Physics 438A – Lecture #1
Stern-Gerlach Experiments

Aaron Wirthwein – wirthwei@usc.edu

Last Updated: January 1, 2026

1 Introduction

1.1 Welcome to Quantum Mechanics

This course marks a turning point in your study of physics. Up to now, you have mostly been working within the framework of classical mechanics. Developed over centuries by Newton, Lagrange, Hamilton, and others, classical mechanics is a triumph of human imagination and reasoning. We assume that particles have definite positions and momenta, regardless of how accurately we can measure them, and that the state of a closed system deterministically evolves over time. These are powerful ideas, and they remain useful. In many systems, *classical mechanics is still the best tool we have.* But it's not the final word.

In the early 20th century, physicists encountered phenomena that classical mechanics couldn't explain, as you may be well-aware. Blackbody radiation, the photoelectric effect, atomic spectra, etc. all pointed to a deeper structure underlying matter and energy. Quantum mechanics was born from the minds of people working hard to solve these puzzles, and what emerged was not just a set of ad hoc rules, but a radically different theoretical framework, grounded in a strange but consistent mathematical formalism. Many classical concepts like energy, momentum, and conservation laws remain central to quantum mechanics, but they are *redefined and reinterpreted*. Classical mechanics becomes an approximation–a limiting case of quantum mechanics–and this continuity is crucial. Quantum mechanics reveals where and how the assumptions of classical mechanics break down and what stands in their place.

1.2 Disclaimer

These notes are intended for pedagogical purposes and are designed to guide students through the core concepts of quantum mechanics in a way that emphasizes understanding and intuition. They make assumptions about what students already know in order to focus attention on the new ideas being introduced and make the transition to these ideas as comfortable as possible (a fool's errand). The presentation favors conceptual development over formal rigor, and as such, the notes are not meant to serve as a mathematical treatise. Rather than functioning as a comprehensive reference, they are structured as an interactive guide—something

to be worked through actively, not merely consulted. Students seeking more formal or detailed treatments are encouraged to supplement these notes with standard textbooks and references (see the recommendations on Brightspace).

We will introduce quantum mechanics using the *Stern-Gerlach experiment* as a concrete entry point. It's important to understand, however, that quantum mechanics is not merely induced from experimental observations, nor is it derived from deeper theoretical principles. It is built on a set of postulates, which are statements informed by experiments but not proven by them. Postulates are creative, conceptual starting points that organize and give meaning to observations and allow us to explain and predict a wide range of phenomena. If the postulates of quantum theory follow from a deeper framework, that framework is not currently known or widely accepted.

2 Stern-Gerlach Experiment

The Stern-Gerlach experiment, published in 1922 by Otto Stern and Walther Gerlach, demonstrated that particles possess quantized intrinsic angular momentum, or spin [1]. The experiment consisted of an oven that produced a beam of neutral atoms, a region of space with an inhomogeneous magnetic field, and a photographic plate to serve as a detector. When a beam of silver atoms passed through the magnetic field, it split into two distinct beams; one deflected upwards and the other deflected downwards, relative to the direction of the magnetic field gradient.

Let's analyze the experiment classically to understand why this was a surprising result. A neutral particle will interact with a magnetic field if it possesses a magnetic dipole moment[1]. The potential energy of this interaction is $U = -\boldsymbol{\mu} \cdot \mathbf{B}$ which results in a force given by $\mathbf{F} = \nabla(\boldsymbol{\mu} \cdot \mathbf{B})$. The magnetic field must have a nonzero gradient in some direction to exert a force in that direction, and this was accomplished with an asymmetrical arrangement of permanent magnets–one sharp-pointed pole (like a wedge or cone) and one broad, flat pole. Let's assume the magnetic field is

$$\mathbf{B} = -\alpha x \hat{\mathbf{x}} + (B_0 + \alpha z)\hat{\mathbf{z}},$$

where $\alpha = \partial B_z/\partial z$ and B_0 are constants. The force on the magnetic dipole is

$$\mathbf{F} = -\alpha \mu_x \hat{\mathbf{x}} + \alpha \mu_z \hat{\mathbf{z}}$$

[1] See Griffiths, *Introduction to Electrodynamics*

According to Maxwell's equations, **B** must be divergence-less, so we end up with a force having both x and z components. It turns out that we can ignore the x component as long as the constant B_0 is very large–so large that μ_x will oscillate due to Larmor precession and average to zero (we'll investigate this phenomenon later in the context of quantum mechanics, but it's a well-established result of classical electromagnetism).

A beam of silver atoms, initially traveling in the y direction, feels a net force in the z direction given by

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

Every atom consists of charged particles which can produce loops of current that give rise to a magnetic moment. A loop of area A and current I produces a magnetic moment $\mu = IA$. For instance, if this loop of current arises from a particle with charge q and mass m traveling in a circle with radius r at speed v, then

$$\mu = \frac{q}{T}\pi r^2 = \frac{q}{2\pi r/v}\pi r^2 = \frac{qrv}{2}.$$

On the other hand, the angular momentum of such a particle about the center of the circle is $L = mvr$, so

$$\mu = \frac{q}{2m}L$$

In the same way that the earth revolves around the sun and rotates around its own axis, we can also imagine a charged particle in an atom having orbital angular momentum **L** and a new property, the intrinsic angular momentum, which we label **S** and call spin. The relationship between $\boldsymbol{\mu}$ and **S** is very similar to that between $\boldsymbol{\mu}$ and **L**, except an integration over current loops must be performed. This contributes an overall constant factor g called the gyroscopic ratio that depends on the geometry of the particle. Hence,

$$\boldsymbol{\mu} = g\frac{q}{2m}\mathbf{S}$$

It turns out that the magnetic moment of a neutral silver atom is almost entirely due to the spin of its outermost, unpaired electron. The magnetic moment of this electron, and therefore of the entire neutral silver atom, is

$$\boldsymbol{\mu} = -g\frac{e}{2m_e}\mathbf{S}$$

where e is the magnitude of the electron's charge. The classical force on the atom is

$$F_z = g\frac{e}{2m_e}S_z\frac{\partial B_z}{\partial z}$$

Hence, the deflection of the beam in the Stern-Gerlach experiment is a measure of the z-component of spin, which is the orientation of the magnetic field gradient. In the thermal environment of the oven, we expect a random distribution of spin directions, so the z-component of spin should be a continuous distribution between $-|\mathbf{S}|$ and $+|\mathbf{S}|$, which would yield a continuous spread in deflections of the silver atomic beam.

As stated earlier, Stern and Gerlach observed two deflections, indicating only two values for the z-component of the electron spin. The magnitudes of deflection are consistent with

$$S_z = \pm\frac{\hbar}{2}$$

where $\hbar$ is Planck's constant h divided by 2π and has the numerical value

$$\hbar = 1.054\ 571\ 817 \times 10^{-34}\ \mathrm{J \cdot s} = 6.582\ 119\ 569 \times 10^{-16}\ \mathrm{eV \cdot s}$$

The factor of $1/2$ in the values of S_z is why we say the electron is a spin-1/2 particle. When $S_z = +\hbar/2$, we say the electron is spin up, and when $S_z = -\hbar/2$ the electron is spin down.

3 Preparing a Quantum State

As it turns out, the concept of intrinsic spin can be fully described within the framework of classical field theory [2]. The *quantization of spin* is entirely quantum mechanical and can be understood in the context of quantum field theory [3]. While the magnitude of the electron's spin is always a fixed value, *there is no way to predict which component of spin you will observe for a single electron* in a Stern-Gerlach experiment. Quantization of spin is certainly interesting, but this is our first key insight into quantum mechanics.

Your physics education has mostly dealt with the certainty and exactness of classical mechanics. The predictions of quantum mechanics, on the other hand, are intrinsically probabilistic. In most areas of science, probabilities reflect ignorance about the exact state of a system. For example, in statistical mechanics, we assign probabilities because we cannot realistically keep track of the overwhelming number of the microstates. In quantum mechanics, probabilities are intrinsic to the theory; even with complete knowledge of the state of a system, the outcomes are probabilistic. Philosophers make a distinction between *epistemic* probability and *ontological* probability; the former has to do with one's knowledge whereas the latter has to do with existence.

When a radioactive nucleus decays, it does so spontaneously; there simply is no rule that will tell you when it decays or in what direction the particles will travel afterwards. When

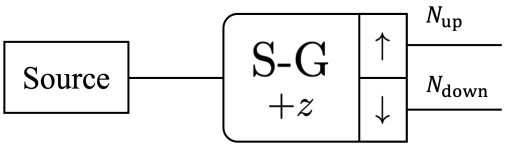

Figure 3.1: A simplified diagram of the Stern-Gerlach experiment. N_{up} is the number of atoms that are spin up, and N_{down} is the number of atoms that are spin down, as recorded by the counter.

a particle of light strikes a half-silvered mirror, it either reflects off it or passes through, and the best we can do is assign probabilities for each occurrence. Truly random numbers have been generated using Geiger counters and quantum optics, and in 2025, a team at NIST built the first random number generator that uses quantum entanglement to produce verifiable random numbers [4]. We'll talk more about entanglement later, but the idea is that quantum systems can exhibit correlations that are much stronger than fluctuating classical systems. We can verify a system is "truly quantum" by measuring these strong correlations.

Figure 3.1 illustrates a simplified version of the Stern-Gerlach experiment consisting of an oven that produces the beam of atoms, a Stern-Gerlach device with two output ports for the possible values of the spin component, and two counters to detect the atoms leaving the output ports. Analyzing the atoms that come directly from the source involves a longer story than we're ready to tell at the moment (involving the density matrix, a topic reserved for Physics 438B). Just know that the ratio of N_{up} to N_{down} is not easily determined because the oven creates a statistical ensemble of what we call "quantum states" in this class.

We can simplify the situation by blocking one of the output ports and keeping track of the number of atoms that exit the other port. This is called **state preparation** because the quantum state of the atoms that exit the free port can be easily determined. The output beams are labeled with symbols called **kets**. We use the ket $|+z\rangle$ to represent the state of atoms that exit with spin up and $|-z\rangle$ to represent the state of atoms that exit with spin down. The ket notation we will use in this course was developed by Paul Dirac. We will discuss the mathematics of kets in full detail later. **Informally, the first postulate of quantum mechanics states that kets describe the quantum state–they contain all the information that we can know about the system.** Kets are elements of a vector space, which is a concept from linear algebra.

4 Ideal S-G Experiments

Now let's consider a series of simple Stern-Gerlach experiments with slight variations that help to illustrate important features of quantum mechanics. The experiments are summarized in Figure 4.1. We first describe the experiments and their results and draw some qualitative and general conclusions about the nature of quantum mechanics. Then we introduce the formal mathematics show how it can be used to describe and predict the results of each experiment.

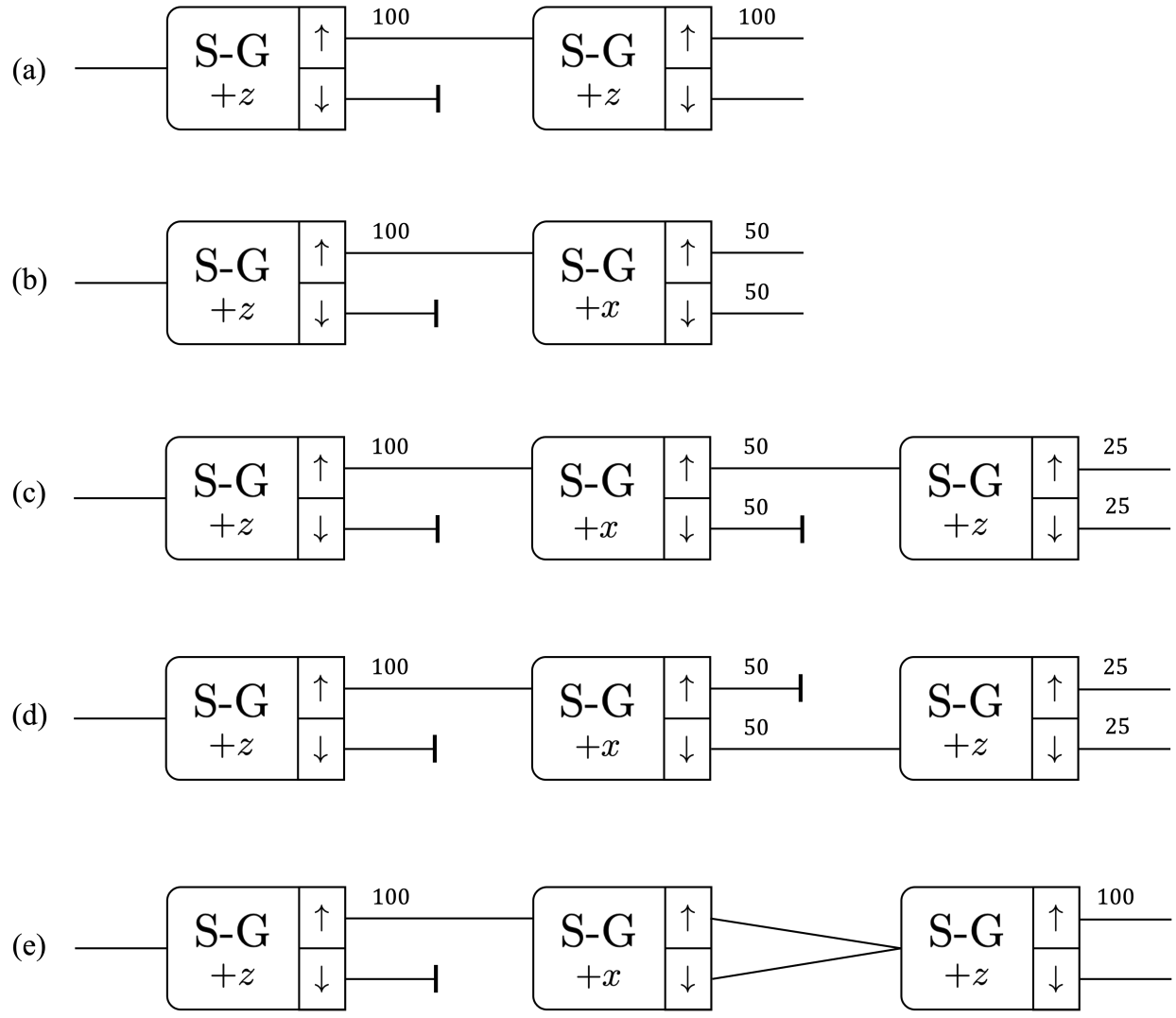

Figure 4.1: Ideal Stern-Gerlach thought experiments.

4.1 Experiment 1

The first experiment is depicted in Figure 4.1(a). Atoms from the source enter a Stern-Gerlach analyzer with a magnetic field gradient in the $+z$ direction. The lower port is blocked and the upper port is free. Atoms that exit the upper port are fed into a second analyzer oriented in the same direction as the first. For every 100 atoms that exit the upper port of the first analyzer, 100 atoms will exit the upper port of the second. This might not be surprising, but it sets the stage for the following experiments, and it illustrates the concept of state preparation. We find a similar result if we block the upper port of the first analyzer and leave the lower port free–all atoms that exit the second analyzer will exit the lower port.

4.2 Experiment 2

Figure 4.1(b) shows the second experiment, which is similar to the first except for the fact that the second analyzer is rotated 90° relative to the first. The magnetic field gradient of the second analyzer is aligned with the $+x$-axis. We find that atoms appear from both output ports of the second analyzer this time, and in equal proportion. For every 100 atoms that exit the upper port of the first analyzer, 50 will exit the upper port of the second and 50 will exit the lower port. We find the same result if we block the upper port of the first analyzer–for every 100 that exit the lower port of the first, 50 exit the upper port of the second and 50 exit the lower port.

The most important feature of this experiment is revealed when we look at the data collected by the counters. Assuming the atoms go through one at a time, there is no way to predict which port of the second analyzer any particular atom will exit. It is only by letting the experiment unfold after many repetitions that a pattern begins to emerge, namely that atoms will exit both ports in roughly equal amounts. In what sense can we claim to know the state of the system if we can't determine exactly what happens to it? One might be tempted to say we simply don't know enough about the system to predict which port an atom will exit. Perhaps one day we'll have a fully deterministic theory, but all signs indicate that nature is fundamentally probabilistic. We'll see that knowing the state of the atom after the first analyzer, which is $|+z\rangle$, we can predict the statistical distribution of outcomes from the second analyzer–that is the sense in which we claim to know the state of the system.

4.3 Experiment 3

The experiments depicted in Figures 4.1(c) and 4.1(d) extend experiment 2 by adding a third analyzer aligned along the z-axis. With the lower ports of the first and second analyzers blocked, for every 100 atoms that exit the upper port of the first, 50 will exit the upper port of the second, 25 will exit the upper of the third, and 25 will exit the lower port of the third. Naively, we might expect all the atoms to exit the upper port of the third analyzer due to the state preparation of the first, but that is not what's observed. The system behaves as if the second analyzer "resets" the quantum state. By analogy with the first analyzer, the state of the atoms that exit the upper port of the second analyzer is $|+x\rangle$, and the state of the atoms that exit the lower port is $|-x\rangle$.

This result demonstrates another key feature of quantum mechanics: a measurement disturbs the system. The second analyzer has disturbed the system in such a way to make the spin component along the z-axis unknown, even though we measured it with the first analyzer. There is a fundamental incompatibility in trying to measure the spin component of the atom along two different directions. We say that S_x and S_z are "incompatible observables."

4.4 Experiment 4

Given the results of experiment 3, it would be interesting to study what happens if the two beams exiting the second analyzer are combined and then fed into the third analyzer, as illustrated in Figure 4.1(e). The results depend on exactly how that's done. If we keep the counters in place and record which atoms exit which port of the second analyzer, then we get a 50/50 split just like experiment 3 (except this time 50 atoms exit the upper port of the third analyzer and 50 exit the lower port). If instead we remove the counters and simply combine the beams with a suitable arrangement of magnets[1], all the atoms exit the upper port of the third analyzer and *none will exit the lower port.* The atoms appear to "remember" that they were initially measured to have spin up along the z-axis. They behave as if the second analyzer isn't even there.

This result demonstrates *quantum mechanical interference*–the chance that an atom will exit the lower port of the third analyzer is zero despite giving the atom more routes through the experiment. As long as there is no record of which path the atoms take through the second analyzer, the two paths seemingly interfere with one another in a way that is reminiscent of wave interference. The upper port experiences constructive interference while the lower port shows destructive interference. It's important to understand that there is no literal wave interference necessarily happening, and this is entirely due to way states and probabilities

[1]See Feynman's *Lectures on Physics*, Volume 3.

work in quantum mechanics. Early attempts to understand quantum mechanical interference fell for this trap, and the terminology developed during this period persists even today.

The existence of quantum mechanical interference makes it difficult to say that kets are simply about our knowledge of a physical system (meaning they are epistemic quantities). Kets encode possibilities, and different possible outcomes can apparently "interfere with one another" like classical waves as long as nothing disturbs them. Any attempt to determine which outcome will occur spoils the interference. Surprisingly, the postulates of quantum mechanics explain interference as a generic feature of physical systems. Interference effects have been observed for photons, electrons, atoms, and even molecules as large as carbon-60 and carbon-70. As systems grow larger, so does the chance that something will spoil interference, and this partially explains why you don't diffract through a doorway...

5 Quantum State Vectors

The kets we introduced earlier are abstract objects belonging to a vector space and they obey many of the familiar rules you know about ordinary geometric vectors. Kets are related to **quantum state vectors** and they are elements of a complex vector space with important properties (we'll talk about the properties as they become relevant). The dimension of the vector space depends on the mathematical model of the physical system and, most importantly, on the number of measurement outcomes. In the Stern-Gerlach experiments we discussed, there are only two possible results for a spin component measurement, so the vector space is two dimensional.

In the two-dimensional vector space of a spin-1/2 system, the two kets $|\pm z\rangle$ form a basis, just like the Cartesian unit vectors $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$ form a basis for vectors in three-dimensional Euclidean space. In particular, the Cartesian unit vectors form an orthonormal basis with the follow properties:

$$\begin{aligned}
&\text{normalization:} && \hat{\mathbf{x}} \cdot \hat{\mathbf{x}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} = 1 \\
&\text{orthogonality:} && \hat{\mathbf{x}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{x}} \cdot \hat{\mathbf{z}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{z}} = 0 \\
&\text{completeness:} && \mathbf{A} = a_x \hat{\mathbf{x}} + a_y \hat{\mathbf{y}} + a_z \hat{\mathbf{z}}
\end{aligned}$$

Notice that the scalar product (dot product) is central to the description of these properties. In light of normalization and orthogonality, we can express completeness as

$$\mathbf{A} = (\hat{\mathbf{x}} \cdot \mathbf{A})\hat{\mathbf{x}} + (\hat{\mathbf{y}} \cdot \mathbf{A})\hat{\mathbf{y}} + (\hat{\mathbf{z}} \cdot \mathbf{A})\hat{\mathbf{z}}$$

since each component can be found by simply taking the scalar product of $\mathbf{A}$ with the corresponding unit vector, i.e. $a_x = (\hat{\mathbf{x}} \cdot \mathbf{A})$, etc.

By analogy with spatial vectors, the kets $|\pm z\rangle$ form a complete orthonormal basis. Completeness implies that a general vector $|\psi\rangle$ is a linear combination of the two basis kets:

$$|\psi\rangle = \alpha|+z\rangle + \beta|-z\rangle \tag{5.1}$$

where α and β are complex scalar numbers multiplying each basis ket. The sense in which $|\pm z\rangle$ are normalized and orthogonal depends on the definition of the inner product, which is a generalization of the familiar scalar product that applies to geometric vectors.

An **inner product** is a map that takes any two elements from the vector space and returns a complex number. For kets $|\psi\rangle$ and $|\phi\rangle$, the inner product is written as $\langle\psi|\phi\rangle$ and it has the following properties:

1. **Linearity** (in the second argument):

$$\langle u|\Big(a|v\rangle + b|w\rangle\Big) = a\langle u|v\rangle + b\langle u|w\rangle \tag{5.2}$$

for complex numbers a, b and kets $|u\rangle, |v\rangle, |w\rangle$.

2. **Conjugate symmetry**:

$$\langle v|w\rangle = \langle w|v\rangle^* \tag{5.3}$$

where * denotes the complex conjugate. Since $a = a^*$ if and only if a is a real number, $\langle\psi|\psi\rangle$ is always a real number.

3. **Positive semi-definiteness**:

$$\langle\psi|\psi\rangle \geq 0 \tag{5.4}$$

and $\langle\psi|\psi\rangle = 0$ if and only if $|\psi\rangle = 0$.

5.1 Bra Vectors and the Induced Norm

The inner product allows us to define a new object called a **bra vector**, which is an element of a dual vector space. Each ket $|\psi\rangle$ can be identified with a bra vector $\langle\psi|$, which is a linear functional that takes a single ket and returns a complex number. Anytime you see the notation $\langle\psi|$, you should see it as a linear map awaiting a ket to produce a complex number in a manner that is consistent with the inner product defined above. When a bra and ket are

combined to form an inner product, we get a bracket or bra-ket (cute huh?). More precisely, for every ket $|\psi\rangle$ in the vector space, there exists a unique bra $\langle\psi|$ such that

$$\langle\psi|\big(|\phi\rangle\big) = \langle\psi|\phi\rangle \qquad \text{(bra vector induced by inner product)}$$

for all kets $|\phi\rangle$. The inner product also allows us to define the **norm** of a ket, which is a generalization of length. The norm of $|\psi\rangle$ is defined as

$$\|\psi\| = \sqrt{\langle\psi|\psi\rangle} \qquad \text{(norm induced by inner product)}$$

All quantum state vectors are normalized kets in the sense that $\|\psi\| = \sqrt{\langle\psi|\psi\rangle} = 1$. This is crucial for the probabilistic interpretation that we'll explain soon.

Given the state $|\psi\rangle$ in Eq. (5.1) we can construct the corresponding bra vector $\langle\psi|$ using the properties of the inner product. Let $|\phi\rangle$ be an arbitrary ket. Then, $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$ by conjugate symmetry, and by linearity,

$$\langle\phi|\Big(\alpha|{+z}\rangle + \beta|{-z}\rangle\Big) = \alpha\langle\phi|{+z}\rangle + \beta\langle\phi|{-z}\rangle.$$

Complex conjugation is itself a linear operation on complex numbers, so the complex conjugate of the preceding expression is

$$\alpha^*\langle\phi|{+z}\rangle^* + \beta^*\langle\phi|{-z}\rangle^* = \alpha^*\langle{+z}|\phi\rangle + \beta^*\langle{-z}|\phi\rangle$$

Hence, $\langle\psi| = \alpha^*\langle{+z}| + \beta^*\langle{-z}|$ for all kets in the vector space, since $|\phi\rangle$ was arbitrary.

5.2 Probabilistic Interpretation

The properties of normalization, orthogonality, and completeness can be expressed in the case of a two-state spin-1/2 quantum system as:

$$\begin{aligned}
&\text{normalization:} && \langle{+z}|{+z}\rangle = \langle{-z}|{-z}\rangle = 1 \\
&\text{orthogonality:} && \langle{+z}|{-z}\rangle = \langle{-z}|{+z}\rangle = 0 \\
&\text{completeness:} && |\psi\rangle = \alpha|{+z}\rangle + \beta|{-z}\rangle
\end{aligned}$$

The inner product can be used to find the coefficients of a ket in the $|\pm z\rangle$ basis. For instance,

$$\langle{+z}|\psi\rangle = \alpha\langle{+z}|{+z}\rangle + \beta\langle{+z}|{-z}\rangle = \alpha$$

$$\langle -z|\psi\rangle = \alpha\langle -z|+z\rangle + \beta\langle -z|-z\rangle = \beta$$

If we apply normalization to $|\psi\rangle$, we obtain a constraint on the coefficients:

$$\langle\psi|\psi\rangle = \Big(\alpha^*\langle +z| + \beta^*\langle -z|\Big)\Big(\alpha|+z\rangle + \beta|-z\rangle\Big) = 1$$

$$\alpha^*\alpha\langle +z|+z\rangle + \beta^*\beta\langle -z|-z\rangle = 1$$

$$|\alpha|^2 + |\beta|^2 = 1$$

or in terms of the expressions above,

$$|\langle +z|\psi\rangle|^2 + |\langle -z|\psi\rangle|^2 = 1 \tag{5.5}$$

In other words, the sum of the absolute squares is equal to 1. Since each term in the sum is greater than or equal to zero, the absolute squares can be used to define a discrete probability distribution. *Now comes a crucial element of quantum mechanics.* **We postulate that each term in the sum is equal to the probability that the quantum state described by the normalized ket $|\psi\rangle$ is measured to be in the corresponding basis state**. The quantity

$$p(+z) = |\langle +z|\psi\rangle|^2 \tag{5.6}$$

is the probability that atoms in the state $|\psi\rangle$ are found in the state $|+z\rangle$ when a measurement of S_z is made, meaning $S_z = +\hbar/2$, and

$$p(-z) = |\langle -z|\psi\rangle|^2 \tag{5.7}$$

is the probability that atoms in the state $|\psi\rangle$ are found in the state $|-z\rangle$ when a measurement of S_z is made, meaning $S_z = -\hbar/2$ for those atoms.

Because the quantum mechanical probability is found by squaring an inner product, we refer to the inner product $\langle +z|\psi\rangle$, for instance, as a **probability amplitude** or just an amplitude–much like a classical wave intensity is found by squaring the wave amplitude. The probabilistic interpretation gives us physical insight into the inner product. For instance, the inner product $\langle +z|+z\rangle = 1$ means that atoms in the state $|+z\rangle$ have a 100% chance of being measured in the $|+z\rangle$ state, and $\langle -z|+z\rangle = 0$ means that atoms in the state $|+z\rangle$ have a 0% chance of being measured in the $|-z\rangle$ state. The possible outcomes are *mutually exclusive*. By convention, the input or initial state appears on the right and the output or final state goes on the left, although $|\langle \text{out}|\text{in}\rangle|^2 = |\langle \text{in}|\text{out}\rangle|^2$, so it doesn't really matter.

Example 5.1: Measurement of S_z

A measurement of S_z is carried out on a particle in the state

$$|\psi\rangle = \frac{1}{2}|+z\rangle + \frac{i\sqrt{3}}{2}|-z\rangle$$

What are the possible results of this measurement and with what probability do these results occur?

Solution: Start by computing the inner products with $|\pm z\rangle$.

$$\langle +z|\psi\rangle = \frac{1}{2} \quad \text{and} \quad \langle -z|\psi\rangle = \frac{i\sqrt{3}}{2}$$

The probability to obtain $S_z = \hbar/2$ is

$$\text{prob}(S_z = \hbar/2) = |\langle +z|\psi\rangle|^2 = \frac{1}{4}$$

and the probability to obtain $S_z = -\hbar/2$ is

$$\text{prob}(S_z = -\hbar/2) = |\langle +z|\psi\rangle|^2 = \langle -z|\psi\rangle^* \langle -z|\psi\rangle = \left(\frac{-i\sqrt{3}}{2}\right)\left(\frac{i\sqrt{3}}{2}\right) = \frac{3}{4}$$

Since the state $|\psi\rangle$ is normalized, the probabilities sum to one.

Example 5.2: Normalizing a Ket

Normalize the ket $|\psi\rangle = C(|+z\rangle + 2i|-z\rangle)$ by finding an appropriate value for the constant C, often referred to as a normalization constant.

Solution: Compute the inner product of $|\psi\rangle$ with itself and set it equal to one.

$$\begin{aligned}
1 &= \langle\psi|\psi\rangle \\
&= C^*\Big(\langle +z| - 2i\langle -z|\Big) \cdot C\Big(|+z\rangle + 2i|-z\rangle\Big) \\
&= |C|^2\Big(\underbrace{\langle +z|+z\rangle}_{=1} + 2i\,\underbrace{\langle -z|+z\rangle}_{=0} - 2i\,\underbrace{\langle -z|+z\rangle}_{=0} - 4i^2\,\underbrace{\langle -z|-z\rangle}_{=1}\Big) \\
&= 5|C|^2
\end{aligned}$$

Hence, $|C|^2 = 1/5$, and a properly normalized ket would be $|\psi\rangle = \frac{1}{\sqrt{5}}(|+z\rangle + 2i|-z\rangle)$.

6 Observables and Measurements

The Stern-Gerlach experiment is understood within a quantum mechanical framework where we no longer assume that all physical quantities have definite values at all times. Instead, we focus on how those quantities manifest as a result of measurement. In quantum theory, an **observable** is any measurable physical quantity like position, velocity, or spin, and we use linear operators to describe observables and model the impact of measurement on a system.

A **linear operator** is a map between vectors that preserves addition and scalar multiplication. For example,

$$A\left[c_1|\psi\rangle + c_2|\phi\rangle\right] = c_1 A|\psi\rangle + c_2 A|\phi\rangle \tag{6.1}$$

where A is an operator (usually clear from context, but sometimes not), c_1 and c_2 are scalars (complex numbers), and $|\psi\rangle$ and $|\phi\rangle$ are kets. Additionally, $A|\psi\rangle$ and $A|\phi\rangle$ are both elements of the same vector space as $|\psi\rangle$ and $|\phi\rangle$.

6.1 Matrix Elements

In the same way that vectors can be written as a linear combination of basis vectors, linear operators can also be written as a linear combination of a basis formed by *outer products*.

The **outer product** of vectors $|\phi\rangle$ and $|\psi\rangle$, denoted $|\phi\rangle\langle\psi|$, is a linear operator on vectors that takes any vector $|\varphi\rangle$, "applies" the bra vector $\langle\psi|$ to compute the inner product $\langle\psi|\varphi\rangle$, and outputs the vector $\langle\psi|\varphi\rangle|\phi\rangle$. In other words,

$$|\phi\rangle\langle\psi|\big(|\varphi\rangle\big) = \langle\psi|\varphi\rangle|\phi\rangle = |\phi\rangle\langle\psi|\varphi\rangle \tag{6.2}$$

For example, any linear operator acting on a two-dimensional vector space can be written as

$$A = a|{+z}\rangle\langle{+z}| + b|{-z}\rangle\langle{+z}| + c|{+z}\rangle\langle{-z}| + d|{-z}\rangle\langle{-z}| \tag{6.3}$$

where a, b, c, d are all complex numbers. Since any vector is a linear combination of basis states, an operator is uniquely specified by its action on basis states. In other words, if we know how an operator acts on basis states, we know how it acts on any vector. By linearity and orthonormality,

$$A|{+z}\rangle = a|{+z}\rangle + b|{-z}\rangle$$

$$A|-z\rangle = c|+z\rangle + d|-z\rangle$$

and it follows that

$$\langle +z|A|+z\rangle = a$$

$$\langle -z|A|+z\rangle = b$$

$$\langle +z|A|-z\rangle = c$$

$$\langle -z|A|-z\rangle = d$$

These are called the matrix elements of the operator A, even though we haven't introduce matrix notation at this point. Notice the matrix elements are basis dependent.

6.2 Eigenvalues and Projections

There are special kets that are not changed by the operation of a particular operator, except for a multiplicative constant. Such kets are known as eigenvectors of the operator and the multiplicative constants are the eigenvalues. **We postulate that the only possible result of measuring a physical observable is one of the eigenvalues of the corresponding operator. After the measurement, the state of the system is the eigenstate corresponding to the measured eigenvalue.** The relationship between an observable, a possible measurement result, and the ket corresponding to that result is called an eigenvalue equation. For example, the eigenvalue equations for the S_z operator are

$$\begin{aligned} S_z|+z\rangle &= \tfrac{\hbar}{2}|+z\rangle \\ S_z|-z\rangle &= -\tfrac{\hbar}{2}|+z\rangle \end{aligned} \tag{6.4}$$

and the matrix elements are $\langle +z|S_z|+z\rangle = \hbar/2$, $\langle -z|S_z|-z\rangle = -\hbar/2$, and the rest are zero. Hence,

$$S_z = \tfrac{\hbar}{2}|+z\rangle\langle +z| - \tfrac{\hbar}{2}|-z\rangle\langle -z| \tag{6.5}$$

Observe that S_z is a linear combination of operators that square to themselves; for instance $(|+z\rangle\langle +z|)^2 = |+z\rangle\langle +z|$. This is analogous to finding the z-component of a Euclidean vector. Applying the projection twice gives you the same result as applying it once.

A **projection operator** P is a linear operator such that $P^2 = P$.

In case it isn't clear that $\langle +z| + z\rangle$ is a projection, let $|\phi\rangle$ be an arbitrary vector:

$$\begin{aligned}(|+z\rangle\langle +z|)^2\,|\phi\rangle &= |+z\rangle\langle +z|\Big[|+z\rangle\langle +z|\big(|\phi\rangle\big)\Big] \\ &= |+z\rangle\langle +z|\Big[\langle +z|\phi\rangle|+z\rangle\Big] \\ &= \langle +z|\phi\rangle|+z\rangle \\ &= |+z\rangle\langle +z|\big(|\phi\rangle\big)\end{aligned}$$

Hence, $\big(|+z\rangle\langle +z|\big)^2 = |+z\rangle\langle +z|$ since $|\phi\rangle$ was arbitrary. Notice that you get the same result by naively multiplying the two operators and "pulling apart" the outer product notation:

$$\big(|+z\rangle\langle +z|\big)^2 = |+z\rangle\ \langle +z| + z\rangle\langle +z| = |+z\rangle\langle +z|.$$

Although this "proof" is much shorter, it invokes unjustified assumptions about how the outer product works, and you should be careful when manipulating new mathematical objects. When in doubt, go back to the definitions, but in this case it works just fine.

Informally, a projection operator projects a vector onto a single vector or combination of vectors. If $P_+ = |+z\rangle\langle +z|$ and $P_- = |-z\rangle\langle -z|$, then

$$S_z = \tfrac{\hbar}{2}P_+ - \tfrac{\hbar}{2}P_-$$

and

$$P_+ + P_- = |+z\rangle\langle +z| + |-z\rangle\langle -z| = \mathbb{1} \tag{6.6}$$

where $\mathbb{1}$ is the identity operator. Eq. (6.6) is called a completeness relation, closure, or resolution of the identity. To see why, consider an arbitrary vector $|\psi\rangle$:

$$\begin{aligned}|\psi\rangle &= |+z\rangle\langle +z|\psi\rangle + |-z\rangle\langle -z|\psi\rangle \\ &= \Big(|+z\rangle\langle +z| + |-z\rangle\langle -z|\Big)|\psi\rangle\end{aligned}$$

The quantity in parentheses is a sum of the projection operators $P_+ + P_-$ and clearly the sum is equal to the identity by comparing the left- and right-hand sides of the last equation. Notice that P_+ and P_- project onto orthogonal subspaces since P_+ always returns a vector proportional to $|+z\rangle$ and P_- always returns a vector proportional to $|-z\rangle$, i.e. $P_+P_- = 0$.

Recall that quantum measurements disturb the state of the system. Projection operators

allow us to model measurements mathematically. For instance, if an atom with a normalized input state $|\psi\rangle$ is measured to have $S_z = \hbar/2$, then the output state is $|+z\rangle$. The post-measurement state is determined by a *non-linear operation* on the pre-measurement state:

$$|\psi'\rangle = \frac{P_+|\psi\rangle}{\sqrt{\langle\psi|P_+|\psi\rangle}} = |+z\rangle \tag{6.7}$$

The probability that $S_z = \hbar/2$ can be written in terms of the corresponding projector as

$$|\langle+z|\psi\rangle|^2 = \langle\psi|+z\rangle\langle+z|\psi\rangle = \langle\psi|P_+|\psi\rangle \tag{6.8}$$

The "projection postulate" is often referred to as *state vector collapse*, and it clearly states that quantum measurements cannot be made without disturbing the system in an irreversible way. We do not really know what is going on in the measurement process, so we cannot explain the mechanism of the collapse of the quantum state vector. This lack of understanding makes some people uncomfortable with quantum mechanics and has been the source of much controversy. Trying to better understand the measurement process is an ongoing research problem.

6.3 Hermitian Adjoint

An operator A acts on kets as well as bras, but we have to be careful. For instance, the ket $|\phi\rangle = A|\psi\rangle$ does not correspond to the bra found by computing $\langle\psi|A$, where A acts on the right. The adjoint of an operator is defined precisely for this reason.

The **Hermitian adjoint** (or just adjoint for brevity) of a linear operator A is the unique operator denoted $A^\dagger$ such that

$$\langle\phi|A\psi\rangle = \langle A^\dagger\phi|\psi\rangle \tag{6.9}$$

for all kets $|\psi\rangle$ and $|\phi\rangle$ in a complex vector space.

(i) An operator is called **self-adjoint** if it is equal to its adjoint, i.e. $A = A^\dagger$, and the domain of A is the same as the domain of $A^\dagger$.

(ii) An operator is **Hermitian** if $A^\dagger = A$ and **anti-Hermitian** if $A^\dagger = -A$.

By conjugate symmetry of the inner product,

$$\langle A^\dagger\phi|\psi\rangle^* = \langle\psi|A^\dagger|\phi\rangle = \langle\phi|A|\psi\rangle^*$$

which relates the matrix elements of A to the matrix elements of $A^\dagger$. We can also find the bra vector corresponding to $|\phi\rangle = A|\psi\rangle$. If $|\varphi\rangle$ is an arbitrary ket, then

$$\langle\phi|\varphi\rangle = \langle\varphi|\phi\rangle^* = \langle\varphi|A|\psi\rangle^* = \langle\psi|A^\dagger|\varphi\rangle$$

Hence, $\langle\psi|A^\dagger$ is the bra corresponding to the ket $A|\psi\rangle$. Here are more properties of $A^\dagger$:

(a) The adjoint is antilinear with respect to scalar multiplication: $(\alpha A)^\dagger = \alpha^* A^\dagger$

(b) Reversal of products: $(AB)^\dagger = B^\dagger A^\dagger$

(c) The adjoint of the adjoint is the original operator: $(A^\dagger)^\dagger = A$

(d) The adjoint of the identity is the identity: $\mathbb{1}^\dagger = \mathbb{1}$

(e) The adjoint of a sum is the sum of the adjoints: $(A+B)^\dagger = A^\dagger + B^\dagger$

(f) Technically, the adjoint is only defined over linear operators, but we often write the relationship between a bra and its ket as if the bra is the "adjoint" of its ket

$$\langle\psi| = (|\psi\rangle)^\dagger \quad \text{and} \quad |\psi\rangle = (\langle\psi|)^\dagger$$

This abuse of notation makes more sense in the context of matrices, where the adjoint corresponds to the conjugate transpose of an operator's matrix representation.

7 Analysis of S-G Experiments

Armed with new mathematical tools and rules, we will analyze the ideal Stern-Gerlach experiments discussed earlier. In the process, we'll learn more about the mathematics of kets, observables, and projectors. We'll also discuss their physical significance.

7.1 Analysis of Experiment 1

The atoms coming from the upper port of the first analyzer are in the state $|+z\rangle$, and the projector corresponding to the upper port of the second analyzer is $P_+ = |+z\rangle\langle+z|$. Hence, the probability that atoms will exit the upper port of the second analyzer is

$$\langle\psi|P_+|\psi\rangle = |\langle+z|+z\rangle|^2 = 1$$

and the post-measurement state for atoms that exit the upper port is

$$\frac{P_+|+z\rangle}{\sqrt{\langle +z|P_+|+z\rangle}} = |+z\rangle$$

No atoms exit the lower port because $\langle +z|P_-|+z\rangle = |\langle -z|+z\rangle|^2 = 0$.

7.2 Analysis of Experiment 2

The second analyzer effectively performs a measurement of the S_x operator. Following the same reasoning as for the S_z operator, it follows that

$$S_x = \tfrac{\hbar}{2}|+x\rangle\langle +x| - \tfrac{\hbar}{2}|+x\rangle\langle +x| \tag{7.1}$$

where $|\pm x\rangle$ also form an orthonormal basis for the vector space. In other words, there was nothing special about the z-axis in experiment 1, meaning we could have rotated the entire apparatus by 90° and obtained the same results along the x-axis.

The results of experiment 2 would be the same if we blocked the upper port instead. Including orthonormality, we have the following results:

$$\begin{aligned} |\langle +x|+z\rangle|^2 &= |\langle -x|+z\rangle|^2 = \tfrac{1}{2} \\ |\langle +x|-z\rangle|^2 &= |\langle -x|-z\rangle|^2 = \tfrac{1}{2} \\ \langle +x|+x\rangle &= \langle -x|-x\rangle = 1 \\ \langle +x|-x\rangle &= \langle -x|+x\rangle = 0 \end{aligned} \tag{7.2}$$

This is almost enough information to determine an expression for $|\pm x\rangle$ as linear combinations of $|\pm z\rangle$. With a, b, c, d all complex numbers,

$$\begin{aligned} |+x\rangle &= a|+z\rangle + b|-z\rangle \\ |-x\rangle &= c|+z\rangle + d|-z\rangle \end{aligned}$$

Each complex number has a real and imaginary part, so there are actually 8 numbers to determine. The experimental results provide us with 7 independent conditions, so we'll need to make an arbitrary choice at some point, and for that we use the following somewhat cryptic convention: the states $|+x\rangle$, $|+y\rangle$, $|+z\rangle$ form a "right-handed triplet." We'll leave that discussion for later, but for now it is straightforward to verify that the following conventional

expressions match the experimental results:

$$|+x\rangle = \tfrac{1}{\sqrt{2}}\Big(|+z\rangle + |-z\rangle\Big) \tag{7.3}$$

$$|-x\rangle = \tfrac{1}{\sqrt{2}}\Big(|+z\rangle - |-z\rangle\Big) \tag{7.4}$$

Consider Experiment 2 in the case where the second Stern-Gerlach analyzer is aligned along the y-axis. The experimental results are identical to when the second analyzer was aligned along the x-axis, and we can use that information to determine the $|\pm y\rangle$ states in terms of $|\pm z\rangle$ as

$$|+y\rangle = \tfrac{1}{\sqrt{2}}\Big(|+z\rangle + e^{i\alpha}|-z\rangle\Big)$$

$$|-y\rangle = \tfrac{1}{\sqrt{2}}\Big(|+z\rangle - e^{i\alpha}|-z\rangle\Big)$$

where α is a real number yet to be determined. Consider a further modification to experiment 2 where the first Stern-Gerlach analyzer is aligned along the x-axis and the second analyzer is aligned along the y-axis. As you might expect, we find

$$|\langle +x|+y\rangle|^2 = \tfrac{1}{2},$$

and after expanding in the $|\pm z\rangle$ basis,

$$|\langle +x|+y\rangle|^2 = \tfrac{1}{4}\left|1 + e^{i\alpha}\right|^2 = \tfrac{1}{2}\left(1 + \cos\alpha\right) = \tfrac{1}{2}$$

This requires $\alpha = 0$ or $\alpha = \pm\pi/2$. We cannot choose $\alpha = 0$ because that would imply $|+x\rangle$ is the same as $|+y\rangle$. By convention, we choose $\alpha = \pi/2$ such that

$$|+y\rangle = \tfrac{1}{\sqrt{2}}\Big(|+z\rangle + i|-z\rangle\Big) \tag{7.5}$$

$$|-y\rangle = \tfrac{1}{\sqrt{2}}\Big(|+z\rangle - i|-z\rangle\Big) \tag{7.6}$$

Notice that the first coefficient, the one in front of $|+z\rangle$, is a real positive constant in each of the final expressions for $|\pm x\rangle, |\pm y\rangle, |\pm z\rangle$. This was not by accident, and it turns out that the state vector has an extra degree of freedom we needed to account for with another convention. Any normalized ket $|\psi\rangle$ is *physically indistinguishable* from $e^{i\alpha}|\psi\rangle$ for any real number α. We'll prove this later!

7.3 Analysis of Experiment 3

With the lower port of the second analyzer blocked, the state of the atoms coming from the upper port that go into the third analyzer is

$$|+x\rangle = \tfrac{1}{\sqrt{2}}\Big(|+z\rangle + |-z\rangle\Big)$$

Hence, the probability an atom from the upper port of the second will exit the upper port of the third is $|\langle +z|+x\rangle|^2 = 1/2$. Recall that the probability that an atom from the upper port of the first analyzer will exit the upper port of the second is $|\langle +x|+z\rangle|^2 = 1/2$. Overall, the probability that an atom from the upper port of the first analyzer will exit the upper port of the second and the upper port of the third is

$$\begin{aligned}\mathcal{P}_{\text{upper},+} &= \langle +x|P_+|+x\rangle\langle +z|P_{+x}|+z\rangle \\ &= |\langle +z|+x\rangle|^2|\langle +x|+z\rangle|^2 = \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{4}\end{aligned}$$

Likewise, the probability that an atom from the upper port of the first analyzer will exit the upper port of the second and the lower port of the third is

$$\begin{aligned}\mathcal{P}_{\text{upper},-} &= \langle +x|P_-|+x\rangle\langle +z|P_{+x}|+z\rangle \\ &= |\langle -z|+x\rangle|^2|\langle +x|+z\rangle|^2 = \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{4}\end{aligned}$$

We get the same results if we block the upper port of the second analyzer and measure the state of the atoms to be $|-x\rangle$ from the lower port before passing them off to the third analyzer. Hence,

$$\begin{aligned}\mathcal{P}_{\text{lower},+} &= \langle -x|P_+|-x\rangle\langle +z|P_{-x}|+z\rangle \\ &= |\langle +z|-x\rangle|^2|\langle -x|+z\rangle|^2 = \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{4}\end{aligned}$$

$$\begin{aligned}\mathcal{P}_{\text{lower},-} &= \langle -x|P_-|-x\rangle\langle +z|P_{-x}|+z\rangle \\ &= |\langle -z|-x\rangle|^2|\langle -x|+z\rangle|^2 = \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{4}\end{aligned}$$

7.4 Analysis of Experiment 4

Experiment #4 is a much simpler example of quantum interference than the famous single-electron (or single-photon) double-slit experiment that you may have learned in a previous course. In my opinion, the simplicity of this experimental design does a much better job at exposing some of the most fascinating aspects of quantum theory and the fundamental nature of reality itself.

Let's consider two experimental designs–one that demonstrates quantum interference and one that doesn't. In the design that *does not* demonstrate interference, we record which atoms exit the output ports of each analyzer. It doesn't matter if we actually keep a record, only that some interaction is taking place between the atom and its environment that would in principle allow us to distinguish its path. In this problem, we discuss the probability that an atom leaving the first analyzer in the $|+z\rangle$ state is detected in one of the counters connected to the output ports of the third analyzer. Such a probability involves two measurements at the second and third analyzers. The total probability is the product of the individual probabilities of each measurement.

For each output port of the third analyzer, there are two mutually exclusive paths that an atom could take. There is a probability that an atom from the upper port of the first analyzer exits the upper port of the second and the upper port of the third, and there is also a probability an atom from the upper port of the first analyzer exits the lower port of the second and the upper port of the third. The rules of probability (axiomatic in fact) dictate that the probabilities of mutually exclusive events must add. In other words, the probability that an atom will exit the upper port of the third analyzer is

$$p(+z) = \mathcal{P}_{\text{upper},+} + \mathcal{P}_{\text{lower},+} = \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{1}{2}$$

Now suppose we do not keep track of the atoms exiting the two ports of the second analyzer and simply arrange magnets to carefully recombine the beams. Classically it seems like there's no difference, but the rules of quantum mechanics say otherwise. After the second analyzer, it is possible to measure each atom in the state $|+x\rangle$ or $|-x\rangle$, so the state vector after the second analyzer must be the linear combination

$$|\psi\rangle = \tfrac{1}{\sqrt{2}}|+x\rangle + \tfrac{1}{\sqrt{2}}|-x\rangle$$

since this state would give equal probability of measuring $|\pm x\rangle$. In fact, since the initial state of the atoms coming from the upper port of the first analyzer is $|+z\rangle$, we can also write

the state as

$$|\psi\rangle = \langle +x|+z\rangle|+x\rangle + \langle -x|+z\rangle|-x\rangle$$

Depending on how we combine the beams, there could be some relative phase between the coefficients but we'll ignore that for now. The magnets are arranged to combine the beams and introduce no other effects. Eqs. (7.3) & (7.4) can be used to show that $|\psi\rangle = |+z\rangle$ and the analysis is pretty straightforward, but let's pretend we don't know this shortcut and proceed anyways.

The probability that an atom from the first analyzer will exit the upper port of the third analyzer, given that it could go through either port of the second is

$$\begin{aligned} p(+z) &= \langle\psi|P_+|\psi\rangle \\ &= \Big|\langle +x|+z\rangle\langle +z|+x\rangle + \langle -x|+z\rangle\langle +z|-x\rangle\Big|^2 \\ &= |\langle +x|+z\rangle|^2|\langle +z|+x\rangle|^2 + |\langle -x|+z\rangle|^2|\langle +z|-x\rangle|^2 \\ &\qquad + 2\mathrm{Re}\Big[\langle +x|+z\rangle\langle +z|+x\rangle\langle -x|+z\rangle^*\langle +z|-x\rangle^*\Big] \\ &= \mathcal{P}_{\mathrm{upper},+} + \mathcal{P}_{\mathrm{lower},+} + \text{interference terms} \end{aligned}$$

In this case, $p(+z) = \frac{1}{4} + \frac{1}{4} + \frac{1}{2} = 1$, so we have *constructive quantum mechanical interference*. The probability that an atom from the first analyzer will exit the lower port of the third, given that it could go through either port of the second is $p(-z) = 0$, so there is destructive interference.

We can summarize our conclusions by saying that if no measurement is made on the intermediate state, then we add amplitudes and then square to find the probability, while if an intermediate measurement is performed, we square the amplitudes first and then add to find the probability. One is the square of a sum and the other is the sum of squares, and only the former exhibits interference.

8 Matrix Notation

So far we have mathematically defined kets in terms of their inner products with other kets. If we choose a basis such as $|\pm z\rangle$, then any ket may be written as

$$|\psi\rangle = a|+z\rangle + b|-z\rangle \tag{8.1}$$

and the coefficients $a = \langle +z|\psi\rangle$ and $b = \langle -z|\psi\rangle$ uniquely specify the state. We can simplify the notation by representing the ket as a column vector. In order to do this, we first need to establish the column vectors corresponding to the basis kets, and for that we choose the "standard basis" by convention. In other words,

$$|+z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |-z\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{8.2}$$

An arbitrary state may now be written as a column vector of the form

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \tag{8.3}$$

For instance,

$$|+x\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |-x\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{8.4}$$

$$|+y\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad |-y\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \tag{8.5}$$

Let $|\psi\rangle = a|+z\rangle + b|-z\rangle$ and $|\phi\rangle = c|+z\rangle + d|-z\rangle$. Their inner product is

$$\langle\psi|\phi\rangle = \left(a^*\langle +z| + b^*\langle -z|\right)\left(c\langle +z| + d\langle -z|\right) = a^*c + b^*d. \tag{8.6}$$

In order for this result to be consistent with the rules of matrix multiplication, we use row vectors to represent bras. For instance,

$$\langle\psi| = \begin{pmatrix} a^* & b^* \end{pmatrix} \tag{8.7}$$

which is the conjugate transpose of the corresponding column vector. In matrix notation,

$$\langle\psi|\phi\rangle = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = a^*c + b^*d \tag{8.8}$$

The simplest way to understand the matrix representation of operators is to use the standard basis where $|i\rangle$ is represented by a column vector with zeros everywhere except

where it has a one in the ith row:

$$|1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix}, \quad |2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \end{pmatrix}, \quad |i\rangle = \begin{pmatrix} \vdots \\ 1 \\ \vdots \end{pmatrix}$$

The standard orthonormal basis obeys a convenient inner product relation

$$\langle j|i\rangle = \delta_{ij} \tag{8.9}$$

where $|i\rangle$ and $|j\rangle$ belong to the standard basis and δ_{ij} is the Kronecker delta symbol. The standard basis allows us to easily compute the matrix elements of any linear operator. Suppose

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

We can expand the kets as linear combinations of the standard orthonormal basis vectors as

$$|\psi'\rangle = \sum_i c'_i|i\rangle, \quad |\psi\rangle = \sum_i c_i|i\rangle$$

where c_i and c'_i are the components of the vectors in matrix representation. Since A is linear,

$$\sum_i c'_i|i\rangle = A\sum_i c_i|i\rangle = \sum_i c_i A|i\rangle.$$

Using Eq. 8.9, we can solve for each component of $|\psi'\rangle$:

$$\langle j|\psi'\rangle = \sum_i c'_i\langle j|i\rangle = \sum_i c'_i\delta_{ij} = c'_j \tag{8.10}$$

$$\Rightarrow c'_j = \sum_i c_i\langle j|A|i\rangle = \sum_i A_{ji}c_i \tag{8.11}$$

In other words, the jth component of $|\psi'\rangle$ is a linear combination of the components of $|\psi\rangle$ in a manner that is consistent with matrix multiplication if we define **the matrix elements of the operator** A as

$$A_{ji} = \langle j|A|i\rangle \tag{8.12}$$

Eq. 8.11 can be written in matrix form as

$$\begin{pmatrix} c_1' \\ c_2' \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle 1|A|1\rangle & \langle 1|A|2\rangle & \cdots \\ \langle 2|A|1\rangle & \langle 2|A|2\rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix} \tag{8.13}$$

Even in cases where the bra and ket are not elements of the standard basis, we still refer to an expression like $\langle\phi|A|\psi\rangle$ as a matrix element.

Let's compute the matrix elements of the operator S_z acting on spin-1/2 states, with $|+z\rangle = |1\rangle$ and $|-z\rangle = |2\rangle$ as our basis vectors:

$$S_z = \begin{pmatrix} \langle +z|S_z|+z\rangle & \langle +z|S_z|-z\rangle \\ \langle -z|S_z|+z\rangle & \langle -z|S_z|-z\rangle \end{pmatrix} = \begin{pmatrix} \frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{8.14}$$

Similarly, we find that

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Alternatively, we could have used $|\pm x\rangle$ as our basis, in which case

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

This is why it's important to think of operators as more abstract quantities than their matrix representations–the matrix is basis dependent but the operator is not. Changing the basis does not affect any measurement outcomes.

Let's now compute the matrix elements of the tensor product operator $|i\rangle\langle j|$ when $|i\rangle$, $|j\rangle$ are from the standard basis.

$$\big(|i\rangle\langle j|\big)_{kl} = \langle k|\big(|i\rangle\langle j|\big)|l\rangle = \langle k|\big(\langle j|l\rangle|i\rangle\big) = \langle k|i\rangle\langle j|l\rangle = \delta_{ki}\delta_{jl}$$

This matrix therefore has zero elements everywhere except when $k = i$ and $l = j$. In other words, the matrix $|i\rangle\langle j|$ is zero everywhere except for a 1 in the ith row and jth column:

$$|1\rangle\langle 1| = \begin{pmatrix} 1 & 0 & \cdots \\ 0 & 0 & \\ \vdots & & \ddots \end{pmatrix}, \quad |1\rangle\langle 2| = \begin{pmatrix} 0 & 1 & \cdots \\ 0 & 0 & \\ \vdots & & \ddots \end{pmatrix}.$$

Hence, any linear operator A can be written as

$$A = \sum_{i,j} A_{ij} |i\rangle\langle j| \tag{8.15}$$

In particular, when $i = j$, we have an elementary projection operators $|i\rangle\langle i|$. The completeness relation can be written as

$$\sum_i |i\rangle\langle i| = \mathbb{1} \tag{8.16}$$

The completeness relation is very useful in many calculations.

Given the matrix elements of an operator, we can easily compute the matrix elements of its adjoint:

$$A^\dagger_{ij} = \langle i|A^\dagger|j\rangle = \langle j|A|i\rangle^* = A^*_{ji}$$

In other words, the adjoint matrix is the conjugate transpose of the operator matrix: compute the transpose and apply complex conjugation to every matrix element. For example,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}$$

Additionally, every bra vector is the conjugate transpose of the corresponding ket, and we often use the same symbol to denote the adjoint even though this is technically an abuse of notation:

$$\begin{pmatrix} a \\ b \end{pmatrix}^\dagger = \begin{pmatrix} a^* & b^* \end{pmatrix}$$

9 Statistical Analysis

In experiment 2, we perform measurements of S_x given the input state $|+z\rangle$ and collect statistical data on the measurement outcomes. We calculate the **expectation value** of S_x as a weighted average of the measurement outcomes, weighted by the measurement-outcome probability:

$$\langle S_x\rangle = \tfrac{\hbar}{2} \cdot \tfrac{1}{2} + \left(-\tfrac{\hbar}{2}\right) \cdot \tfrac{1}{2} = 0 \tag{9.1}$$

Notice that the term "expectation value" is a bit misleading, since we do not expect nor do we ever observe $S_x = 0$ as a measurement outcome. A better name for the quantity might be "expected mean value" because it's the average we expect of a large number of experiments. It is not a time average, but an average over many identical experiments.

There is clearly **uncertainty** in the measurements, since they do not all yield the same

value, and we quantify the uncertainty by computing the standard deviation:

$$\begin{aligned}(\Delta S_x)^2 &= \langle (S_x - \langle S_x \rangle)^2 \rangle \\ &= \langle S_x^2 - 2S_x\langle S_x\rangle + S_x\rangle^2\rangle \\ &= \langle S_x^2\rangle - 2\langle S_x\rangle\langle S_x\rangle + \langle S_x\rangle^2 \\ &= \langle S_x^2\rangle - \langle S_x\rangle^2 \end{aligned} \tag{9.2}$$

The expectation value of S_x^2 is again a weighted average of each value of S_x^2 measured:

$$\langle S_x^2\rangle = \left(\tfrac{\hbar}{2}\right)^2 \cdot \tfrac{1}{2} + \left(-\tfrac{\hbar}{2}\right)^2 \cdot \tfrac{1}{2} = \tfrac{\hbar^2}{4} \tag{9.3}$$

Hence, the uncertainty of S_x given the state $|+z\rangle$ is $\Delta S_x = \hbar/2$.

We often write the expectation value of an operator A as

$$\langle A\rangle = \langle \psi | A | \psi \rangle \tag{9.4}$$

where $|\psi\rangle$ is a normalized ket representing the quantum state that A is being measured on. For instance, consider the expectation value of S_z in some arbitrary state $|\psi\rangle$:

$$\begin{aligned}\langle S_z\rangle &= \tfrac{\hbar}{2}|\langle +z|\psi\rangle|^2 - \tfrac{\hbar}{2}|\langle -z|\psi\rangle|^2 \\ &= \tfrac{\hbar}{2}\langle\psi|+z\rangle\langle +z|\psi\rangle - \tfrac{\hbar}{2}\langle\psi|-z\rangle\langle -z|\psi\rangle \\ &= \langle\psi|\Big(\tfrac{\hbar}{2}|+z\rangle\langle +z| - \tfrac{\hbar}{2}|-z\rangle\langle -z|\Big)|\psi\rangle \\ &= \langle\psi|S_z|\psi\rangle\end{aligned}$$

The uncertainty, or standard deviation, of an operator A is

$$\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2} \tag{9.5}$$

Notice that the uncertainty only makes sense mathematically if $\langle A^2\rangle - \langle A\rangle^2 \geq 0$.

Example 9.1: Expectation Value and Uncertainty

A spin-1/2 particle is in the state

$$|\psi\rangle = \frac{1}{\sqrt{5}}(|+z\rangle + 2i|-z\rangle)$$

What are the expectation value $\langle S_z\rangle$ and uncertainty ΔS_z for this state?

Solution: First we calculate the expectation value of S_z as

$$\begin{aligned}
\langle S_z\rangle &= \langle\psi|S_z|\psi\rangle \\
&= \left[\frac{1}{\sqrt{5}}(\langle+z| - 2i\langle-z|)\right] S_z \left[\frac{1}{\sqrt{5}}(|+z\rangle + 2i|-z\rangle)\right] \\
&= \frac{1}{5}\Big(\langle+z|S_z|+z\rangle + 2i\langle+z|S_z|-z\rangle - 2i\langle-z|S_z|+z\rangle + 4\langle-z|S_z|-z\rangle\Big) \\
&= \frac{1}{5}\left(\frac{\hbar}{2} - \frac{4\hbar}{2}\right) = -\frac{3\hbar}{10}
\end{aligned}$$

Now we calculate the expectation value of S_z^2.

$$\langle S_z^2\rangle = \langle\psi|S_z^2|\psi\rangle = \frac{1}{5}\left(\frac{\hbar^2}{4} + \frac{4\hbar^2}{4}\right) = \frac{\hbar^2}{4}$$

The uncertainty is then

$$\Delta S_z = \sqrt{\langle S_z^2\rangle - \langle S_z\rangle^2} = \sqrt{\frac{25\hbar^2}{100} - \frac{9\hbar^2}{100}} = \frac{4\hbar}{10}$$

Each measurement of S_z will produce either $\hbar/2$ or $-\hbar/2$, but after repeated measurements of S_z on identically prepared systems in the state $|\psi\rangle$, we find that the expectation value of S_z is $-0.3\hbar$ and the uncertainty is $0.4\hbar$.

10 Commuting Observables

In experiment 3, we learned the two observables S_x and S_z could not be known or simultaneously measured, and we called them "incompatible." Let's explore what it means for two observables to be incompatible and how incompatibility affects measurement results.

Measuring S_x projects the input state to an eigenstate of S_x, and a subsequent measurement of the observable S_z projects an eigenstate of S_x to an eigenstate of S_z. In the end, we have an eigenstate of S_z which is a superposition of S_x eigenstates. If we measure S_z and then S_x, we end up with an eigenstate of S_x which is a superpositon of S_z eigenstates. Notice that the order of operations matters with respect to the final state, and the eigenstates of S_x are different than the eigenstates of S_z. We use a mathematical operation called the commutator to measure the degree to which two operators fail to commute.

The **commutator** of two operators is defined as the difference between the products of the two operators taken in alternate orders:

$$[A, B] = AB - BA \tag{10.1}$$

If the commutator is equal to zero, we say that the operators commute.

Let A be diagonalizable with a set of non-degenerate eigenvalues. If a is an eigenvalue of A, there exists an eigenvector $|a\rangle$ such that $A|a\rangle = a|a\rangle$. If B is an operator that commutes with A,

$$A\left(B|a\rangle\right) = BA|a\rangle = B(a|a\rangle) = a\left(B|a\rangle\right).$$

Since A is non-degenerate, the vector $B|a\rangle$ must be some scalar multiple of $|a\rangle$. Let's call the scalar b such that $B|a\rangle = b|a\rangle$; this is an eigenvalue equation for B with eigenvalue a and eigenvector $|a\rangle$. Even when there is degeneracy, it turns out that **if A and B commute and are both diagonalizable, then they can be simultaneously diagonalized**, i.e., there exists a basis of common eigenvectors for both operators. Hence, we can know the eigenvalues of any two commuting observables simultaneously. It is common to extend this language and say that these two observables can be measured simultaneously. What we mean is that we can measure one observable without erasing our knowledge of the previous results of the other observable. Observables A and B are said to be **compatible**.

If two operators do not commute, then they are incompatible observables and cannot be

measured or known simultaneously. For example,

$$[S_z, S_x] = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} - \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = i\hbar S_y$$

In fact, none of the spin component operators commute and the complete set of *commutation relations* is

$$[S_x, S_y] = i\hbar S_z, \quad [S_z, S_x] = i\hbar S_y, \quad [S_y, S_z] = i\hbar S_x \tag{10.2}$$

We'll learn the origins of these commutation relations in lecture #4.

11 Uncertainty Principle

The **uncertainty principle** of quantum mechanics states that if A and B are *self-adjoint linear operators* on a complex vector space, then

$$\Delta A \Delta B \geq \tfrac{1}{2}\Big|\big\langle [A, B]\big\rangle\Big| \tag{11.1}$$

where $\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2}$ is the uncertainty (standard deviation) and

$$\langle [A, B]\rangle = \langle\psi|[A, B]|\psi\rangle$$

for all states $|\psi\rangle$ in the vector space.

We will prove this theorem after we've learned more about self-adjoint operators (also called Hermitian operators), which play a significant role in quantum theory. For now, let's consider what it says about a simple Stern-Gerlach experiment. For example,

$$\begin{aligned}\Delta S_x \Delta S_y &\geq \tfrac{1}{2}\Big|\big\langle [S_x, S_y]\big\rangle\Big| \\ &\geq \tfrac{1}{2}\Big|i\hbar\langle S_z\rangle\Big| \\ &\geq \tfrac{\hbar}{2}\Big|\langle S_z\rangle\Big|\end{aligned} \tag{11.2}$$

If the quantum state is an eigenstate of either S_x or S_y, then either ΔS_x or ΔS_y will vanish, and it's easy to show that the expectation value of S_z vanishes as well. Anytime S_z has a nonzero expectation value, the uncertainty principle tells us that neither ΔS_x nor ΔS_y can

be zero. For instance, suppose we know that $S_z = \hbar/2$ such that the state of the system is $|+z\rangle$. Then,

$$\Delta S_x \Delta S_y \geq \left(\tfrac{\hbar}{2}\right)^2$$

The uncertainties can be calculated and it turns out that, in this case, $\Delta S_x = \hbar/2$ and $\Delta S_y = \hbar/2$ such that

$$\Delta S_x \Delta S_y = \left(\tfrac{\hbar}{2}\right)^2$$

and the uncertainty principle is satisfied. We say that the spin-up state $|+z\rangle$ is a **minimum uncertainty state** for the observables S_x and S_y, meaning the uncertainty relation is saturated in this state. Although S_x and S_y are individually uncertain, their uncertainties are balanced in such a way that their product attains the minimum allowed by quantum mechanics.

Example 11.1: Uncertainty Principle

A spin-1/2 particle is in the state

$$|\psi\rangle = \frac{1}{\sqrt{5}}(|+z\rangle + 2i|-z\rangle)$$

Repeated measurements of S_x and S_y are made on the particle prepared in this state. Verify the uncertainty relation $\Delta S_x \Delta S_y \geq \frac{\hbar}{2}|\langle S_z\rangle|$.

Solution: Let's use matrix notation to speed this along.

$$\langle S_x\rangle = \langle\psi|S_x|\psi\rangle = \frac{1}{\sqrt{5}}\begin{pmatrix}1 & -2i\end{pmatrix}\cdot\frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\cdot\frac{1}{\sqrt{5}}\begin{pmatrix}1\\ 2i\end{pmatrix}$$

$$= \frac{\hbar}{10}\begin{pmatrix}1 & -2i\end{pmatrix}\begin{pmatrix}2i\\ 1\end{pmatrix} = 0$$

$$\langle S_x^2\rangle = \langle\psi|S_x^2|\psi\rangle = \frac{1}{\sqrt{5}}\begin{pmatrix}1 & -2i\end{pmatrix}\cdot\frac{\hbar^2}{4}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\cdot\frac{1}{\sqrt{5}}\begin{pmatrix}1\\ 2i\end{pmatrix}$$

$$= \frac{\hbar^2}{20}\begin{pmatrix}1 & -2i\end{pmatrix}\begin{pmatrix}1\\ 2i\end{pmatrix} = \frac{\hbar^2}{4}$$

The uncertainty $\Delta S_x = 0.5\hbar$.

$$\langle S_y \rangle = \langle \psi | S_y | \psi \rangle = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2i \end{pmatrix} \cdot \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2i \end{pmatrix}$$

$$= \frac{\hbar}{10} \begin{pmatrix} 1 & -2i \end{pmatrix} \begin{pmatrix} 2 \\ i \end{pmatrix} = \frac{4\hbar}{10}$$

$$\langle S_y^2 \rangle = \langle \psi | S_y^2 | \psi \rangle = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -2i \end{pmatrix} \cdot \frac{\hbar^2}{4} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2i \end{pmatrix}$$

$$= \frac{\hbar^2}{20} \begin{pmatrix} 1 & -2i \end{pmatrix} \begin{pmatrix} 1 \\ 2i \end{pmatrix} = \frac{\hbar^2}{4}$$

The uncertainty $\Delta S_y = 0.3\hbar$. In an earlier example, we found $\langle S_z \rangle = -0.3\hbar$.

$$\Delta S_x \Delta S_y = (0.5\hbar)(0.3\hbar) = 0.15\hbar^2 \quad \text{and} \quad \frac{\hbar}{2} |\langle S_z \rangle| = 0.15\hbar^2$$

Evidently, $|\psi\rangle$ is a "minimum uncertainty state" for S_x and S_y. **Your turn**. Verify that the following uncertainty relation is satisfied: $\Delta S_y \Delta S_z \geq \frac{\hbar}{2} |\langle S_x \rangle|$.

References

[1] Martin Bauer. *Translation of: "Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld"*. 2023. arXiv: 2301.11343 [physics.hist-ph]. URL: https://arxiv.org/abs/2301.11343.

[2] Jacob A. Barandes. *Manifestly Covariant Lagrangians, Classical Particles with Spin, and the Origins of Gauge Invariance*. 2023. arXiv: 1911.08892 [physics.gen-ph]. URL: https://arxiv.org/abs/1911.08892.

[3] Hans C. Ohanian. "What is spin?" In: *American Journal of Physics* 54.6 (June 1986), pp. 500–505. ISSN: 0002-9505. DOI: 10.1119/1.14580. eprint: https://pubs.aip.org/aapt/ajp/article-pdf/54/6/500/11833257/500_1_online.pdf. URL: https://doi.org/10.1119/1.14580.

[4] Gautam A. Kavuri et al. "Traceable random numbers from a non-local quantum advantage". In: *Nature* 642.8069 (2025), pp. 916–921. DOI: 10.1038/s41586-025-09054-3. URL: https://doi.org/10.1038/s41586-025-09054-3.