\langle\psi|\{A,B\}|\psi\rangle|^2 = 4|\langle\psi|AB|\psi\rangle|^2. \tag{3}$$

By the Cauchy–Schwarz inequality

$$|\langle\psi|AB|\psi\rangle|^2 \le \langle\psi|A^2|\psi\rangle\,\langle\psi|B^2|\psi\rangle\,, \tag{4}$$

which combined with Equation (3) and dropping a non-negative term gives

$$|\langle\psi|[A,B]|\psi\rangle|^2 \le 4\,\langle\psi|A^2|\psi\rangle\,\langle\psi|B^2|\psi\rangle\,. \tag{5}$$

Suppose C and D are two observables. Substituting $A = C - \langle C\rangle$ and $B = D - \langle D\rangle$ into the last equation, we obtain Heisenberg's uncertainty principle as it is usually stated:

$$\Delta(C)\Delta(D) \ge \frac{|\langle\psi|[C,D]|\psi\rangle|}{2}. \tag{6}$$

Interpreting the statement of the Uncertainty Principle. What equation (6) says is that if you prepare a large number of quantum systems in exactly the same state, and perform a measurement corresponding C on some of them and D on the others, such that their standard deviations are $\Delta(C)$ and $\Delta(D)$, then the product of these follow the inequality in equation (6). *The uncertainty is of mathematical origin, not of experimental negligence.*

Postulate 4: Post-Measurement State

Key concepts: projective measurements, measurement operators, state update, normalization, measurement-induced disturbance.

Postulate 3 specifies the probabilities of different measurement outcomes. We now complete the description of a measurement by stating how the quantum state changes *conditioned on the outcome obtained*.

In the special case of a *projective measurement*[11], immediately after the measurement of an observable $\hat{A}$ yields the outcome a_n, the state of the system collapses to the corresponding normalized eigenstate $|a_n\rangle$.

[9] We provide the treatment provided in Nielsen & Chuang (2011), pg. 89

[10] $[A,B] = AB - BA$, $\{A,B\} = AB + BA$

[11] A projective measurement is one in which the measurement operators are orthogonal projection operators onto the eigenstates of the observable, i.e., $M_n = |a_n\rangle\langle a_n|$.

More generally, in the measurement-operator formalism introduced in Postulate 3, each possible outcome m is associated with a measurement operator M_m. If the system is initially in the state $|\psi\rangle$ and outcome m is obtained, the post-measurement state is

$$|\psi_m\rangle = \frac{M_m |\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}.$$

The denominator ensures normalization, while the action of M_m reflects the physical disturbance induced by the measurement.

Projective measurements correspond to the special case where the measurement operators are orthogonal projectors onto the eigenstates of the observable. It is important to note that the Born rule is a special case of the more general measurement-operator formalism introduced above. In a projective measurement, each outcome corresponds to an orthogonal projector $P_n = |a_n\rangle\langle a_n|$ onto an eigenstate of the observable $\hat{A}$. The probability of obtaining outcome a_n is then

$$p(n) = \langle\psi|P_n|\psi\rangle = |\langle a_n|\psi\rangle|^2,$$

which is precisely the Born rule.

More generally, the measurement-operator formalism allows for non-projective measurements, where the operators M_m need not be orthogonal projectors and may even be non-Hermitian, capturing a wider class of measurement processes that still satisfy $\sum_m M_m^\dagger M_m = \mathbb{1}$. Thus the Born rule emerges naturally when the measurement operators are projectors onto the eigenstates of the measured observable.

There are two more postulates, one of them regarding how the state vector evolves with time, and the other regarding composite systems. These two will be introduced in the next chapter. We now move on to use an example to illustrate the postulates introduced so far.

1.2.3 Describing the Stern-Gerlach Experiment with Quantum Mechanics

Identifying the state space and state vector

Remember that our classical intuitions failed us when we applied them to the Stern-Gerlach experiment. Let us see how the postulates stated above hold up.

Let us consider the state of one particular silver atom[12] to be $|\psi\rangle$ (Quantum State Postulate). This state vector has all the information of everything we could possibly know about that atom.

In the experiments outlined, we see two kinds of observables–splitting to two spots along the z-direction, and splitting to two spots along the x-direction. Let these observables be denoted by S_z and S_x respectively. (Observables are associated with Hermitian Operators, Postulate 2.)

Since we see two distinct measurement outcomes for either orientation, we realize that the state ket $|\psi\rangle$ lives in a vector space with at least 2 dimensions, with one basis vector corresponding to s_z+ and one corresponding to s_z-[13]. Let these be $|+\rangle$ and $|-\rangle$ respectively. Explicitly, if you choose to work with $|+\rangle$ and $|-\rangle$ as basis vectors, you could set:

$$|+\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$|-\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Note that $|+\rangle$ and $|-\rangle$ would be the eigenvectors of the S_z operator. We will use $|+\rangle_x$ and $|-\rangle_x$ as basis vectors corresponding to s_x+ and s_x- respectively (which would be the eigenvectors of S_x).

[12] Please note that we are not dealing with the entire beam right now, since that would need some more mathematical machinery, which will be introduced at a later stage.

[13] The eigenvectors of a Hermitian operator form an orthogonal basis set; measurement collapses the state ket strictly to one of the eigenvectors of the associated operator (Measurement Postulate (3) and Post-Measurement Postulate (4))

Unfortunately, there is no way to know whether the Hilbert space is larger, or if the eigenspaces associated with each eigenvalue (here being "up" spin or "down" spin) have dimensions >1. We assume the non-degenerate case to begin with, and will return to handling degeneracy later in the text when needed.

Defining measurement operators for the problem
Now let us think about how we could define the projective measurements that correspond to the Stern-Gerlach experiment. We could think of the SG$\hat{z}$ apparatus as some kind of filter that separates s_z+ from s_z-, hence we define the projector operators:

$$\begin{aligned} P_{z+} &= |+\rangle \langle +| \\ P_{z-} &= |-\rangle \langle -| \\ P_{x+} &= |+\rangle_x \langle +|_x \\ P_{x-} &= |-\rangle_x \langle -|_x \end{aligned}$$

Verifying projectors introduced above
For a sanity check, we know that in this experiment, you can either be deflected one way or the other and not anything else. That means there are only 2 possible outcomes, so the probability of either of them occurring is 1, requiring that $\sum_m M_m^\dagger M_m = \mathbb{1}$.

$$\begin{aligned} P_{z+}^\dagger P_{z+} + P_{z-}^\dagger P_{z-} &= P_{z+} + P_{z-} \\ &= |+\rangle \langle +| + |-\rangle \langle -| \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \mathbb{1} \end{aligned}$$

Finding the state after the first measurement
Let us look at what happens when one silver atom from the oven is passed through the first SG$\hat{z}$ apparatus, and suppose that atom is observed to be deflected upward. This operation constitutes a measurement of the spin of the atom, which corresponds to applying the measurement operator P_{z+}, and the post measurement state is the eigenvector of the operator S_z corresponding to the observable "spin-up", which would be $|+\rangle$.

Concretely, if the original state of a particular silver atom before measurement is $|\psi\rangle$ and the state after measurement is $|\phi\rangle$, when we apply the projection operator to 'measure':

$$\begin{aligned} |\phi\rangle &= \frac{P_{z+}|\psi\rangle}{\sqrt{\langle\psi|P_{z+}^\dagger P_{z+}|\psi\rangle}} \\ &= \frac{|+\rangle\langle+|\psi\rangle}{\sqrt{\langle\psi|\left(|+\rangle\langle+|+\rangle\langle+|\right)|\psi\rangle}} \\ &= \frac{|+\rangle\langle+|\psi\rangle}{\sqrt{\langle\psi|\left(|+\rangle\langle+|\right)|\psi\rangle}} \\ &= |+\rangle\left(\frac{\langle+|\psi\rangle}{\sqrt{|\langle+|\psi\rangle|^2}}\right) \\ &= e^{i\theta}|+\rangle \end{aligned}$$

Comments about the Global Phase
The global phase is harmless. In the eigenvector equation, you have the freedom to choose the global phase since, if $A|a_n\rangle = a_n|a_n\rangle$, then $Ae^{i\theta}|a_n\rangle = a_n e^{i\theta}|a_n\rangle$ is also true. Further, we will see below that the global phase doesn't affect the probability of the observations.

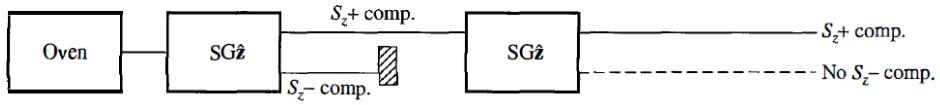

Figure 7: The sequential Stern-Gerlach Experiment, repeating SG$\hat{z}$ on the s_z+ beam. The s_z+ component would have state vector as $e^{i\theta}\,|+\rangle$ as derived above.

Verifying with experimental observations

Now we bring our attention to the first of the sequential SG experiments. We see that $|\phi\rangle$ is measured once again. Calculating the probability of being deflected upwards or downwards is given by the Born Rule (Postulate 3):

$$\begin{aligned} p(s_z+) &= \langle\phi|P_{z+}^{\dagger}P_{z+}|\phi\rangle \\ &= (e^{-i\theta}\,\langle+|)P_{z+}^{\dagger}P_{z+}(e^{i\theta}\,|\phi\rangle) \\ &= (e^{-i\theta}\,\langle+|)\,|+\rangle\,\langle+|+\rangle\,\langle+|\,(e^{i\theta}\,|\phi\rangle) \\ &= 1 \end{aligned}$$

and,

$$\begin{aligned} p(s_z-) &= \langle\phi|P_{z-}^{\dagger}P_{z-}|\phi\rangle \\ &= (e^{-i\theta}\,\langle+|)P_{z-}^{\dagger}P_{z-}(e^{i\theta}\,|\phi\rangle) \\ &= (e^{-i\theta}\,\langle+|)\,|-\rangle\,\langle-|-\rangle\,\langle-|\,(e^{i\theta}\,|\phi\rangle) \\ &= e^{-i\theta}\,\langle+|-|\,\langle-|+\rangle\,e^{i\theta} \\ &= 0 \end{aligned}$$

This perfectly matches with the experimental observation of seeing only the s_z+ deposition on the glass plate! Another point to note is that the "global" phase factor ($e^{i\theta}$) does not affect the probability, which is what we can physically monitor (by creating a multitude of identical systems and repeating the experiment and logging the results for each one individually).

Introducing a different basis

Let us now look at the second sequential SG experiment, where we first measure using the SG$\hat{z}$ apparatus, and pass the s_z+ beam to the SG$\hat{x}$ apparatus.

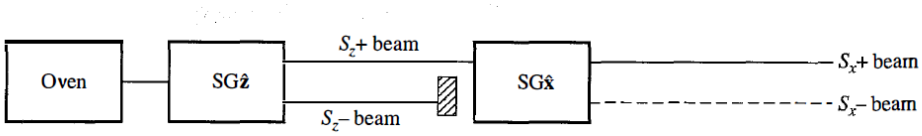

Figure 8: The sequential Stern-Gerlach Experiment, using the SG$\hat{x}$ on the s_z+ beam. As discussed in the previous section, we consider a particular silver atom in s_z+ beam, which would be in state $|\phi\rangle$

Like in the previous case, the probability of a deflection along the x-direction is given by:

$$\begin{aligned} p(s_x+) &= \langle\phi|P_{x+}^\dagger P_{x+}|\phi\rangle \\ &= (e^{-i\theta}\langle+|)(|+\rangle_x\langle+|_x|+\rangle_x\langle+|_x)(e^{i\theta}|+\rangle) \\ &= \langle+|+\rangle_x\langle+|_x|x\rangle \\ &= |\langle+|+\rangle_x|^2 \qquad \text{(eq. 1)} \\ p(s_x-) &= \langle\phi|P_{x-}^\dagger P_{x-}|\phi\rangle \\ &= (e^{-i\theta}\langle+|)(|-\rangle_x\langle-|_-|+\rangle_x\langle-|_x)(e^{i\theta}|+\rangle) \\ &= \langle+|-\rangle_x\langle-|_x|x\rangle \\ &= |\langle+|-\rangle_x|^2 \qquad \text{(eq. 2)} \end{aligned}$$

From the experiment, we observe that $p(s_x+) = \frac{1}{2}$ and $p(s_x-) = \frac{1}{2}$. This implies that the transition coefficients have magnitude $1/\sqrt{2}$.

Phase Convention
At this stage, the phases of the basis vectors $|+\rangle_x$ and $|-\rangle_x$ are not yet fixed. To align our mathematical description with standard conventions, we *choose* the phase of the $|+\rangle_x$ basis vector such that its overlap with $|+\rangle$ is real and positive. Thus:

$$\langle+|+\rangle_x = \frac{1}{\sqrt{2}}$$

Using the Resolution of Identity $\mathbb{1} = |+\rangle_x\langle+|_x + |-\rangle_x\langle-|_x$, we expand the state $|+\rangle$:

$$|+\rangle = \frac{1}{\sqrt{2}}|+\rangle_x + \frac{1}{\sqrt{2}}e^{i\delta}|-\rangle_x$$

(where δ is some relative phase).

Now, we use **orthogonality** to determine the structure of the second state, $|-\rangle$. Since $|+\rangle$ and $|-\rangle$ are orthogonal eigenstates of the Hermitian operator S_z:

$$\langle-|+\rangle = 0$$

Expanding $|-\rangle$ in the x-basis as $|-\rangle = a|+\rangle_x + b|-\rangle_x$, and substituting into the inner product, the orthogonality condition forces the coefficients a and b to have opposite signs relative to the expansion of $|+\rangle$.

To maintain the symmetry and normalization found in the experiment, the unique solution (up to an overall global phase) that satisfies orthogonality is:

$$|-\rangle = \frac{1}{\sqrt{2}}|+\rangle_x - \frac{1}{\sqrt{2}}|-\rangle_x$$

The minus sign in the definition of $|-\rangle$ is not an arbitrary choice; it is the mathematical consequence of enforcing that $|-\rangle$ must be orthogonal to $|+\rangle$ while sharing the same experimental transition probabilities.

Finally, we will invert the relation to express $|+\rangle_x$ and $|-\rangle_x$ in terms of $|+\rangle$ and $|-\rangle$. This is straight forward, and can be done by adding and subtracting the expansions derived above to give:

$$\begin{aligned} |+\rangle_x &= \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T \end{aligned}$$

$$\begin{aligned} |-\rangle_x &= \frac{1}{\sqrt{2}}|+\rangle - \frac{1}{\sqrt{2}}|-\rangle \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}^T \end{aligned}$$

Thus we see that $|+\rangle_x$ and $|-\rangle_x$ are another basis set in the same state space!

Exercises for the Reader

- Find the elements of the S_z and S_x operators in the $|+\rangle, |-\rangle$ basis. Now find them in the $|+\rangle_x, |-\rangle_x$ basis.
- Verify the result of the third sequential Stern-Gerlach set-up using the formalism introduced.

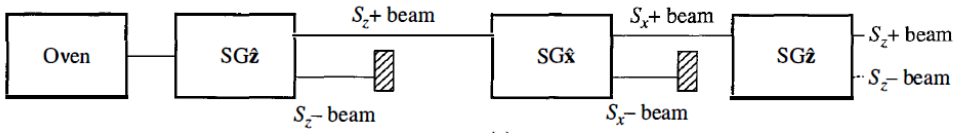

Figure 9: The sequential Stern-Gerlach Experiment, first using SG$\hat{x}$ on the s_z+ beam then SG$\hat{x}$ on the s_x+ beam. Image taken from Sakurai

References

Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 10th anniversary edition, 2011. ISBN 978-1-107-00217-3.

J. J. Sakurai and Jim Napolitano. *Modern Quantum Mechanics*. Addison-Wesley, San Francisco, CA, 2nd edition, 2011. ISBN 978-0-8053-8291-4.