Spin and Quantum Measurement

David H. McIntyre
Oregon State University

PH 425 Paradigm 5
19 December 2002

PREFACE

This text was developed as part of a course on Spin and Quantum Measurement at Oregon State University. The course is part of the Paradigms in Physics project, which entailed a reform of the junior level physics curriculum. The Spin and Quantum Measurement course is an introduction to quantum mechanics through the analysis of sequential Stern-Gerlach spin measurements. The approach and material are based upon previous presentations of spin systems by Feynman, Sakurai, Cohen-Tannoudji, and Townsend. The postulates of quantum mechanics are illustrated through their manifestation in the simple spin 1/2 Stern-Gerlach experiments. The organization of the postulates follows the presentation of Cohen-Tannoudji. The table below lists the postulates briefly and their manifestation in the spin 1/2 system as presented in the course.

Postulates of Quantum Mechanics		Spin 1/2 manifestation
1)	State defined by ket	$\|+\rangle,\ \|-\rangle$
2)	Operators, observables	S_i, $\mathbf{S}$, H
3)	Measure eigenvalues	$\pm\hbar/2$
4)	Probability	$\left\|\langle+\|\psi\rangle\right\|^2$
5)	State reduction	$\|\psi\rangle \rightarrow \|+\rangle$
6)	Schrödinger equation evolution	Larmor precession

The specific examples covered are: sequential Stern-Gerlach measurements of spin 1/2 and spin 1 systems, spin precession in a magnetic field, spin resonance in an oscillating magnetic field, neutrino oscillations, and the EPR experiment. The tools of Dirac notation and matrix notation are used throughout the course. General two- and three-state quantum mechanical systems are also covered as simple extensions of the spin systems.

The Stern-Gerlach experiments are discussed in class and are performed by the students using a software program that simulates the experiments on spin 1/2 and spin 1 systems (also SU(3) for those ambitious enough!). The program permits the students to study any configuration of sequential Stern-Gerlach measurements, interferometers, spin precession in a magnetic field, and which path (Welcher Weg) detection to destroy interference. The program provides the student with unknown quantum states that he must determine through experiment. The program is available on the web at http://www.physics.orst.edu/~mcintyre/ph425/spins/.

The aim of the text is twofold: (1) To give the students an immersion into the **quantum** spookiness of quantum mechanics by focusing on simple measurements that have no classical explanation, and (2) To give the students experience with the **mechanics** of quantum mechanics in the form of Dirac and matrix notation. Since these two goals are so at odds with classical mechanics, the simplicity of the spin 1/2 system allows the students to focus on these new features instead of the complexity of the problem at hand.

Chapter 1 STERN-GERLACH EXPERIMENTS

1.1 Introduction

Quantum mechanics is based upon a set of postulates that dictates how to treat a quantum mechanical system mathematically and how to interpret the mathematics to learn about the physical system in question. These postulates cannot be proven, but they have been successfully tested by many experiments, and so we accept them as an accurate way to describe quantum mechanical systems. New results could force us to reevaluate these postulates at some later time. The postulates are listed below to give you an idea where we are headed and a framework into which you can place the new concepts as we confront them.

Postulates of Quantum Mechanics

1. The state of a quantum mechanical system is described mathematically by a normalized ket $|\psi\rangle$ that contains all the information we can know about the system.
2. A physical observable is described mathematically by an operator A that acts on kets.
3. The only possible result of a measurement of an observable is one of the eigenvalues a_n of the corresponding operator A.
4. The probability of obtaining the eigenvalue a_n in a measurement of the observable A on the system in the state $|\psi\rangle$ is

$$\mathcal{P}(a_n) = \left|\langle a_n|\psi\rangle\right|^2,$$

where $|a_n\rangle$ is the normalized eigenvector of A corresponding to the eigenvalue a_n.

5. After a measurement of A that yields the result a_n, the quantum system is in a new state that is the normalized projection of the original system ket onto the ket (or kets) corresponding to the result of the measurement:

$$|\psi'\rangle = \frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}}.$$

6. The time evolution of a quantum system is determined by the Hamiltonian or total energy operator $H(t)$ through the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle.$$

As you read these postulates for the first time, you will undoubtedly encounter new terms and concepts. Rather than explain them all here, the plan of this text is to explain them through their manifestation in one of the simplest yet most instructive examples in quantum mechanics – the Stern-Gerlach spin 1/2 experiment. We choose this example because it is inherently quantum mechanical and forces us to break away from reliance on classical intuition or concepts. Moreover, this simple example is a paradigm for many other quantum mechanical systems. By studying it in detail, we can appreciate much of the richness of quantum mechanics.

1.2 Stern-Gerlach experiment

The Stern-Gerlach experiment is a conceptually simple experiment that demonstrates many basic principles of quantum mechanics. Studying this example has two primary benefits: (1) It demonstrates how quantum mechanics works in principle by illustrating the postulates of quantum mechanics, and (2) It demonstrates how quantum mechanics works in practice through the use of Dirac notation and matrix mechanics to solve problems. By using an extremely simple example, we can focus on the principles and the new mathematics, rather than having the complexity of the physics obscure these new aspects.

In 1922 Otto Stern and Walter Gerlach performed a seminal experiment in the history of quantum mechanics. In its simplest form, the experiment consists of an oven that produces a beam of neutral atoms, a region of inhomogeneous magnetic field, and a detector for the atoms, as depicted in Fig. 1.1. Stern and Gerlach used a beam of silver atoms and found that the beam was split into two in its passage through the magnetic field. One beam was deflected upwards and one downwards in relation to the direction of the magnetic field gradient.

To understand why this result is so at odds with our classical expectations, we

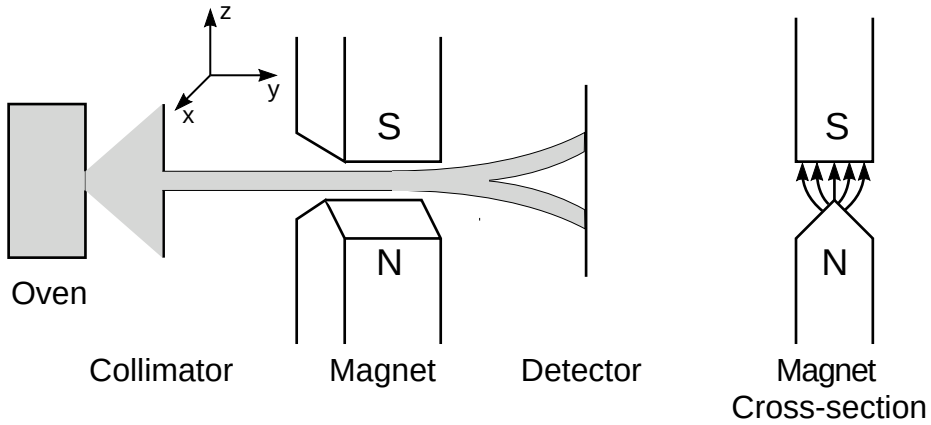

Figure 1.1. Stern-Gerlach experiment to measure spin projection of neutral particles along the z-axis. The magnetic cross-section at right shows the inhomogeneous field used in the experiment.

must first analyze the experiment classically. The results of the experiment suggest an interaction between a neutral particle and a magnetic field. We expect such an interaction if the particle possesses a magnetic moment **μ**. The energy of this interaction is given by $E = -\boldsymbol{\mu} \bullet \mathbf{B}$, which results in a force $\mathbf{F} = \nabla(\boldsymbol{\mu} \bullet \mathbf{B})$. In the Stern-Gerlach experiment, the magnetic field gradient is primarily in the z-direction, and the resulting z–component of the force is

$$\begin{aligned} F_z &= \frac{\partial}{\partial z}(\boldsymbol{\mu} \bullet \mathbf{B}) \\ &\cong \mu_z \frac{\partial B_z}{\partial z} \end{aligned} \quad . \tag{1.1}$$

This force is perpendicular to the direction of motion and deflects the beam in proportion to the magnitude of the magnetic moment in the direction of the magnetic field gradient.

Now consider how to understand the origin of the atom's magnetic moment from a classical viewpoint. The atom consists of charged particles, which, if in motion, can produce loops of current that give rise to magnetic moments. A loop of area *A* and current *I* produces a magnetic moment

$$\mu = \frac{IA}{c} \tag{1.2}$$

in cgs units. If this loop of current arises from a charge *q* traveling at speed *v* in a circle of radius *r*, then

$$\begin{aligned} \mu &= \frac{1}{c}\frac{q}{2\pi r/v}\pi r^2 \\ &= \frac{qrv}{2c} \\ &= \frac{q}{2mc}L \end{aligned} \quad , \tag{1.3}$$

where $L = mrv$ is the orbital angular momentum of the particle. 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 intrinsic rotational angular momentum, which we call **S**. The intrinsic angular momentum also creates current loops, so we expect a similar relation between the magnetic moment **μ** and **S**. The exact calculation involves an integral over the charge distribution, which we will not do. We simply assume that we can relate the magnetic moment to the intrinsic angular momentum in the same fashion as Eq. (1.3), giving

$$\boldsymbol{\mu} = g\frac{q}{2mc}\mathbf{S}, \tag{1.4}$$

where the **gyroscopic ratio** *g* contains the details of that integral.

A silver atom has 47 electrons, 47 protons, and 60 or 62 neutrons (for the most common isotopes). Since the magnetic moments depend on the inverse of the particle

mass, we expect the heavy protons and neutrons ($\approx 2000\, m_e$) to have little effect on the magnetic moment of the atom and so we neglect them. From your study of the periodic table in chemistry, you recall that silver has an electronic configuration $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 4d^{10} 5s$, which means that there is only the lone 5*s* electron outside of the closed shells. The electrons in the closed shells can be represented by a spherically symmetric cloud with no orbital or intrinsic angular momentum (unfortunately we are injecting some quantum mechanical knowledge of atomic physics into this classical discussion). That leaves the lone 5*s* electron as a contributor to the magnetic moment of the atom as a whole. An electron in an *s* state has no orbital angular momentum, but it does have intrinsic angular momentum, which we call **spin**. Hence the magnetic moment of this electron, and therefore of the entire neutral silver atom, is

$$\boldsymbol{\mu} = -\frac{eg}{2m_e c}\mathbf{S}, \tag{1.5}$$

where *e* is the magnitude of the electron charge. The classical force on the atom can now be written as

$$F_z \cong -\frac{eg}{2m_e c} S_z \frac{\partial B_z}{\partial z}. \tag{1.6}$$

The deflection of the beam in the Stern-Gerlach experiment is thus a measure of the component or projection S_z of the spin along the z-axis, which is the orientation of the magnetic field gradient.

If we assume that each electron has the same magnitude $|\mathbf{S}|$ of the intrinsic angular momentum or spin, then classically we would write the projection as $S_z = |\mathbf{S}|\cos\theta$, where θ is the angle between the z-axis and the direction of the spin **S**. In the thermal environment of the oven, we expect a random distribution of spin directions and hence all possible angles θ. Thus we expect some continuous distribution (the details are not important) of spin projections from $S_z = -|\mathbf{S}|$ to $S_z = +|\mathbf{S}|$, which would yield a continuous spread in deflections of the silver atomic beam. Rather, the experimental result is that there are only two deflections, indicating that there are only two possible values of the spin projection of the electron. The magnitudes of these deflections are consistent with values of the spin projection of

$$S_z = \pm\frac{\hbar}{2}, \tag{1.7}$$

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

$$\begin{aligned} \hbar &= 1.0546\times10^{-27}\,\text{erg}\cdot\text{s} \\ &= 6.5821\times10^{-16}\,\text{eV}\cdot\text{s} \end{aligned}. \tag{1.8}$$

This result of the Stern-Gerlach experiment is evidence of the **quantization** of the electron's spin angular momentum projection along an axis. This quantization is at odds with our classical expectations for this measurement. The factor of 1/2 in Eq. (1.7) leads us to refer to this as a **spin 1/2** system. In this example, we have chosen the z-axis along

which to measure the spin projection, but we could have chosen any other axis and would have obtained the same results.

Now that we know the fine details of the Stern-Gerlach experiment, we simplify the experiment for the rest of our discussions by focusing on the essential features. A simplified schematic representation of the experiment is shown in Fig. 1.2, which depicts an oven that produces the beam of atoms, a Stern-Gerlach device with two output ports for the two possible values of the spin projection, and two counters to detect the atoms leaving the output ports of the Stern-Gerlach device. The Stern-Gerlach device is labeled with the axis along which the magnetic field is oriented. The up and down arrows indicate the two possible measurement results for the device; they correspond respectively to the results $S_z = \pm\hbar/2$ in the case where the field is oriented along the z–axis. Since there are only two possible results in this case, they are generally referred to simply as **spin up** and **spin down**. The physical quantity that is measured, S_z in this case, is called an **observable**. In our detailed discussion of the experiment above, we chose the field gradient in such a manner that the spin up states were deflected upwards. In this new simplification, the deflection is not an important issue. We simply label the output port with the desired state and count the particles leaving that port.

In Fig. 1.2, the output beams have also been labeled with a new symbol called a **ket**. We use the ket $|+\rangle$ as a mathematical representation of the quantum state of the atoms that exit the upper port corresponding to $S_z = +\hbar/2$. The lower output beam is labeled with the ket $|-\rangle$, which corresponds to $S_z = -\hbar/2$. According to postulate 1, which is repeated below, these kets contain all the information that we can know about the system. Since there are only two possible results of the measurement, there are only two kets for this system (we are ignoring the position and velocity of the atoms in the beam). This ket notation was developed by P. A. M. Dirac and is central to the approach to quantum mechanics that we will take in this text. We will discuss the mathematics of these kets in full detail later. For now, it is sufficient for us to consider the ket as simply labeling the quantum state. With regard to notation, you will find many different ways of writing the $|\pm\rangle$ kets ($|\pm\rangle$ refers to both the $|+\rangle$ and $|-\rangle$ kets). The information contained within the ket symbol is used merely to label the ket and to distinguish the ket from other different kets. For example, the kets $|+\rangle$, $|+\hbar/2\rangle$, $|S_z = +\hbar/2\rangle$, $|+\hat{\mathbf{z}}\rangle$, and $\left|\uparrow\right\rangle$ are all equivalent ways of writing the same thing, and they all behave the same mathematically.

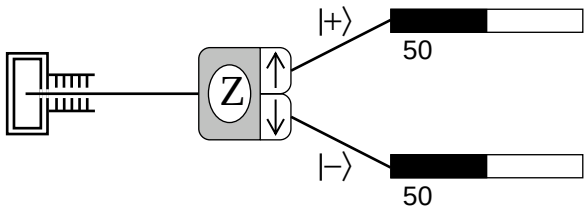

Figure 1.2. Simplified schematic of Stern-Gerlach experiment, depicting source of atoms, Stern-Gerlach analyzer, and counters.

> Postulate 1
> The state of a quantum mechanical system is described mathematically by a normalized ket $|\psi\rangle$ that contains all the information we can know about the system.

We have chosen the particular simplified schematic representation of Stern-Gerlach experiments shown in Fig. 1.2 because it is the same representation used in the SPINS software program that you may use to simulate these experiments. The SPINS program allows you to perform all the experiments described in this text. In the program, the components are simply connected together to represent the paths the atoms take. The directions and deflections of the beams in the program are not relevant, and so we follow that lead in our depiction of the experiment hereafter. That is, whether the spin up output beam is drawn as deflected upwards, or downwards, or not all is not relevant. The labeling on the output port is enough to tell us what that state is. Thus the extra ket label $|+\rangle$ on the spin up output beam in Fig. 1.2 is redundant and will be dropped soon.

The SPINS program permits alignment of Stern-Gerlach analyzing devices along all three axes and also at any angle ϕ measured from the x-axis in the x-y plane. This would appear to be difficult, if not impossible, given that the atomic beam in Fig. 1.1 is directed along the y-axis, making it unclear how to align the magnet in the y-direction and measure a deflection. In our depiction and discussion of Stern-Gerlach experiments, we ignore this technical complication.

In the SPINS program, as in real Stern-Gerlach experiments, the numbers of atoms detected in particular states are determined by probability rules that we will discuss later. To simplify our schematic depictions of Stern-Gerlach experiments, the numbers shown for detected atoms are obtained by simply using the calculated probabilities without any regard to possible statistical uncertainties. That is, if the probabilities of two possibilities are each 50%, then our schematics will display equal numbers for those two possibilities, whereas in a real experiment, statistical uncertainties might yield a 55%/45% split in one experiment and a 47%/53% split in another, etc. In your SPINS program simulations, you will note these statistical uncertainties and so will need to perform enough experiments to convince yourself that you have a sufficiently good estimate of the probability (see Appendix A for more information on statistics).

Now consider a series of simple Stern-Gerlach experiments with slight variations that help to illustrate the main features of quantum mechanics. We first describe the experiments and their results and draw some qualitative conclusions about the nature of quantum mechanics. Then we introduce the formal mathematics of the ket notation and show how it can be used to predict the results of each of the experiments.

1.2.1 Experiment 1

The first experiment is shown in Fig. 1.3 and consists of a source of atoms, two Stern-Gerlach devices both aligned along the z-axis, and counters for some of the output

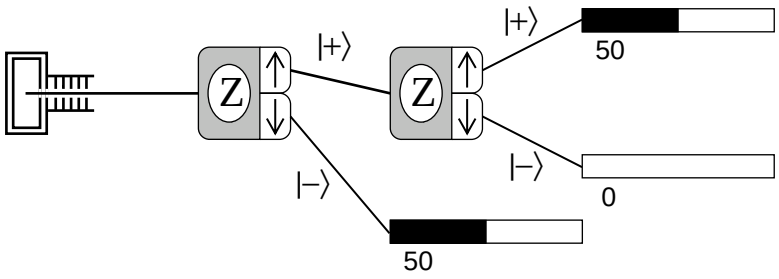

Figure 1.3. Experiment 1 measures the spin projection along the z-axis twice in succession.

ports of the analyzers. The atomic beam coming into the 1st Stern-Gerlach device is split into two beams at the output, just like the original experiment. Now instead of counting the atoms in the upper output beam, the spin projection is measured again by directing those atoms into the 2nd Stern-Gerlach device. The result of this experiment is that no atoms are ever detected coming out of the lower output port of the 2nd Stern-Gerlach device. All atoms that are output from the upper port of the 1st device also pass through the upper port of the 2nd device. Thus we say that when the 1st Stern-Gerlach device measures an atom to have $S_z = +\hbar/2$, then the 2nd device also measures $S_z = +\hbar/2$ for that atom.

Though both devices are identical, the 1st device is often referred to as the **polarizer** and the 2nd one as the **analyzer**, since the 1st one "polarizes" the beam along the z-axis and the second one "analyzes" the resultant beam. This is analogous to what can happen with optical polarizers. Some also refer to the 1st analyzer as a **state preparation device**, since it prepares the quantum state that is then measured with the analyzer. By preparing the state in this manner, the details of the source of atoms can be ignored. Thus our main focus in Experiment 1 is what happens at the analyzer, since we know that any atom entering the analyzer is described by the $|+\rangle$ ket prepared by the polarizer. All the experiments we will describe employ a polarizer to prepare the state, though the SPINS program has a feature where the state of the atoms coming from the oven is determined but unknown and the user can perform experiments to figure out the unknown state.

1.2.2 Experiment 2

The second experiment is shown in Fig. 1.4 and is identical to Experiment 1 except that the analyzer has been rotated by 90˚ to be aligned with the x-axis. Now the analyzer measures the spin projection along the x-axis rather the z-axis. Atoms input to the analyzer are still described by the ket $|+\rangle$ since the polarizer is unchanged. The result of this experiment is that atoms appear at both possible output ports of the analyzer. Atoms leaving the upper port of the analyzer have been measured to have $S_x = +\hbar/2$ and atoms leaving the lower port have $S_x = -\hbar/2$. On average, each of these ports has 50% of the atoms that left the upper port of the analyzer. As shown in Fig. 1.4, the output

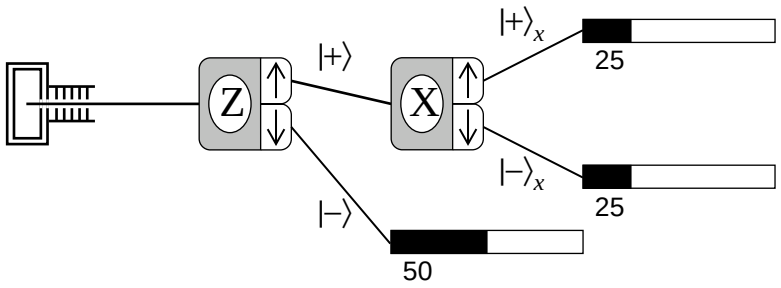

Figure 1.4. Experiment 2 measures the spin projection along the z-axis and then along the x-axis.

states of the 2nd analyzer have new labels $|+\rangle_x$ and $|-\rangle_x$, where the x subscript denotes that the spin projection has been measured along the x-axis. We assume that if no subscript is present on the quantum ket, then the spin projection is along the z-axis. This use of the z–axis as the default is common throughout our work and also in much of physics.

A few items are noteworthy about this experiment. First, we notice that there are still only two possible outputs of the Stern-Gerlach analyzer. The fact that it is aligned along a different axis doesn't affect the fact that we can only ever get two possible results for the case of a spin 1/2 particles. Second, note that the results of this experiment would be unchanged if we used the lower port of the polarizer. That is, atoms entering the analyzer in state $|-\rangle$ would also result in half the atoms in each of the $|\pm\rangle_x$ output ports. Finally, note that we cannot predict which of the analyzer output ports any particular atom will come out. This can be demonstrated experimentally by recording the counts out of each port. The arrival sequences at any counter are completely random. We can only say that there is a 50% probability that an atom from the polarizer will exit the upper analyzer port and a 50% probability that it will exit the lower port. The random arrival of atoms at the detectors can be seen clearly in the SPINS program simulations.

This probabilistic nature is at the heart of quantum mechanics. One might be tempted to say that we just don't know enough about the system to predict which port the atom will be registered in. That is to say, there may be some other variables, of which we are ignorant, that would allow us to predict the results. Such a viewpoint is know as a hidden variable theory, and such theories have been proven to be incompatible with quantum mechanics. John Bell proved that such a quantum mechanical system cannot be described by a hidden variable theory, which amounts to saying that the system cannot have things we don't know about. It is a pretty powerful statement to be able to say that there are not things that we cannot know about a system. The conclusion to draw from this is that even though quantum mechanics is a probabilistic theory, it is a complete description of reality. We will have more to say about this later.

Note that the 50% probability referred to above is the probability that an atom input to the analyzer exits one particular output port. It is not the probability for an atom

to pass through the whole system of Stern-Gerlach devices. Later we will have occasion to ask about such a probability and then we will say so. Note also that the results of this experiment (the 50/50 split at the analyzer) would be the same for any combination of two orthogonal axes of the polarizer and analyzer.

1.2.3 Experiment 3

Now consider Experiment 3, shown in Fig. 1.5, which extends Experiment 2 by adding a third Stern-Gerlach device aligned along the z-axis. (In this case, we refer to each device as an analyzer and label them first, second, or third.) Atoms entering the new third analyzer have been measured by the first Stern-Gerlach analyzer to have spin projection up along the z-axis, and by the second analyzer to have spin projection up along the x-axis. The third analyzer then measures how many atoms have spin projection up or down along the z-axis. Classically, one would expect that the final measurement would yield the result spin up along the z-axis, since that was measured at the first analyzer. That is to say: classically the first 2 analyzers tell us that the atoms have $S_z = +\hbar/2$ and $S_x = +\hbar/2$, so the third measurement must yield $S_z = +\hbar/2$. But that doesn't happen. The quantum mechanical result is that the atoms are split with 50% probability into each output port at the third analyzer. Thus the last two analyzers behave like the two analyzers of Experiment 2 (except with the order reversed), and the fact that there was an initial measurement that yielded $S_z = +\hbar/2$ is somehow forgotten or erased.

This result demonstrates another key feature of quantum mechanics: the measurement perturbs the system. One might ask: Can I be more clever in designing the experiment such that I don't perturb the system? The short answer is no. There is a fundamental incompatibility in trying to measure the spin projection of the atom along two different directions. So we say that S_x and S_z are **incompatible observables**. We cannot know the values of both simultaneously. The state of the system can be described by the ket $|+\rangle = |S_z = +\hbar/2\rangle$ or by the ket $|+\rangle_x = |S_x = +\hbar/2\rangle$, but it cannot be described by a ket $|S_z = +\hbar/2, S_x = +\hbar/2\rangle$ that specifies values of both projections. Having said this, it should be noted that not all pairs of quantum mechanical observables are incompatible. It is possible to do some experiments without perturbing some other aspects of the system. And we will see later that whether two observables are compatible

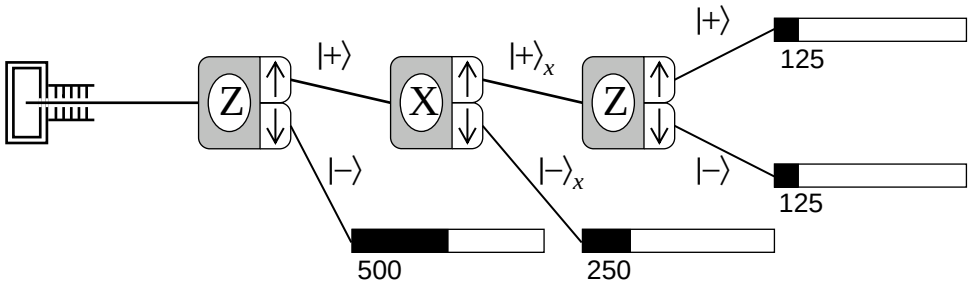

Figure 1.5. Experiment 3 measures the spin projection three times in succession.

or not is very important in how we analyze a quantum mechanical system.

Not being able to measure both the S_z and S_x spin projections is clearly distinct from the classical case whereby we can measure all three components of the spin vector, which tells us which direction the spin is pointing. In quantum mechanics, we cannot know which direction the spin is pointing. So when we say the spin is up, we really mean only that the spin projection along that one axis is up (vs. down). The spin is not really pointing in any given direction. This is an example of where you must check your classical intuition at the door.

1.2.4 Experiment 4

Experiment 4 is depicted in Fig. 1.6 and is a slight variation on Experiment 3. Before we get into the details, note a few changes in the schematic drawings. As promised, we have dropped the ket labels on the beams since they are redundant. We have deleted the counters on all but the last analyzer and instead simply block the unwanted beams and give the average number of atoms passing from one analyzer to the next. Note also that in Experiment 4c two output beams are combined as input to the following analyzer. This is simple in principle and in the SPINS program, but can be difficult in practice. The recombination of the beams must be done properly so as to

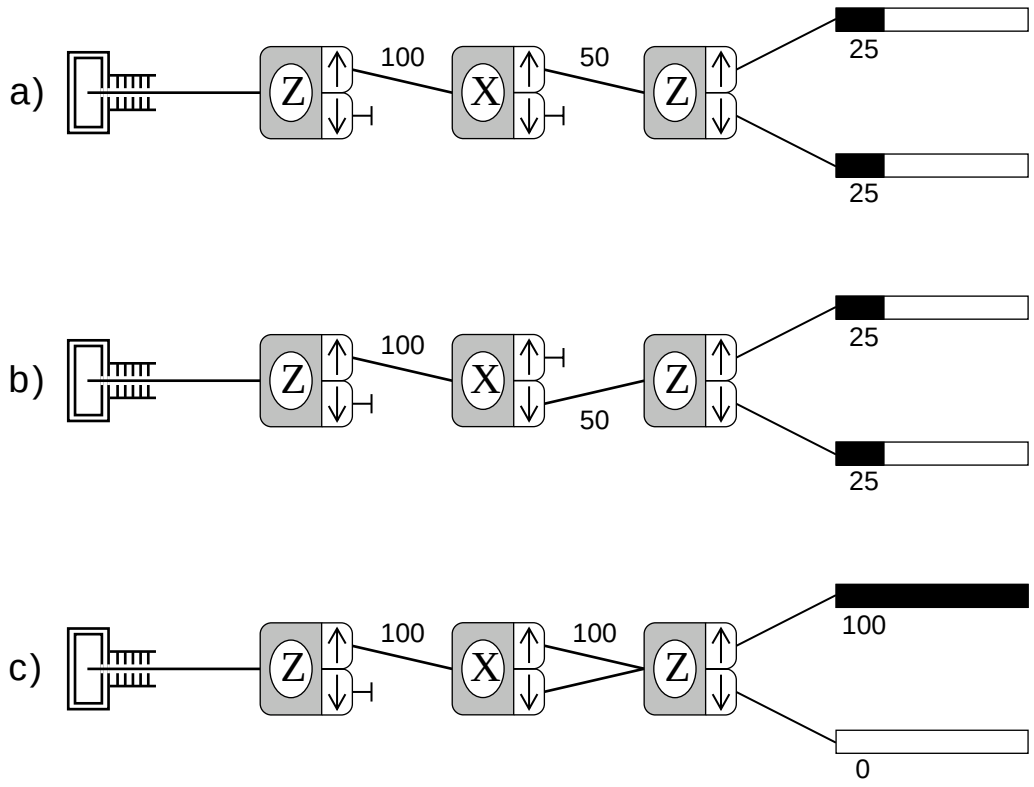

Figure 1.6. Experiment 4 measures the spin projection three times in succession and uses one (a and b) or two beams (c) from the middle analyzer.

avoid "disturbing" the beams. If you care to read more about this problem, see Feynman's Lectures on Physics, volume 3. We will have more to say about the "disturbance" later. For now we simply assume that the beams can be recombined in the proper manner.

Experiment 4a is identical to Experiment 3. In Experiment 4b the upper beam of the middle analyzer is blocked and the lower beam is sent to the third analyzer. In Experiment 4c, both beams are combined with our new method and sent to the third analyzer. It should be clear from our previous experiments that Experiment 4b has the same results as Experiment 4a. We now ask what the results of Experiment 4c are. If we were to use classical probability analysis, then Experiment 4a would indicate that the probability for an atom leaving the first analyzer to take the upper path through the second analyzer and then exit through the upper port of the third analyzer is 25%, where we are now referring to the total probability for those two steps. Likewise, Experiment 4b would indicate that the probability to take the lower path through the second analyzer and exit through the upper port of the third analyzer is also 25%. Hence the total probability to exit from the upper port of the third analyzer when both paths are available, which is simply Experiment 4c, would be 50%, and likewise for the exit from the lower port.

However, the quantum mechanical result in Experiment 4c is that all the atoms exit the upper port of the third analyzer and none exits the lower port. The atoms now appear to "remember" that they were initially measured to have spin up along the z-axis. By combining the two beams from the middle analyzer, we have avoided the quantum mechanical perturbation that was evident in Experiment 3. The result is now the same as Experiment 1, which means it is as if the middle analyzer is not there.

To see how odd this is, look carefully at what happens at the lower port of the third analyzer. In this discussion, we refer to percentages of atoms leaving the first analyzer, since that analyzer is the same in all three experiments. In Experiments 4a and 4b, 50% of the atoms are blocked after the middle analyzer and 25% of the atoms exit the lower port of the third analyzer. In Experiment 4c, 100% of the atoms pass from the second analyzer to the third analyzer, yet fewer atoms come out of the lower port. In fact, no atoms make it through the lower port! So we have a situation where allowing more ways or paths to reach a counter results in fewer counts. Classical probability theory cannot explain this aspect of quantum mechanics.

However, you may already know of a way to explain this effect. Imagine a procedure whereby combining two effects leads to cancellation rather than enhancement. The concept of wave interference, especially in optics, comes to mind. In the Young's double slit experiment, light waves pass through two narrow slits and create an interference pattern on a distant screen, as shown in Fig. 1.7. Either slit by itself produces a nearly uniform illumination of the screen, but the two slits combined produced bright and dark fringes. We explain this by adding together the electric field vectors of the light from the two slits, then squaring the resultant vector to find the light intensity. We say that we add the amplitudes and then square the total amplitude to find the resultant intensity.

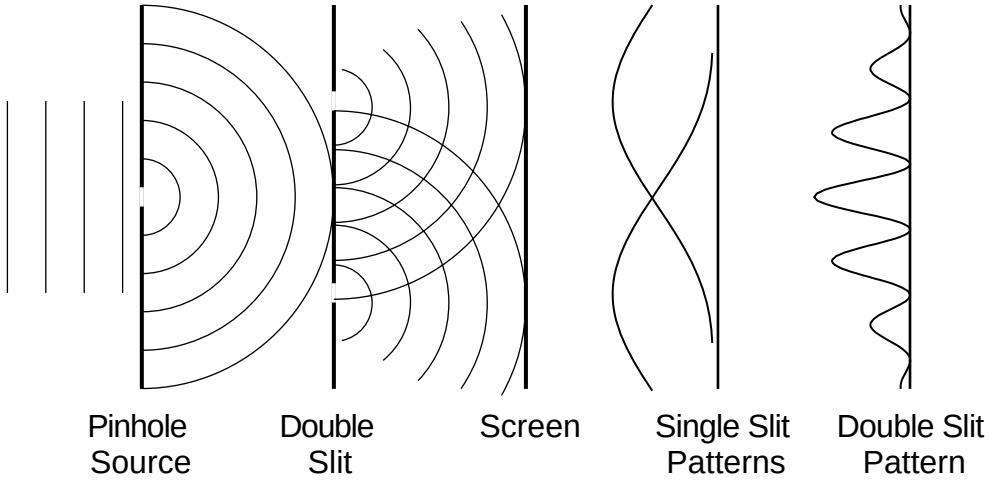

Figure 1.7. Young's double slit interference experiment.

We follow a similar prescription in quantum mechanics. We add together amplitudes and then take the square to find the resultant probability. Before we do this, we need to explain what we mean by an amplitude in quantum mechanics and how we calculate it.

1.3 Quantum State Vectors

Postulate 1 stipulates that kets are to be used for a mathematical description of a quantum mechanical system. These kets are abstract vectors that obey many of the rules you know about ordinary spatial vectors. Hence they are often called **quantum state vectors**. As we will show later, these vectors must employ complex numbers in order to properly describe quantum mechanical systems. Quantum state vectors are part of a vector space whose dimensionality is determined by the physics of the system at hand. In the Stern-Gerlach example, the two possible results for a spin projection measurement dictate that the vector space has only two dimensions. That makes this problem simple, which is why we have chosen to study it. Since the quantum state vectors are abstract, it is hard to say much about what they are, other than how they behave mathematically and how they lead to physical predictions.

In the two-dimensional vector space of a spin 1/2 system, the two kets $|\pm\rangle$ form a basis, just like the unit vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ form a basis for describing vectors in three dimensional space. However, the analogy we want to make with these spatial vectors is only mathematical, not physical. The spatial unit vectors have three important mathematical properties that are characteristic of a basis: the basis vectors are **orthogonal**, **normalized**, and **complete** (meaning any vector in the space can be written as a linear superposition of the basis vectors). These properties of spatial basis vectors can be summarized as follows:

$$\begin{aligned} \hat{\mathbf{i}} \bullet \hat{\mathbf{i}} = \hat{\mathbf{j}} \bullet \hat{\mathbf{j}} = \hat{\mathbf{k}} \bullet \hat{\mathbf{k}} = 1 \qquad & \text{normalization} \\ \hat{\mathbf{i}} \bullet \hat{\mathbf{j}} = \hat{\mathbf{i}} \bullet \hat{\mathbf{k}} = \hat{\mathbf{j}} \bullet \hat{\mathbf{k}} = 0 \qquad & \text{orthogonality} \\ \mathbf{A} = a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} + a_z \hat{\mathbf{k}} \qquad & \text{completeness} \end{aligned}, \tag{1.9}$$

where $\mathbf{A}$ is any vector.

Continuing the mathematical analogy between spatial vectors and abstract vectors, we require that these same properties (at least conceptually) apply to quantum mechanical basis vectors. The completeness of the kets $|\pm\rangle$ implies that any quantum state vector $|\psi\rangle$ can be written as a linear combination of the two basis kets:

$$|\psi\rangle = a|+\rangle + b|-\rangle, \tag{1.10}$$

where a and b are complex scalar numbers multiplying each ket. This addition of two kets yields another ket in the same abstract space. The complex scalar can appear either before or after the ket without affecting the mathematical properties of the ket (i.e., $a|+\rangle = |+\rangle a$). Note that is customary to use the symbol ψ for a generic quantum state. You may have seen $\psi(x)$ used before as a wave function. However, the state vector $|\psi\rangle$ is not a wave function. It has no spatial dependence as a wave function does.

To discuss orthogonality and normalization (known together as **orthonormality**) we must first define scalar products as they apply to these new kets. As we said above, the machinery of quantum mechanics requires the use of complex numbers. You may have seen other fields of physics use complex numbers. For example, sinusoidal oscillations can be described using the complex exponential $e^{i\omega t}$ rather than $\cos(\omega t)$. However, in such cases, the complex numbers are not required, but are rather a convenience to make the mathematics easier. When using complex notation to describe classical vectors like electric and magnetic fields, dot products are changed slightly such that one of the vectors is complex conjugated. A similar approach is taken in quantum mechanics. The analog to the complex conjugated vector in classical physics is called a **bra** in the Dirac notation of quantum mechanics. Thus corresponding to the ket $|\psi\rangle$ is a bra, or bra vector, which is written as $\langle\psi|$. The bra $\langle\psi|$ is defined as

$$\langle\psi| = a^*\langle+| + b^*\langle-|, \tag{1.11}$$

where the basis bras $\langle+|$ and $\langle-|$ correspond to the basis kets $|+\rangle$ and $|-\rangle$, respectively, and the coefficients a and b have been complex conjugated.

The scalar product in quantum mechanics is defined as the combination of a bra and a ket, such as $\langle+|+\rangle$, which as you would guess is equal to one since we want the basis vectors to be normalized. Note that the bra and ket must occur in the proper order. Hence bra and ket make bracket – physics humor. The scalar product is often also called an **inner product** or a **projection** in quantum mechanics. Using this notation, orthogonality is expressed as $\langle+|-\rangle = 0$. Hence the properties of normalization, orthogonality, and completeness can be expressed in the case of a two-state quantum system as:

$$\left.\begin{aligned}\langle +|+\rangle &= 1\\ \langle -|-\rangle &= 1\end{aligned}\right\} \quad \text{normalization}$$

$$\left.\begin{aligned}\langle +|-\rangle &= 0\\ \langle -|+\rangle &= 0\end{aligned}\right\} \quad \text{orthogonality.} \tag{1.12}$$

$$|\psi\rangle = a|+\rangle + b|-\rangle \quad \text{completeness}$$

Note that a product of kets (e.g., $|+\rangle|+\rangle$) or a product of bras (e.g., $\langle +|\langle +|$) is meaningless in this new notation, while a product of a ket and a bra in the "wrong" order (e.g., $|+\rangle\langle +|$) has a meaning that we will define later. Equations (1.12) are sufficient to define how the basis kets behave mathematically. Note that the inner product is defined using a bra and a ket, though it is common to refer to the inner product of two kets, where it is understood that one is converted to a bra first. The order does matter as we will see shortly.

Using this new notation, we can learn a little more about general quantum states and derive some expressions that will be useful later. Consider the general state vector $|\psi\rangle = a|+\rangle + b|-\rangle$. Take the inner product of this ket with the bra $\langle +|$ and obtain

$$\begin{aligned}\langle +|\psi\rangle &= \langle +|a|+\rangle + \langle +|b|-\rangle\\ &= a\langle +|+\rangle + b\langle +|-\rangle\ ,\\ &= a\end{aligned} \tag{1.13}$$

using the property that scalars can be moved freely through bras or kets. Likewise, it can be shown that $\langle -|\psi\rangle = b$. Hence the coefficients multiplying the basis kets are simply the inner products or projections of the general state $|\psi\rangle$ along each basis ket, albeit in an abstract complex vector space, rather than the concrete three dimensional space of normal vectors. Using these results, we can rewrite the general state as

$$\begin{aligned}|\psi\rangle &= \langle +|\psi\rangle|+\rangle + \langle -|\psi\rangle|-\rangle\\ &= |+\rangle\langle +|\psi\rangle + |-\rangle\langle -|\psi\rangle\end{aligned}\ , \tag{1.14}$$

where the rearrangement of the second equation again uses the property that scalars (e.g., $\langle +|\psi\rangle$) can be moved through bras or kets.

For a general state vector $|\psi\rangle = a|+\rangle + b|-\rangle$ we defined the corresponding bra to be $\langle\psi| = a^*\langle +| + b^*\langle -|$. Thus, the inner product of the state $|\psi\rangle$ with the basis ket $|+\rangle$ taken in the reverse order compared to Eq. (1.13) yields

$$\begin{aligned}\langle\psi|+\rangle &= \langle +|a^*|+\rangle + \langle -|b^*|+\rangle\\ &= a^*\langle +|+\rangle + b^*\langle -|+\rangle\ .\\ &= a^*\end{aligned} \tag{1.15}$$

Thus we see that an inner product with the states reversed results in a complex conjugation of the inner product:

$$\langle +|\psi\rangle = \langle\psi|+\rangle^*. \tag{1.16}$$

This important property holds for any inner product.

Now we come to a new aspect of quantum vectors that differs from our use of vectors in classical mechanics. The rules of quantum mechanics (postulate 1) require that all state vectors describing a quantum system be normalized, not just the basis kets. This is clearly different from ordinary spatial vectors, where the length or magnitude means something. This new rule means that in the quantum mechanical state space only the direction is important. If we apply this normalization requirement to the general state $|\psi\rangle$, then we obtain

$$\begin{aligned} &\langle\psi|\psi\rangle = \left\{a^*\langle+| + b^*\langle-|\right\}\left\{a|+\rangle + b|-\rangle\right\} = 1 \\ &\Rightarrow a^*a\langle+|+\rangle + a^*b\langle+|-\rangle + b^*a\langle-|+\rangle + b^*b\langle-|-\rangle = 1 \\ &\Rightarrow a^*a + b^*b = 1 \\ &\Rightarrow |a|^2 + |b|^2 = 1 \end{aligned}, \tag{1.17}$$

or using the expressions for the coefficients obtained above,

$$\left|\langle+|\psi\rangle\right|^2 + \left|\langle-|\psi\rangle\right|^2 = 1. \tag{1.18}$$

Now comes the crucial element of quantum mechanics. We postulate that each term in the sum of Eq. (1.18) is equal to the **probability** that the quantum state described by the ket $|\psi\rangle$ is measured to be in the corresponding basis state. Thus

$$\mathcal{P}(+) = \left|\langle+|\psi\rangle\right|^2 \tag{1.19}$$

is the probability that the state $|\psi\rangle$ is found to be in the state $|+\rangle$ when a measurement of S_z is made, meaning that the result $S_z = +\hbar/2$ is obtained. Likewise,

$$\mathcal{P}(-) = \left|\langle-|\psi\rangle\right|^2 \tag{1.20}$$

is the probability that the measurement yields the result $S_z = -\hbar/2$. Since this simple system has only two possible measurement results, the probabilities must add up to one, which is why the rules of quantum mechanics require that state vectors be properly normalized before they are used in any calculation of probabilities. This is an application of the 4th postulate of quantum mechanics, which is repeated below.

Postulate 4
The probability of obtaining the eigenvalue a_n in a measurement of the observable A on the system in the state $|\psi\rangle$ is

$$\mathcal{P}(a_n) = \left|\langle a_n|\psi\rangle\right|^2,$$

where $|a_n\rangle$ is the eigenvector of A corresponding to the eigenvalue a_n.

This formulation of the 4th postulate uses some terms we have not defined yet. A simpler version employing the terms we know at this point would read:

> Postulate 4 (Spin 1/2 system)
> The probability of obtaining the value $\pm\hbar/2$ in a measurement of the observable S_z on a system in the state $|\psi\rangle$ is
>
> $$\mathcal{P}(\pm) = |\langle \pm|\psi\rangle|^2,$$
>
> where $|\pm\rangle$ is the basis ket of S_z corresponding to the result $\pm\hbar/2$.

The inner product, $\langle +|\psi\rangle$ for example, is called the **probability amplitude** or sometimes just the **amplitude**. Note that the convention is to put the input or initial state on the right and the output or final state on the left: $\langle out|in\rangle$, so one would read from right to left in describing a problem. Since the probability involves the complex square of the amplitude, and $\langle out|in\rangle = \langle in|out\rangle^*$, this convention is not critical for calculating probabilities. Nonetheless, it is the accepted practice and is important in situations where several amplitudes are combined.

Armed with these new quantum mechanical rules and tools, we now return to analyze the experiments discussed earlier.

1.3.1 Analysis of Experiment 1

In Experiment 1, the initial Stern-Gerlach analyzer prepared the system in the $|+\rangle$ state and the second analyzer measured this state to always be in the $|+\rangle$ state and never in the $|-\rangle$ state. Our new tools would predict the results of these measurements as

$$\begin{aligned} \mathcal{P}(+) &= |\langle +|+\rangle|^2 = 1 \\ \mathcal{P}(-) &= |\langle -|+\rangle|^2 = 0 \end{aligned}, \tag{1.21}$$

which agree with the experiment and are also consistent with the normalization and orthogonality properties of the basis vectors $|+\rangle$ and $|-\rangle$.

1.3.2 Analysis of Experiment 2

In Experiment 2, the initial Stern-Gerlach analyzer prepared the system in the $|+\rangle$ state and the second analyzer performed a measurement of the spin projection along the x-axis, finding 50% probabilities for each of the two possible states $|+\rangle_x$ and $|-\rangle_x$. In this case, we cannot predict the results of the measurements, since we do not yet have enough information about how the states $|+\rangle_x$ and $|-\rangle_x$ behave mathematically. Rather, we can use the results of the experiment to determine these states. Recalling that the experimental results would be the same if the first analyzer prepared the system to be in the $|-\rangle$ state, we have four results:

$$\begin{aligned} \mathcal{P}_1(+) &= |{}_x\langle +|+\rangle|^2 = \tfrac{1}{2} \\ \mathcal{P}_1(-) &= |{}_x\langle -|+\rangle|^2 = \tfrac{1}{2} \\ \mathcal{P}_2(+) &= |{}_x\langle +|-\rangle|^2 = \tfrac{1}{2} \\ \mathcal{P}_2(-) &= |{}_x\langle -|-\rangle|^2 = \tfrac{1}{2} \end{aligned} \tag{1.22}$$

Since the kets $|+\rangle$ and $|-\rangle$ form a basis, we know that the kets describing the S_x measurement, $|+\rangle_x$ and $|-\rangle_x$, can be written in terms of them. The $|\pm\rangle$ kets are referred to as the S_z basis, and allow us to write

$$\begin{aligned} |+\rangle_x &= a|+\rangle + b|-\rangle \\ |-\rangle_x &= c|+\rangle + d|-\rangle \end{aligned}, \tag{1.23}$$

where we wish to find the coefficients a, b, c, and d. Combining these with the experimental results (Eq. (1.22)), we obtain

$$\begin{aligned} |{}_x\langle +|+\rangle|^2 &= \left|\left\{a^*\langle +| + b^*\langle -|\right\}|+\rangle\right|^2 = \tfrac{1}{2} \\ &= \left|a^*\right|^2 = \tfrac{1}{2} \\ &= |a|^2 = \tfrac{1}{2} \end{aligned} \tag{1.24}$$

Likewise, one can show that $|b|^2 = |c|^2 = |d|^2 = \frac{1}{2}$. Since each coefficient is complex, it has an amplitude and phase. However, since the overall phase of a quantum state vector is not physically meaningful (problem 1.2), we can choose one coefficient of each vector to be real and positive without any loss of generality. This allows us to write the desired states as

$$\begin{aligned} |+\rangle_x &= \tfrac{1}{\sqrt{2}}\left[|+\rangle + e^{i\alpha}|-\rangle\right] \\ |-\rangle_x &= \tfrac{1}{\sqrt{2}}\left[|+\rangle + e^{i\beta}|-\rangle\right] \end{aligned}. \tag{1.25}$$

Note that these are already normalized since we used all of the experimental results, which reflect the fact that the probability for all possible results of an experiment must sum to one.

The $|\pm\rangle_x$ kets also form a basis, the S_x basis, since they correspond to the distinct results of a spin projection measurement. Thus we also must require that they are orthogonal to each other, which leads to

$$\begin{aligned} &{}_x\langle -|+\rangle_x = 0 \\ &\tfrac{1}{\sqrt{2}}\left[\langle +| + e^{-i\beta}\langle -|\right]\tfrac{1}{\sqrt{2}}\left[|+\rangle + e^{i\alpha}|-\rangle\right] = 0 \\ &\tfrac{1}{2}\left[1 + e^{i(\alpha-\beta)}\right] = 0 \\ &e^{i(\alpha-\beta)} = -1 \\ &e^{i\alpha} = -e^{i\beta} \end{aligned} \quad . \tag{1.26}$$

where the complex conjugation of the second coefficient of the ${}_x\langle -|$ bra should be noted. At this point, we are free to choose the value of the phase α since there is no more information that can be used to constrain it. This freedom comes from the fact that we have required only that the x-axis be perpendicular to the z-axis, which limits it only to a

plane rather than a single line. We follow convention here and choose the phase $\alpha = 0$. Thus we can express the S_x basis kets in terms of the S_z basis kets as

$$\begin{aligned} |+\rangle_x &= \tfrac{1}{\sqrt{2}}\left[|+\rangle + |-\rangle\right] \\ |-\rangle_x &= \tfrac{1}{\sqrt{2}}\left[|+\rangle - |-\rangle\right] \end{aligned}. \tag{1.27}$$

We generally use the S_z basis, but could use any basis we choose. If we were to use the S_x basis, then we could write the $|\pm\rangle$ kets in terms of the $|\pm\rangle_x$ kets. This can be done by simply solving Eqs. (1.27) for the $|\pm\rangle$ kets, yielding

$$\begin{aligned} |+\rangle &= \tfrac{1}{\sqrt{2}}\left[|+\rangle_x + |-\rangle_x\right] \\ |-\rangle &= \tfrac{1}{\sqrt{2}}\left[|+\rangle_x - |-\rangle_x\right] \end{aligned}. \tag{1.28}$$

In terms of the measurement performed in Experiment 2, these equations tells us that the $|+\rangle$ state is a combination of the states $|+\rangle_x$ and $|-\rangle_x$. The coefficients tell us that there is a 50% probability for measuring the spin projection to be up along the x-axis, and likewise for the down possibility, which is what was measured. A combination of states is usually referred to as a **superposition** state.

1.3.3 Superposition states

To understand the importance of a quantum mechanical superposition of states, consider the $|+\rangle_x$ state found above. This state can be written in terms of the S_z basis states as

$$|+\rangle_x = \tfrac{1}{\sqrt{2}}\left[|+\rangle + |-\rangle\right]. \tag{1.29}$$

If we measure the spin projection along the x-axis for this state, then we record only the result $S_x = +\hbar/2$ (Experiment 1 with both analyzers along the x-axis). If we measure the spin projection along the orthogonal z-axis, then we record the two results $S_z = \pm\hbar/2$ with 50% probability each (Experiment 2 with the first and second analyzers along the x– and z-axes, respectively). Based upon these results, one might be tempted to consider the $|+\rangle_x$ state as describing a beam that contains a mixture of atoms with 50% of the atoms in the $|+\rangle$ state and 50% in the $|-\rangle$ state.

Let's now carefully examine the results of experiments on this proposed mixture beam. If we measure the spin projection along the z-axis, then each atom in the $|+\rangle$ state yields the result $S_z = +\hbar/2$ with 100% certainty and each atom in the $|-\rangle$ state yields the result $S_z = -\hbar/2$ with 100% certainty. The net result is that 50% of the atoms yield $S_z = +\hbar/2$ and 50% yield $S_z = -\hbar/2$. This is exactly the same result as that obtained with all atoms in the $|+\rangle_x$ state. If we instead measure the spin projection along the x-axis, then each atom in the $|+\rangle$ state yields the two results $S_x = \pm\hbar/2$ with 50% probability each (Experiment 2 with the first and second analyzers along the z- and x-axes, respectively). The atoms in the $|-\rangle$ state yield the same results. The net result is that 50% of the atoms yield $S_x = +\hbar/2$ and 50% yield $S_x = -\hbar/2$. This is in stark contrast to

the results of Experiment 1, which tells us that once we have measured the state to be $|+\rangle_x$, then subsequent measurements yield $S_x = +\hbar/2$ with certainty.

Hence we must conclude that the system described by the $|+\rangle_x$ state is not the same as a mixture of atoms in the $|+\rangle$ and $|-\rangle$ states. This means that each atom in the beam is in a state that itself is a combination of the $|+\rangle$ and $|-\rangle$ states. A superposition state is often called a **coherent superposition** since the relative phase of the two terms is important. For example, if the beam were in the $|-\rangle_x$ state, then there would be a relative minus sign between the two coefficients, which would result in a $S_x = -\hbar/2$ measurement but would not affect the S_z measurement.

We will not have any further need to speak of mixtures, so any combination of states is a superposition. Note that we cannot even write down a ket for the mixture case. So, if someone gives you a quantum state written as a ket, then it must be a superposition and not a mixture. The random option in the SPINS program produces a mixture, while the unknown states are all superpositions.

1.4 Matrix notation

Up to this point we have defined kets mathematically in terms of their inner products with other kets. Thus in the general case we can write a ket as

$$|\psi\rangle = \langle +|\psi\rangle|+\rangle + \langle -|\psi\rangle|-\rangle, \tag{1.30}$$

or in a specific case, we can write

$$\begin{aligned} |+\rangle_x &= \langle +|+\rangle_x|+\rangle + \langle -|+\rangle_x|-\rangle \\ &= \tfrac{1}{\sqrt{2}}|+\rangle + \tfrac{1}{\sqrt{2}}|-\rangle \end{aligned} \quad . \tag{1.31}$$

In both of these cases, we have chosen to write the kets in terms of the $|+\rangle$ and $|-\rangle$ basis kets. If we agree on that choice of basis as a convention, then we really only need to specify the coefficients, and we can simply the notation by merely using those numbers. Thus, we can represent a ket as a column vector containing the two coefficients multiplying each basis ket. For example, we represent $|+\rangle_x$ as

$$|+\rangle_x \doteq \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \tag{1.32}$$

where we have used the symbol $\doteq$ to signify "is represented by", and it is understood that we are using the $|+\rangle$ and $|-\rangle$ basis or the S_z basis. We cannot say that the ket equals the column vector, since the ket is an abstract vector in the state space and the column vector is just two complex numbers. We also need to have a convention for the ordering of the amplitudes in the column vector. The standard convention is to put the spin up amplitude first (at top). Thus the representation of the $|-\rangle_x$ state (Eq. (1.28)) is

$$|-\rangle_x \doteq \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}. \tag{1.33}$$

Using this convention, it should be clear that the basis kets themselves can be written as

$$\begin{aligned} |+\rangle &\doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |-\rangle &\doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}. \tag{1.34}$$

This expression of a ket simply as the coefficients multiplying the basis kets is referred to as a **representation**. Since we have assumed the S_z basis kets, this is called the S_z representation. It is always true that basis kets have the simple form shown in Eq. (1.34) when written in their own representation. A general ket $|\psi\rangle$ is written as

$$|\psi\rangle \doteq \begin{pmatrix} \langle +|\psi\rangle \\ \langle -|\psi\rangle \end{pmatrix}. \tag{1.35}$$

This use of matrices simplifies the mathematics of bras and kets. The advantage is not so evident for the simple 2 dimensional state space of spin 1/2 systems, but is very evident for larger dimensional problems. This notation is indispensable when using computers to calculate quantum mechanical results. For example, the SPINS program employs matrix notation to simulate the Stern-Gerlach experiments.

We saw earlier that the coefficients of a bra are the complex conjugates of the coefficients of the corresponding ket. We also know that an inner product of a bra and a ket yields a single complex number. In order for the matrix rules of multiplication to be used, a bra must be represented by a row vector, with the entries being the coefficients ordered in the same sense as for the ket. For example, if we use the general ket

$$|\psi\rangle = a|+\rangle + b|-\rangle, \tag{1.36}$$

which can be represented as

$$|\psi\rangle \doteq \begin{pmatrix} a \\ b \end{pmatrix}, \tag{1.37}$$

then the corresponding bra

$$\langle\psi| = a^*\langle +| + b^*\langle -| \tag{1.38}$$

can be represented as a row vector as

$$\langle\psi| \doteq \begin{pmatrix} a^* & b^* \end{pmatrix}. \tag{1.39}$$

The rules of matrix algebra can then be applied to find an inner product. For example,

$$\begin{aligned} \langle\psi|\psi\rangle &= \begin{pmatrix} a^* & b^* \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}. \\ &= |a|^2 + |b|^2 \end{aligned} \tag{1.40}$$

So a bra is represented by a row vector that is the complex conjugate and transpose of the column vector representing the corresponding ket.

To get some practice using this new matrix notation, and to learn some more about the spin 1/2 system, consider Experiment 2 in the case where the second Stern-Gerlach analyzer is aligned along the y-axis. We said before that the results will be the same as in the case shown in Fig. 1.2. Thus we have

$$\begin{aligned} \mathcal{P}_1(+) &= \left|{}_y\langle +|+\rangle\right|^2 = \tfrac{1}{2} \\ \mathcal{P}_1(-) &= \left|{}_y\langle -|+\rangle\right|^2 = \tfrac{1}{2} \\ \mathcal{P}_2(+) &= \left|{}_y\langle +|-\rangle\right|^2 = \tfrac{1}{2} \\ \mathcal{P}_2(-) &= \left|{}_y\langle -|-\rangle\right|^2 = \tfrac{1}{2} \end{aligned} \quad . \tag{1.41}$$

This allows us to determine the kets $|\pm\rangle_y$ corresponding to spin projection up and down along the y-axis. The argument and calculation proceeds exactly as it did earlier for the $|\pm\rangle_x$ states up until the point where we arbitrarily choose the phase α to be zero. Having done that for the $|\pm\rangle_x$ states, we are no longer free to make that same choice for the $|\pm\rangle_y$ states. Thus we can write the $|\pm\rangle_y$ states as

$$\begin{aligned} |+\rangle_y &= \tfrac{1}{\sqrt{2}}\left[|+\rangle + e^{i\alpha}|-\rangle\right] \doteq \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ e^{i\alpha} \end{pmatrix} \\ |-\rangle_y &= \tfrac{1}{\sqrt{2}}\left[|+\rangle - e^{i\alpha}|-\rangle\right] \doteq \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -e^{i\alpha} \end{pmatrix} \end{aligned} \tag{1.42}$$

To determine the phase α, we can use some more information at our disposal. Experiment 2 could be performed with the first Stern-Gerlach analyzer along the x-axis and the second along the y-axis. Again the results would be identical (50% at each output port), yielding

$$\mathcal{P}(+) = \left|{}_y\langle +|+\rangle_x\right|^2 = \tfrac{1}{2} \tag{1.43}$$

as one of the measured quantities. Now use matrix algebra to calculate this:

$$\begin{aligned} {}_y\langle +|+\rangle_x &= \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & e^{-i\alpha} \end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \tfrac{1}{2}\left(1 + e^{-i\alpha}\right) \\ \left|{}_y\langle +|+\rangle_x\right|^2 &= \tfrac{1}{2}\left(1 + e^{-i\alpha}\right)\tfrac{1}{2}\left(1 + e^{i\alpha}\right) \\ &= \tfrac{1}{4}\left(1 + e^{i\alpha} + e^{-i\alpha} + 1\right) \\ &= \tfrac{1}{2}(1 + \cos\alpha) = \tfrac{1}{2} \end{aligned} \tag{1.44}$$

This result requires that $\cos\alpha = 0$, or that $\alpha = \pm\pi/2$. The two choices for the phase correspond to the two possibilities for the direction of the y-axis relative to the already determined x- and z-axes. The choice $\alpha = +\pi/2$ can be shown to correspond to a right

handed coordinate system. Since that is the standard convention, we will choose that phase. We thus represent the $|\pm\rangle_y$ kets as

$$|+\rangle_y \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}$$
$$|-\rangle_y \doteq \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix} \tag{1.45}$$

Note that the imaginary components of these kets are required. They are not merely a mathematical convenience as one sees in classical mechanics. In general quantum mechanical state vectors have complex coefficients. But this does not imply any complexity in the results of physical measurements, since we always have to calculate a probability which involves a complex square, so all quantum mechanics predictions are real.

1.5 General Quantum Systems

The machinery we have developed for spin 1/2 systems can be generalized to other quantum systems. For example, if we have an observable *A* that yields quantized measurement results a_n for some finite range of *n*, then we could generalize the schematic depiction of a Stern-Gerlach measurement as shown in Fig. 1.8. The observable *A* labels the measurement device and the possible results label the output ports. The basis kets corresponding to the results a_n are then $|a_n\rangle$. The mathematical rules about kets can then be written in this general case as

$$\langle a_i | a_j \rangle = \delta_{ij} \qquad \text{orthonormality}$$
$$|\psi\rangle = \sum_i \langle a_i | \psi \rangle | a_i \rangle \qquad \text{completeness}$$
$$\langle \psi | \phi \rangle = \langle \phi | \psi \rangle^* \qquad \text{amplitude conjugation}$$

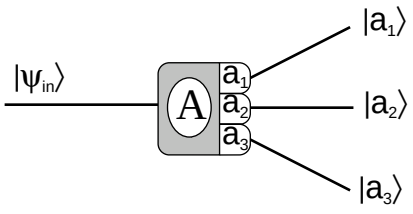

Figure 1.8. Generic depiction of quantum mechanical measurement of observable *A*.

Problems

1.1 Show that a change in the overall phase of a quantum state vector does not change the probability of obtaining a particular result in a measurement. To do this, consider how the probability is affected by changing the state $|\psi\rangle$ to the state $e^{i\delta}|\psi\rangle$.

1.2 Consider a quantum system described by a basis $|a_1\rangle$, $|a_2\rangle$, and $|a_3\rangle$. The system is initially in a state

$$|\psi_i\rangle = \frac{i}{\sqrt{3}}|a_1\rangle + \sqrt{\frac{2}{3}}|a_2\rangle.$$

Find the probability that the system is measured to be in the final state

$$|\psi_f\rangle = \frac{1+i}{\sqrt{3}}|a_1\rangle + \frac{1}{\sqrt{6}}|a_2\rangle + \frac{1}{\sqrt{6}}|a_3\rangle.$$

1.3 Normalize the following state vectors:

a) $|\psi\rangle = 3|+\rangle + 4|-\rangle$

b) $|\phi\rangle = |+\rangle + 2i|-\rangle$

c) $|\gamma\rangle = 3|+\rangle - e^{i\pi/3}|-\rangle$

1.4 (Townsend 1.7) A beam of spin 1/2 particles is sent through a series of three Stern-Gerlach (SG) measuring devices. The first SG device is aligned along the z-axis and transmits particles with $S_z = \hbar/2$ and blocks particles with $S_z = -\hbar/2$. The second device is aligned along the **n** direction and transmits particles with $S_n = \hbar/2$ and blocks particles with $S_n = -\hbar/2$, where the direction **n** makes an angle θ in the *x*-*z* plane with respect to the *z* axis. Thus particles after passage through this second device are in the state $|+\rangle_n = \cos(\theta/2)|+\rangle + \sin(\theta/2)|-\rangle$. A third SG device is aligned along the z-axis and transmits particles with $S_z = -\hbar/2$ and blocks particles with $S_z = \hbar/2$.

a) What fraction of the particles transmitted through the first SG device will survive the third measurement?

b) How must the angle θ of the second SG device be oriented so as to maximize the number of particles that are transmitted by the final SG device? What fraction of the particles survive the third measurement for this value of θ?

c) What fraction of the particles survive the last measurement if the second SG device is simply removed from the experiment?

1.5 Consider the three quantum states:

$$|\psi_1\rangle = \frac{1}{\sqrt{3}}|+\rangle + i\frac{\sqrt{2}}{\sqrt{3}}|-\rangle$$

$$|\psi_2\rangle = \frac{1}{\sqrt{5}}|+\rangle - \frac{2}{\sqrt{5}}|-\rangle$$

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{e^{i\pi/4}}{\sqrt{2}}|-\rangle$$

Use bra and ket notation (not matrix notation) to solve the following problems. Note that $\langle +|+\rangle = 1$, $\langle -|-\rangle = 1$, and $\langle +|-\rangle = 0$.

a) For each of the $|\psi_i\rangle$ above, find the normalized vector $|\phi_i\rangle$ that is orthogonal to it.

b) Calculate the inner products $\langle \psi_i | \psi_j \rangle$ for i and j = 1, 2, 3.

Chapter 2 OPERATORS AND MEASUREMENT

Up until now we have used the results of experiments, both the measured quantities and their probabilities, to deduce mathematical descriptions of the spin 1/2 system in terms of the basis kets of the spin projection observables. For a complete theory, we want to be able to predict the possible values of the measured quantities and the probabilities of measuring them in any specific experiment. In order to do this, we need to learn about the operators of quantum mechanics.

2.1 Operators

In quantum mechanics, physical observables are represented by mathematical operators (postulate 2, repeated below) in the same sense that quantum states are represented by mathematical vectors or kets (postulate 1). An operator is a mathematical object that acts or operates on a ket and transforms it into a new ket, for example $A|\psi\rangle = |\phi\rangle$. However, there are special kets that are not changed by the operation of a particular operator, aside from a multiplicative constant, which we saw before does not change anything measurable about the state. An example of this would be $A|\psi\rangle = a|\psi\rangle$. Such kets are known as **eigenvectors** of the operator A and the multiplicative constants are known as the **eigenvalues** of the operator. These are important because postulate 3 (repeated below) postulates that the only possible result of a measurement of a physical observable is one of the eigenvalues of the corresponding operator.

Postulate 2
A physical observable is described mathematically by an operator A that acts on kets.

Postulate 3
The only possible result of a measurement of an observable is one of the eigenvalues a_n of the corresponding operator A.

Given these postulates and the experimental results of Chap. 1, we can write the eigenvalue equations for the S_z operator:

$$\begin{aligned} S_z|+\rangle &= \frac{\hbar}{2}|+\rangle \\ S_z|-\rangle &= -\frac{\hbar}{2}|-\rangle \end{aligned}, \tag{2.1}$$

which mean that $+\hbar/2$ is the eigenvalue of S_z corresponding to the eigenvector $|+\rangle$ and $-\hbar/2$ is the eigenvalue of S_z corresponding to the eigenvector $|-\rangle$. (Though it is common in some texts, in this text we do not use different symbols for an observable and its corresponding operator; it is generally obvious from the context which is being used.) Equations (2.1) are sufficient to define how the S_z operator acts on kets. However, it is

useful to use matrix notation to represent operators in the same way as we did earlier with bras and kets. In order for Eqs. (2.1) to be satisfied using matrix algebra, the operator S_z must be represented by a 2×2 matrix. The eigenvalue equations (Eqs. (2.1)) provide sufficient information to determine this matrix. Let the matrix representing the operator S_z have the form

$$S_z \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{2.2}$$

and write the eigenvalue equations in matrix form:

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} &= +\frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} &= -\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}, \tag{2.3}$$

where we are still using the convention that the $|\pm\rangle$ kets are used as the basis for the representation. It is crucial that the rows and columns of the operator matrix are ordered in the same manner as used for the ket column vectors; anything else would amount to nonsense. Multiplying these out yields

$$\begin{aligned} \begin{pmatrix} a \\ c \end{pmatrix} &= +\frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} b \\ d \end{pmatrix} &= -\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}, \tag{2.4}$$

which results in

$$\begin{aligned} a &= +\frac{\hbar}{2} & b &= 0 \\ c &= 0 & d &= -\frac{\hbar}{2} \end{aligned}. \tag{2.5}$$

Thus the matrix representation of the operator S_z is

$$\begin{aligned} S_z &\doteq \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \\ &\doteq \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}. \tag{2.6}$$

Note two important features of this matrix: (1) it only has diagonal elements and (2) the diagonal elements are the eigenvalues of the operator, ordered in the same manner as the corresponding eigenvectors. In this example, the basis used for the matrix representation is that formed by the eigenvectors of the operator S_z. That the matrix representation of the operator in this case is a diagonal matrix is a necessary and general result of linear

algebra that will prove valuable as we study quantum mechanics. In simple terms, we can say that an operator is always diagonal in its own basis.

Now consider how matrix representation works in general. Consider a general operator A (still in the two-dimensional spin 1/2 system), which we represent by the matrix

$$A \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{2.7}$$

in the S_z basis. The operation of A on the basis ket $|+\rangle$ yields

$$A|+\rangle \doteq \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}. \tag{2.8}$$

The inner product of this new ket with the ket $|+\rangle$ results in

$$\langle +|A|+\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ c \end{pmatrix} = a, \tag{2.9}$$

which serves to isolate one of the elements of the matrix. Hence an individual element such as $\langle +|A|+\rangle$ is generally referred to as a **matrix element**. Similarly, all four elements of the matrix representation of A can be determined in this manner, with the final result

$$A \doteq \begin{pmatrix} \langle +|A|+\rangle & \langle +|A|-\rangle \\ \langle -|A|+\rangle & \langle -|A|-\rangle \end{pmatrix}. \tag{2.10}$$

In a more general problem with more than 2 dimensions in the complex vector space, one would write the matrix as

$$A \doteq \begin{pmatrix} A_{11} & A_{12} & A_{13} & \cdots \\ A_{21} & A_{22} & A_{23} & \cdots \\ A_{31} & A_{32} & A_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2.11}$$

where

$$A_{ij} = \langle i|A|j\rangle \tag{2.12}$$

and the basis is assumed to be the states labeled $|i\rangle$, with the subscripts i and j labeling the rows and columns respectively.

In the case of the operator S_z above, we used the experimental results and the eigenvalue equations to find the matrix representation of the operator. It is more common to work the other way. That is, one is given the matrix representation of an operator and is asked to find the possible results of a measurement of the corresponding observable. According to the 3rd postulate, the possible results are the eignevalues of the operator, and the eigenvectors are the quantum states representing them. In the case of a general operator A in a two-state system, the eigenvalue equation would be

$$A|a_i\rangle = a_i|a_i\rangle, \tag{2.13}$$

where we have labeled the eigenvalues $a_{\rm i}$ and have labeled the eigenvectors with the corresponding eigenvalues. In matrix notation, this equation is

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} c_{1,i} \\ c_{2,i} \end{pmatrix} = a_i \begin{pmatrix} c_{1,i} \\ c_{2,i} \end{pmatrix}, \tag{2.14}$$

where $c_{1,\rm i}$ and $c_{2,\rm i}$ are the unknown coefficients of the eigenvectors. This matrix equation yields the set of homogeneous equations

$$\begin{aligned} (A_{11} - a_i)c_{1,i} + A_{12}c_{2,i} &= 0 \\ (A_{21} - a_i)c_{1,i} + A_{22}c_{2,i} &= 0 \end{aligned}. \tag{2.15}$$

This set of homogeneous equations will have solutions for the unknowns $c_{1,\rm i}$ and $c_{2,\rm i}$ only if the determinant of the coefficients vanishes:

$$\begin{vmatrix} A_{11} - a_i & A_{12} \\ A_{21} & A_{22} - a_i \end{vmatrix} = 0. \tag{2.16}$$

It is common notation to use the symbol λ for the eigenvalues, in which case this equation can be written as

$$\det(A - \lambda I) = 0, \tag{2.17}$$

where I is the identity matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.18}$$

Equation (2.17) is known as the **secular** or **characteristic** equation. It is a second order equation in the parameter λ and the two roots are identified as the two eigenvalues a_1 and a_2 we are trying to find. Once those eigenvalues are found, they are individually substituted back into Eqs. (2.15), which are solved to find the coefficients of the corresponding eigenvector.

As an example, assume that we know the matrix representation for the operator $S_{\rm y}$ (e.g., from Problem 2.1) and we wish to find the eigenvalues and eigenvectors. The general eigenvalue equation can be written as

$$S_y|\psi\rangle = \lambda|\psi\rangle \tag{2.19}$$

and the possible eigenvalues λ are then found using the secular equation

$$\det|S_y - \lambda I| = 0. \tag{2.20}$$

The matrix for the operator $S_{\rm y}$ is

$$S_y \doteq \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{2.21}$$

which leads to the secular equation

$$\begin{vmatrix} -\lambda & -i\frac{\hbar}{2} \\ i\frac{\hbar}{2} & -\lambda \end{vmatrix} = 0. \tag{2.22}$$

Solving this equation yields

$$\begin{aligned} \lambda^2 + i^2\left(\frac{\hbar}{2}\right)^2 &= 0 \\ \lambda^2 - \left(\frac{\hbar}{2}\right)^2 &= 0 \\ \lambda^2 &= \left(\frac{\hbar}{2}\right)^2 \\ \lambda &= \pm\frac{\hbar}{2} \end{aligned}, \tag{2.23}$$

which was to be expected, since we know that the only possible results of a measurement of the spin projection along any axis are $\pm\hbar/2$. As before, we label the eigenvectors $|\pm\rangle_y$. The eigenvalue equation for the positive eigenvalue can then written as

$$S_y|+\rangle_y = +\frac{\hbar}{2}|+\rangle_y \tag{2.24}$$

or in matrix notation as

$$\frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = +\frac{\hbar}{2}\begin{pmatrix} a \\ b \end{pmatrix}, \tag{2.25}$$

where we must solve for *a* and *b* to determine the eigenvector. Multiplying through and cancelling the common factor yields

$$\begin{pmatrix} -ib \\ ia \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}. \tag{2.26}$$

This results in two equations, but they are not linearly independent. So to solve for both coefficients, we recall that the ket must be normalized. Thus we have 2 equations to solve:

$$\begin{aligned} b &= ia \\ |a|^2 + |b|^2 &= 1 \end{aligned}. \tag{2.27}$$

Solving these yields

$$\begin{aligned} |a|^2 + |ia|^2 &= 1 \\ |a|^2 &= \tfrac{1}{2} \end{aligned}. \tag{2.28}$$

Again we follow the convention of choosing the first coefficient to be real and positive, resulting in

$$\begin{aligned} a &= \tfrac{1}{\sqrt{2}} \\ b &= i\tfrac{1}{\sqrt{2}} \end{aligned}. \tag{2.29}$$

Thus the eigenvector corresponding to the positive eigenvalue is

$$|+\rangle_y \doteq \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}. \tag{2.30}$$

Likewise, one can find the eigenvector for the negative eigenvalue to be

$$|-\rangle_y \doteq \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \tag{2.31}$$

This procedure of finding the eigenvalues and eigenvectors of a matrix is known as **diagonalization** of the matrix. However, we stop short of the mathematical exercise of finding the matrix that transforms the original matrix to its diagonal form. In this case, the S_y matrix is not diagonal, while the S_z matrix is diagonal, since we are using the S_z basis. It is common practice to use the S_z basis as the default basis, so you can assume that is the case unless you are told otherwise.

2.1.1 Spin Projection in General Direction

We now know the eigenvalues and eigenvectors of each of the three operators corresponding to the spin projections along the axes of a coordinate system. It is also useful to discuss the operator for spin projection along a general direction $\hat{\mathbf{n}}$, which is specified by the polar and azimuthal angles θ and ϕ as shown in Fig. 2.1. The unit vector can be written as

$$\hat{\mathbf{n}} = \hat{\mathbf{i}}\sin\theta\cos\phi + \hat{\mathbf{j}}\sin\theta\sin\phi + \hat{\mathbf{k}}\cos\theta. \tag{2.32}$$

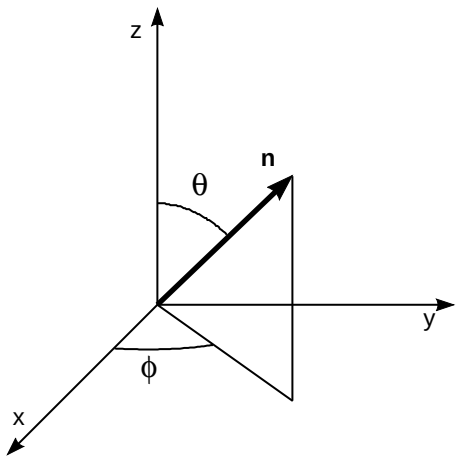

Figure 2.1. General direction along which to measure spin projection.

The spin projection along this direction is given by

$$\begin{aligned} S_n &= \mathbf{S} \bullet \hat{\mathbf{n}} \\ &= S_x \sin\theta\cos\phi + S_y \sin\theta\sin\phi + S_z \cos\theta \end{aligned} \tag{2.33}$$

and has a matrix representation

$$S_n \doteq \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}. \tag{2.34}$$

Following the same diagonalization procedure used above for the S_y matrix, we find that the eigenvalues of S_n are $\pm\hbar/2$ as we expect. The eigenvectors are

$$\begin{aligned} |+\rangle_n &= \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{i\phi}|-\rangle \\ |-\rangle_n &= \sin\frac{\theta}{2}|+\rangle - \cos\frac{\theta}{2}e^{i\phi}|-\rangle \end{aligned}, \tag{2.35}$$

where we again use the convention of choosing the first coefficient to be real and positive. It is important to point out that the $|+\rangle_n$ eigenstate (or equivalently the $|-\rangle_n$ eigenstate) can be used to represent any possible ket in a spin 1/2 system, if one allows for all possible angles $0 \le \theta < \pi$ and $0 \le \phi < 2\pi$. We generally write the most general state as $|\psi\rangle = a|+\rangle + b|-\rangle$, where a and b are complex. Requiring that the state be normalized and using the freedom to choose the first coefficient real and positive reduces this to

$$|\psi\rangle = |a||+\rangle + \sqrt{1-|a|^2}e^{i\phi}|-\rangle. \tag{2.36}$$

If we change the parametrization of $|a|$ to $\cos(\theta/2)$, we see that $|+\rangle_n$ is equivalent to the most general state $|\psi\rangle$. This correspondence between the $|+\rangle_n$ eigenstate and the most general state is only valid in a two-state system such as spin 1/2. In systems with more dimensionality, it does not hold since more parameters are needed to specify the most general state than are afforded by the angles θ and ϕ.

2.1.2 Hermitian Operators

So far we have only discussed how operators act upon kets. An operator acts on a bra from the right side

$$\langle\xi| = \langle\psi|A \tag{2.37}$$

and the result is another bra. If the operator A acting on the ket $|\psi\rangle$ yields the ket $|\phi\rangle = A|\psi\rangle$, then the bra $\langle\xi|$ defined above is not the bra corresponding to the ket $|\phi\rangle$. Rather the bra $\langle\phi|$ is found by defining a new operator A^+ that obeys

$$\langle\phi| = \langle\psi|A^+. \tag{2.38}$$

This new operator is called the **Hermitian adjoint** of the operator A. We can learn something about such an operator by calculating its matrix elements

$$\begin{aligned} \langle\phi|\beta\rangle &= \langle\beta|\phi\rangle^* \\ \left[\langle\psi|A^+\right]|\beta\rangle &= \langle\beta|\left[A|\psi\rangle\right]^*, \\ \langle\psi|A^+|\beta\rangle &= \langle\beta|A|\psi\rangle^* \end{aligned} \tag{2.39}$$

which tells us that the matrix representing the Hermitian adjoint A^+ is found by transposing and complex conjugating the matrix representing A. This is consistent with the definition of Hermitian adjoint used in matrix algebra.

An operator A is said to be **Hermitian** if it is equal to its Hermitian adjoint A^+. In quantum mechanics all operators that correspond to physical observables are Hermitian. This is important in light of two features of Hermitian matrices. (1) Hermitian matrices have real eigenvalues, which ensures that results of measurements are always real. (2) The eigenvectors of a Hermitian matrix comprise a complete set of basis states, which ensures that we can use the eigenvectors of any observable as a valid basis.

2.1.3 Projection Operators

For this simple spin 1/2 system, we now know four operators: S_x, S_y, S_z, and S_n. Let's look for some other operators. Consider the ket $|\psi\rangle$ written in terms of its coefficients in the S_z basis

$$\begin{aligned} |\psi\rangle &= \langle+|\psi\rangle|+\rangle \; + \; \langle-|\psi\rangle|-\rangle \\ &= |+\rangle\langle+|\psi\rangle \; + \; |-\rangle\langle-|\psi\rangle. \\ &= \left[|+\rangle\langle+| \; + \; |-\rangle\langle-|\right]|\psi\rangle \end{aligned} \tag{2.40}$$

The term in brackets is an operator since it acts on a ket to produce another ket. Since the result is the same as the original ket, the operator must be the identity operator. This relationship is often written as

$$|+\rangle\langle+| \; + \; |-\rangle\langle-| \; = \; 1, \tag{2.41}$$

and is generally known as the **completeness relation** or **closure**. It expresses the fact that the basis states $|\pm\rangle$ comprise a complete set of states, meaning any arbitrary ket can be written in terms of them. Note that each piece of this new operator is a product of a ket and a bra, but in the opposite order compared to the inner product defined earlier. This new object is known as an **outer product**. To make it obvious that such an object is an operator, it is useful to express this new operator in matrix notation using standard rules of matrix multiplication. Thus we get

$$
\begin{aligned}
|+\rangle\langle+| \; + \; |-\rangle\langle-| &\doteq \begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0 & 1\end{pmatrix} \\
&\doteq \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} \quad , \\
&\doteq \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}
\end{aligned}
\tag{2.42}
$$

which makes it clear that this operator is the identity operator.

Now consider breaking this new operator into its constituent pieces, of which there are only two in the spin 1/2 case. These operators are called **projection** operators, and for spin 1/2 are given by

$$
\begin{aligned}
P_+ &= |+\rangle\langle+| \doteq \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} \\
P_- &= |-\rangle\langle-| \doteq \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}
\end{aligned}
\; . \tag{2.43}
$$

In terms of these new operators the completeness relation can also be written as

$$P_+ + P_- = 1. \tag{2.44}$$

When a projection operator for a particular eigenstate acts on a state $|\psi\rangle$, it produces a new ket that is aligned along the eigenstate and has a magnitude equal to the amplitude (including the phase) for the state $|\psi\rangle$ to be in that eigenstate. For example,

$$
\begin{aligned}
P_+|\psi\rangle &= |+\rangle\langle+|\psi\rangle = \big(\langle+|\psi\rangle\big)|+\rangle \\
P_-|\psi\rangle &= |-\rangle\langle-|\psi\rangle = \big(\langle-|\psi\rangle\big)|-\rangle
\end{aligned}
\; . \tag{2.45}
$$

Note also that a projector acting on its own eigenstate results in the eigenstate, and a projector acting on an orthogonal state results in zero:

$$
\begin{aligned}
P_+|+\rangle &= |+\rangle\langle+|+\rangle = |+\rangle \\
P_-|+\rangle &= |-\rangle\langle-|+\rangle = 0
\end{aligned}
\; . \tag{2.46}
$$

Since the projection operator produces the probability amplitude, we expect that it must be intimately tied to measurement in quantum mechanics.

We saw in Chap. 1 that the probability of a measurement is given by the square of the inner product of initial and final states (postulate 4). Using our new projection operators, we can rewrite the probability as

$$
\begin{aligned}
\mathcal{P}(+) &= \big|\langle+|\psi\rangle\big|^2 \\
&= \langle+|\psi\rangle^*\langle+|\psi\rangle \\
&= \langle\psi|+\rangle\langle+|\psi\rangle \\
&= \langle\psi|P_+|\psi\rangle
\end{aligned}
\; . \tag{2.47}
$$

Thus we can say that the probability can be calculated as a matrix element of the projection operator, using the input state and the projector corresponding to the result.

The other important aspect of quantum measurement that we learned in Chap. 1 was that the measurement perturbs the system. That is, if an input state $|\psi\rangle$ is measured to have $S_z = +\hbar/2$, then the output state is no longer $|\psi\rangle$, but is changed to $|+\rangle$. We saw above that the projection operator does this operation for us, with a multiplicative constant of the probability amplitude. Thus if we divide by this amplitude, which is the square root of the probability, then we can describe the abrupt change of the input state as

$$|\psi'\rangle = \frac{P_+|\psi\rangle}{\sqrt{\langle\psi|P_+|\psi\rangle}} = |+\rangle, \tag{2.48}$$

where $|\psi'\rangle$ is the output state. This effect is described by the fifth postulate, which is repeated below and is often referred to as the **projection postulate**.

The projection postulate is at the heart of quantum measurement. This effect is often referred to as the **collapse** (or **reduction** or **projection**) of the quantum state vector. We do not attempt to explain the mechanism for this collapse. This effect has been the source of much of the controversy surrounding quantum mechanics. It clearly states that quantum measurements cannot be made without disturbing the system (except in the case where the input state is the same as the output state), in sharp contrast to classical measurements. This collapse of the quantum state makes quantum mechanics irreversible, again in contrast to classical mechanics. We will have more to say about the collapse later.

Postulate 5

After a measurement of A that yields the result a_n, the quantum system is in a new state that is the normalized projection of the original system ket onto the ket (or kets) corresponding to the result of the measurement:

$$|\psi'\rangle = \frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}}.$$

2.1.4 Analysis of Experiments 3 and 4

We can now return to Experiments 3 and 4 from Chap. 1 and analyze them with these new tools. Recall that Experiment 3 is the same as Experiment 4a, and Experiments 4a and 4b are similar in that they each use only one of the output ports of the second Stern-Gerlach analyzer as input to the third. Figure 2.2 depicts these experiments again, with Fig. 2.2(a) showing a hybrid experiment that is essentially Experiment 4a in its upper half and Experiment 4b in its lower half, and Fig. 2.2(b) showing Experiment 4c. As before, it is useful here to discuss the probability that an atom leaving the first

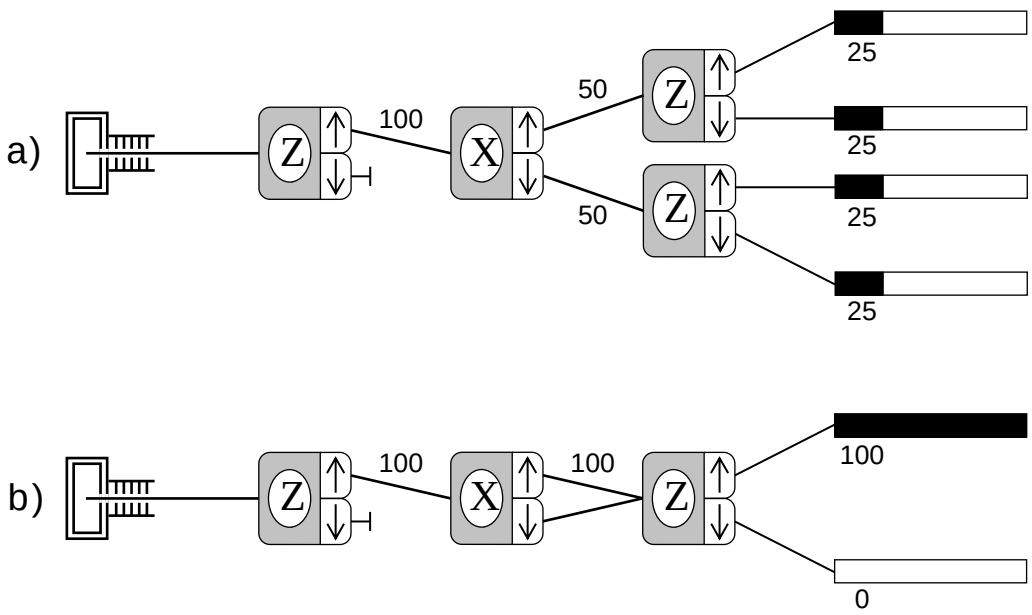

Figure 2.2. (a) Hybrid Experiment 4a and 4b, and (b) Experiment 4c.

analyzer in the $|+\rangle$ state is detected in one of the counters connected to the output ports of the third analyzer. Such a probability involves two measurement processes at the second and third analyzers. The total probability is the product of the individual probabilities of each measurement process.

For the hybrid experiment shown in Fig. 2.2(a), the probability of measuring an atom at the top most counter is the probability of measuring $S_x = +\hbar/2$ at the second analyzer, $|{}_x\langle +|+\rangle|^2$, times the probability of measuring $S_z = +\hbar/2$ at the third analyzer, $|\langle +|+\rangle_x|^2$, giving

$$\mathcal{P}_{\text{upper}}(+) = |\langle +|+\rangle_x|^2 |{}_x\langle +|+\rangle|^2. \tag{2.49}$$

Likewise the probability of measuring the atom to have $S_x = +\hbar/2$ and then $S_z = -\hbar/2$ is

$$\mathcal{P}_{\text{upper}}(-) = |\langle -|+\rangle_x|^2 |{}_x\langle +|+\rangle|^2, \tag{2.50}$$

where we have written the product so as to be read from right to left as is the usual practice with quantum mechanical amplitudes and probabilities. For atoms that take the lower path from the second analyzer, the final probabilities are

$$\begin{aligned} \mathcal{P}_{\text{lower}}(+) &= |\langle +|-\rangle_x|^2 |{}_x\langle -|+\rangle|^2 \\ \mathcal{P}_{\text{lower}}(-) &= |\langle -|-\rangle_x|^2 |{}_x\langle -|+\rangle|^2 \end{aligned}. \tag{2.51}$$

For Experiment 4c, shown in Fig. 2.2(b), we have a new situation at the second analyzer. Both output ports are connected to the third analyzer, which means that the probability of an atom from the first analyzer being input to the third analyzer is 100%. So we need only calculate the probability of passage through the third analyzer. The crucial step is determining the input state, for which we use the projection postulate. Since both states are used, the relevant projection operator is the sum of the two projection operators for each port, $P_{+x} + P_{-x}$, where $P_{+x} = |+\rangle_{x\,x}\langle +|$ and $P_{-x} = |-\rangle_{x\,x}\langle -|$. Thus the state after the second analyzer is

$$\begin{aligned} |\psi_2\rangle &= \frac{(P_{+x} + P_{-x})|\psi_1\rangle}{\sqrt{\langle\psi_1|(P_{+x} + P_{-x})|\psi_1\rangle}} \\ &= \frac{(P_{+x} + P_{-x})|+\rangle}{\sqrt{\langle +|(P_{+x} + P_{-x})|+\rangle}} \end{aligned} . \tag{2.52}$$

In this simple example, the projector $P_{+x} + P_{-x}$ is equal to the identity operator since the two states form a complete basis. This clearly simplifies the calculation, giving $|\psi_2\rangle = |+\rangle$, but in order to illustrate our point let's only simplify the denominator (which equals one), giving

$$\begin{aligned} |\psi_2\rangle &= \left(|+\rangle_{x\,x}\langle +| \;+\; |-\rangle_{x\,x}\langle -|\right)|+\rangle \\ &= |+\rangle_{x\,x}\langle +|+\rangle \;+\; |-\rangle_{x\,x}\langle -|+\rangle \end{aligned} . \tag{2.53}$$

Thus the beam entering the third analyzer can be viewed as a coherent superposition of the eigenstates of the second analyzer. Now calculate the probability of measuring spin up at the third analyzer:

$$\begin{aligned} \mathcal{P}(+) &= \left|\langle +|\psi_2\rangle\right|^2 \\ &= \left|\langle +|+\rangle_{x\,x}\langle +|+\rangle \;+\; \langle +|-\rangle_{x\,x}\langle -|+\rangle\right|^2 \end{aligned} . \tag{2.54}$$

The probability of measuring spin down at the third analyzer is similarly

$$\begin{aligned} \mathcal{P}(-) &= \left|\langle -|\psi_2\rangle\right|^2 \\ &= \left|\langle -|+\rangle_{x\,x}\langle +|+\rangle \;+\; \langle -|-\rangle_{x\,x}\langle -|+\rangle\right|^2 \end{aligned} . \tag{2.55}$$

In each case, the probability is a square of a sum of amplitudes, each amplitude being the amplitude for a successive pair of measurements. For example, in $\mathcal{P}(-)$ the amplitude $\langle -|+\rangle_{x\,x}\langle +|+\rangle$ refers to the upper path that the initial $|+\rangle$ state takes as it is first measured to be in the $|+\rangle_x$ state, and then measured to be in the $|-\rangle$ state (read from right to left). This amplitude is added to the amplitude for the lower path since the beams of the second analyzer are combined, in the proper fashion, to create the input beam to the third analyzer. When the sum of amplitudes is squared, four terms are obtained, two squares and two cross terms, giving

$$\begin{aligned}\mathcal{P}(-) &= \left|\langle -|+\rangle_{x\,x}\langle +|+\rangle\right|^2 \;+\; \left|\langle -|-\rangle_{x\,x}\langle -|+\rangle\right|^2 \\ &\quad +\langle -|+\rangle^*_{x\,x}\langle +|+\rangle^*\langle -|-\rangle_{x\,x}\langle -|+\rangle \\ &\quad +\langle -|+\rangle_{x\,x}\langle +|+\rangle\langle -|-\rangle^*_{x\,x}\langle -|+\rangle^* \\ &= \mathcal{P}_{\text{upper}}(-) + \mathcal{P}_{\text{lower}}(-) + \text{interference terms}\end{aligned} \quad . \tag{2.56}$$

This tells us that the probability of detecting an atom to have spin down when both paths are used is the sum of the probabilities for detecting a spin down atom when either the upper path or the lower path is used alone *plus* additional cross terms involving both amplitudes, which are commonly called interference terms. It is these additional terms, which are not positive definite, that allow the total probability to become zero in this case, illustrating the phenomenon of interference.

This interference arises from the nature of the superposition of states that is input to the third analyzer. To illustrate this, consider what happens if we change the superposition to a mixture as we discussed previously (Sec. 1.3.3). Recall that a superposition implies a beam with each atom in the same state, which is a combination of states, while a mixture implies that the beam consists of atoms in separate states. As we have described it so far, Experiment 4c involves a superposition as input to the third analyzer. We can change this to a mixture by "watching" to see which of the two output ports of the second analyzer each atom travels through. There are a variety of ways to imagine doing this experimentally. The usual idea proposed is to illuminate the paths with light and watch for the scattered light from the atoms. With proper design of the optics, the light can be localized sufficiently to determine which path the atom takes. Hence, such experiments are generally referred to as "**Which Path**" or "**Welcher Weg**" experiments. Such experiments can be performed in SPINS by selecting the Watch feature. Since we know which path the atom takes, the state is not the superposition $|\psi_2\rangle$ described above, but is either $|+\rangle_x$ or $|-\rangle_x$, depending on which path produces the light signal. To find the probability that atoms are detected at the spin down counter of the third analyzer, we add the probabilities for atoms to follow the path $|+\rangle \rightarrow |+\rangle_x \rightarrow |-\rangle$ to the probability for other atoms to follow the path $|+\rangle \rightarrow |-\rangle_x \rightarrow |-\rangle$ since these are independent events, giving

$$\begin{aligned}\mathcal{P}_{\text{watch}}(-) &= \left|\langle -|+\rangle_{x\,x}\langle +|+\rangle\right|^2 \;+\; \left|\langle -|-\rangle_{x\,x}\langle -|+\rangle\right|^2 \\ &= \mathcal{P}_{\text{upper}}(-) + \mathcal{P}_{\text{lower}}(-)\end{aligned} \quad , \tag{2.57}$$

in which no interference terms are present.

This again illustrates the important distinction between a coherent superposition and a statistical mixture. In a coherent superposition, there is a definite relative phase between the different states, which can give rise to interference effects that are dependent on that phase. In a statistical mixture, the phase relationship between the states has been destroyed and the interference is washed out. Now we can understand what it takes to have the beams "properly" combined after the second analyzer of Experiment 4c. The relative phases of the two paths must be preserved. Anything that randomizes the phase

is equivalent to destroying the superposition and leaving only a statistical mixture. If the beams are properly combined to leave the superposition intact, the results of Experiment 4c are the same as if no measurement was made at the second analyzer. So even though we have used a measuring device in the middle of Experiment 4c, we generally say that no measurement was made there. We can summarize our conclusions by saying that if no measurement is made at the intermediate state, then we add amplitudes and then square to find the probability, while if an intermediate measurement is performed (i.e., watching), then 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.

2.2 Measurement

Consider now how the probabilistic nature of quantum mechanics affects the way experiments are performed and compared with theory. In classical physics, a theoretical prediction can be reliably compared to a single experimental result. For example, a prediction of the range of a projectile can be tested by doing an experiment. The experiment may be repeated several times in order to understand and possibly reduce any systematic errors (e.g. wind) and measurement errors (e.g misreading tape measure). In quantum mechanics, a single measurement is meaningless. If we measure an atom to have spin up in a Stern-Gerlach analyzer, we cannot discern whether the original state was $|+\rangle$ or $|-\rangle_x$ or any arbitrary state $|\psi\rangle$ (except $|-\rangle$). Moreover, we cannot repeat the measurement on the same atom, since the original measurement changed the state, per the projection postulate.

Thus, one must, by necessity, perform identical measurements on identically prepared systems. In the spin 1/2 example, an initial Stern-Gerlach analyzer is used to prepare atoms in a particular state $|\psi\rangle$. Then a second Stern-Gerlach analyzer is used to perform the same experiment on each identically prepared atom. Consider performing a measurement of S_z on N identically prepared atoms. Let N_+ be the number of times the result $+\hbar/2$ is recorded and N_- be the number of times the result $-\hbar/2$ is recorded. Since there are only two possible results for each measurement, we must have $N = N_+ + N_-$. The probability postulate (postulate 4) predicts that the probability of measuring $+\hbar/2$ is

$$\mathcal{P}(+) = |\langle + | \psi \rangle|^2 . \tag{2.58}$$

For a finite number N of atoms, we expect that N_+ is only approximately equal to $\mathcal{P}(+)N$ due to the statistical fluctuations inherent in a random process (see Appendix A). Only in the limit of an infinite number N do we expect exact agreement:

$$\lim_{N\to\infty} \frac{N_+}{N} = \mathcal{P}(+) = |\langle + | \psi \rangle|^2 . \tag{2.59}$$

As with any data set, it is useful to characterize it in terms of the **mean** and **standard deviation**. Quantum mechanics predicts a mean value given by the sum of the products of each possible result and its probability:

$$\langle S_z \rangle = +\frac{\hbar}{2}\mathcal{P}(+) + \left(-\frac{\hbar}{2}\right)\mathcal{P}(-), \tag{2.60}$$

which can be rewritten as

$$\begin{aligned}
\langle S_z \rangle &= +\frac{\hbar}{2}\left|\langle +|\psi\rangle\right|^2 + \left(-\frac{\hbar}{2}\right)\left|\langle -|\psi\rangle\right|^2 \\
&= +\frac{\hbar}{2}\langle\psi|+\rangle\langle +|\psi\rangle + \left(-\frac{\hbar}{2}\right)\langle\psi|-\rangle\langle -|\psi\rangle \\
&= \langle\psi|\left[+\frac{\hbar}{2}|+\rangle\langle +|\psi\rangle + \left(-\frac{\hbar}{2}\right)|-\rangle\langle -|\psi\rangle\right]. \\
&= \langle\psi|\left[S_z|+\rangle\langle +|\psi\rangle + S_z|-\rangle\langle -|\psi\rangle\right] \\
&= \langle\psi|S_z\left[|+\rangle\langle +| + |-\rangle\langle -|\right]|\psi\rangle \\
&= \langle\psi|S_z|\psi\rangle
\end{aligned} \tag{2.61}$$

This is commonly called the **expectation value**, but it is not the expected value of any single experiment, but rather the expected mean value of a large number of experiments. It is not a time average, but rather an average over many identical experiments.

To see how this applies to our study of spin 1/2 systems, consider two examples. First consider a system prepared in the state $|+\rangle$. The expectation value of S_z is given by

$$\begin{aligned}
\langle S_z \rangle &= \langle +|S_z|+\rangle \\
&= \langle +|\frac{\hbar}{2}|+\rangle \\
&= \frac{\hbar}{2}\langle +|+\rangle \\
&= \frac{\hbar}{2}
\end{aligned}. \tag{2.62}$$

Since $+\hbar/2$ is the only possible result of a measurement of S_z for this state, this must be the expectation value. Next consider a system prepared in the state $|+\rangle_x$. In this case the expectation value is

$$\begin{aligned}
\langle S_z \rangle &= {}_x\langle +|S_z|+\rangle_x \\
&= \tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} \\
&= \frac{\hbar}{4}\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1\\ -1\end{pmatrix} \\
&= 0
\end{aligned}. \tag{2.63}$$

Here the two possible results $\pm\hbar/2$ each have 50% probability, so the average result is zero.

In addition to the mean value, it is common to characterize a measurement by the standard deviation, which quantifies the spread of measurements about the mean or expectation value. By the way a mean is defined, the deviations from the mean are equally distributed above and below the mean in such a way that the average of the deviations is zero. Thus it is common to define the standard deviation as the square root of the mean of the square of the deviations. By squaring the deviations, they can be averaged meaningfully, and then the square root is taken to arrive back at a measure of the "average" deviation. For a measurement of an observable A, the standard deviation is defined as

$$\Delta A = \sqrt{\left\langle \left(A - \langle A \rangle\right)^2 \right\rangle}, \tag{2.64}$$

where the angle brackets signify average value as used in the definition of an expectation value. This result is also often called the **root-mean-square deviation**, or r.m.s. deviation. This expression can be simplified by expanding the square and performing the averages, resulting in

$$\begin{aligned} \Delta A &= \sqrt{\left\langle \left(A^2 - 2A\langle A \rangle + \langle A \rangle^2\right) \right\rangle} \\ &= \sqrt{\left\langle A^2 \right\rangle - 2\langle A \rangle \langle A \rangle + \langle A \rangle^2}\,, \\ &= \sqrt{\left\langle A^2 \right\rangle - \langle A \rangle^2} \end{aligned} \tag{2.65}$$

where one must be clear to distinguish between the square of the mean and the mean of the square. While the mean of the square may not be a common experimental quantity, it can be calculated using its definition

$$\left\langle A^2 \right\rangle = \langle \psi | A^2 | \psi \rangle. \tag{2.66}$$

To gain experience with the standard deviation, return to the two examples used above. To calculate the standard deviation, we first need to find the mean of the square of the operator S_z. In the first case ($|+\rangle$ initial state), we get

$$\begin{aligned} \left\langle S_z^2 \right\rangle &= \langle + | S_z^2 | + \rangle \\ &= \langle + | \left(\frac{\hbar}{2}\right)^2 | + \rangle. \\ &= \left(\frac{\hbar}{2}\right)^2 \end{aligned} \tag{2.67}$$

This gives a standard deviation of

$$\begin{aligned}\Delta S_z &= \sqrt{\langle S_z^2\rangle - \langle S_z\rangle^2} \\ &= \sqrt{\left(\frac{\hbar}{2}\right)^2 - \left(\frac{\hbar}{2}\right)^2}, \\ &= 0\end{aligned} \tag{2.68}$$

which is to be expected since there is only one possible result, and hence no spread in the results of the measurement, as shown in Fig. 2.3(a). In the second case ($|+\rangle_x$ initial state), we get

$$\begin{aligned}\langle S_z^2\rangle &= {}_x\langle +|S_z^2|+\rangle_x \\ &= \tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} \\ &= \tfrac{1}{2}\left(\frac{\hbar}{2}\right)^2\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1\\ -1\end{pmatrix} \\ &= \tfrac{1}{2}\left(\frac{\hbar}{2}\right)^2\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix} \\ &= \left(\frac{\hbar}{2}\right)^2\end{aligned} \quad . \tag{2.69}$$

This gives a standard deviation of

$$\begin{aligned}\Delta S_z &= \sqrt{\langle S_z^2\rangle - \langle S_z\rangle^2} \\ &= \sqrt{\left(\frac{\hbar}{2}\right)^2 - 0} \\ &= \frac{\hbar}{2}\end{aligned} \tag{2.70}$$

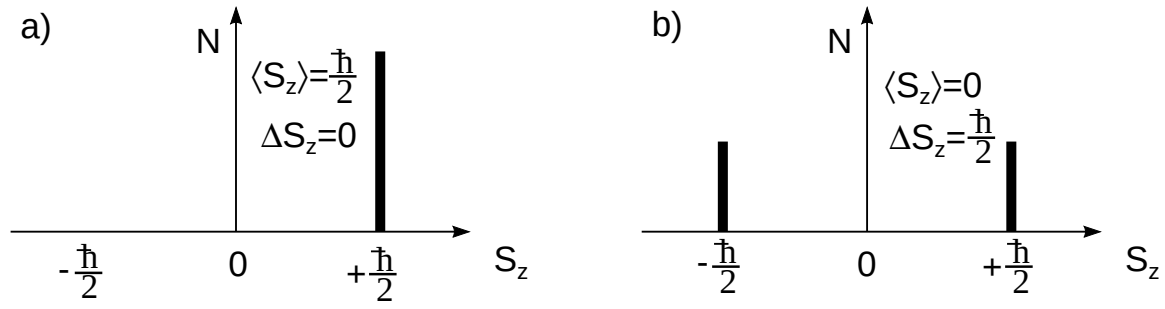

Figure 2.3. Idealized measurements of S_z with (a) $|+\rangle$ input state and (b) with $|+\rangle_x$ input state.

Again this makes sense since each measurement deviates from the mean (0) by the same value of $\hbar/2$, as shown in Fig. 2.3(b).

The standard deviation ΔA represents the uncertainty in the results of an experiment. In quantum mechanics, this uncertainty is fundamental, meaning that you cannot design the experiment any better to improve the result. What we have calculated then, is the minimum uncertainty allowed by quantum mechanics. Any actual uncertainty may be larger due to experimental error. This is another ramification of the probabilistic nature of quantum mechanics.

2.3 Commuting Observables

We saw in Experiment 2 that two incompatible observables could not be known or measured simultaneously, since measurement of one somehow erased the knowledge of the other. Let us now explore more about what it means for two observables to be incompatible and how that affects the results of measurements. First we need to define a new object called a **commutator**. The commutator of two operators is defined as

$$[A,B] = AB - BA. \tag{2.71}$$

If the commutator is equal to zero, we say that the operators or observables **commute**; if it is not zero, we say they don't commute. Now consider what happens when two operators *A* and *B* do commute:

$$\begin{aligned} [A,B] &= 0 \\ AB - BA &= 0 \\ AB &= BA \end{aligned} \quad . \tag{2.72}$$

Thus for commuting operators, the order of operation does not matter, whereas it does for noncommuting operators. Now let $|a\rangle$ be an eigenstate of the operator *A* with eigenvalue *a*:

$$A|a\rangle = a|a\rangle. \tag{2.73}$$

Operate on both sides of this equation with the operator *B* and use the fact that *A* and *B* commute:

$$\begin{aligned} BA|a\rangle &= Ba|a\rangle \\ AB|a\rangle &= aB|a\rangle \\ A\big(B|a\rangle\big) &= a\big(B|a\rangle\big) \end{aligned} \quad . \tag{2.74}$$

The last equation says that the state $B|a\rangle$ is also an eigenstate of the operator *A* with the same eigenvalue *a*. Assuming that each eigenvalue has a unique eigenstate (which is true if there is no degeneracy, which we haven't discussed yet), the state $B|a\rangle$ must be some scalar multiple of the state $|a\rangle$. If we call this multiple *b*, then we can write

$$B|a\rangle = b|a\rangle, \tag{2.75}$$

which is just an eigenvalue equation for the operator *B*. Thus we must conclude that the state $|a\rangle$ is also an eigenstate of the operator *B*, with the eigenvalue *b*. The assumption

that the operators A and B commute has lead us to the result that A and B have common or simultaneous sets of eigenstates.

The ramifications of this result for experiments are very important. Recall that a measurement of the observable A projects the initial state $|\psi\rangle$ onto an eigenstate of A: $|a\rangle$. A subsequent measurement of the observable B then projects the input state $|a\rangle$ onto an eigenstate of B. But since the eigenstates of A and B are the same, the second measurement does not change the state $|a\rangle$. Thus another measurement of A following the measurement of B yields the same result as the initial measurement of A, as illustrated in Fig. 2.4. We say that these two observables can be measured simultaneously, meaning 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.

Conversely, if two operators do not commute, then they are incompatible observables, and cannot be measured simultaneously. This is what we saw in Experiment 2 earlier. In that case, the two observables were S_x and S_z. Let's take a look at their commutator to show that they are not compatible:

$$\begin{aligned}
[S_z, S_x] &\stackrel{\bullet}{=} \frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} - \frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \\
&\stackrel{\bullet}{=} \left(\frac{\hbar}{2}\right)^2\left[\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} - \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\right] \\
&\stackrel{\bullet}{=} \left(\frac{\hbar}{2}\right)^2\begin{pmatrix}0 & 2\\ -2 & 0\end{pmatrix} \\
&= i\hbar S_y
\end{aligned} \tag{2.76}$$

As expected, these two operators do not commute. None of the spin projection observables commute with each other. The complete commutation relations are

$$\begin{aligned}
[S_x, S_y] &= i\hbar S_z \\
[S_y, S_z] &= i\hbar S_x \,, \\
[S_z, S_x] &= i\hbar S_y
\end{aligned} \tag{2.77}$$

which are written to make the cyclic relations clear.

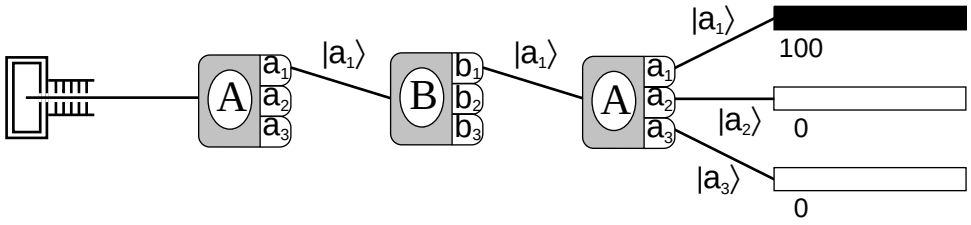

Figure 2.4. Successive measurements of commuting observables.

2.4 Uncertainty Principle

Given this intimate connection between the commutator of two observables and measurements of the two corresponding observables, it is useful and instructive to also consider the product of uncertainties or standard deviations of the two observables. One can show that this product is related to the commutator:

$$\Delta A \Delta B \geq \tfrac{1}{2}\left|\langle [A,B] \rangle\right|. \tag{2.78}$$

This is what we call the **uncertainty principle** of quantum mechanics. Consider what it says about a simple Stern-Gerlach experiment. The uncertainty principle for the S_x and S_y spin projections is

$$\begin{aligned} \Delta S_x \Delta S_y &\geq \tfrac{1}{2}\left|\left\langle [S_x, S_y] \right\rangle\right| \\ &\geq \tfrac{1}{2}\left|\left\langle i\hbar S_z \right\rangle\right| \\ &\geq \frac{\hbar}{2}\left|\left\langle S_z \right\rangle\right| \end{aligned} \quad . \tag{2.79}$$

If the initial state is $|+\rangle$, then a measurement of S_z results in an expectation value $\langle S_z \rangle = \hbar/2$ with an uncertainty $\Delta S_z = 0$. The product of the other uncertainties in this case is

$$\Delta S_x \Delta S_y \geq \left(\frac{\hbar}{2}\right)^2, \tag{2.80}$$

or simply

$$\Delta S_x \Delta S_y \neq 0, \tag{2.81}$$

which implies that

$$\begin{aligned} \Delta S_x &\neq 0 \\ \Delta S_y &\neq 0 \end{aligned}. \tag{2.82}$$

The conclusion to draw from this is that while we can know one spin projection absolutely, we can never know all three, nor even two, simultaneously. This means that the spin does not really point in a given direction, as a classical spin or angular momentum does. So when we say that we have measured "spin up," we really mean only that the spin projection along that axis is up, as opposed to down.

2.5 S^2 Operator

Another argument to illustrate that the spin does not point along the axis along which you measure the projection is obtained by considering a new operator that represents the magnitude of the spin vector but has no information about the direction. It is common to use the square of the spin vector for this task. This new operator can be written as

$$\mathbf{S}^2 = S_x^2 + S_y^2 + S_z^2, \tag{2.83}$$

and can be calculated in the S_z representation as

$$\begin{aligned}\mathbf{S}^2 &\doteq \left(\frac{\hbar}{2}\right)^2\left[\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}+\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}+\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right]\\ &\doteq \left(\frac{\hbar}{2}\right)^2\left[\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}+\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}+\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right]\\ &\doteq \tfrac{3}{4}\hbar^2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\end{aligned}. \qquad (2.84)$$

Thus the $\mathbf{S}^2$ operator is proportional to the identity operator, which means it must commute with all the other operators S_x, S_y, and S_z. This also means that all states are eigenstates of the $\mathbf{S}^2$ operator. Thus we can write

$$\mathbf{S}^2|\psi\rangle = \tfrac{3}{4}\hbar^2|\psi\rangle \qquad (2.85)$$

for any $|\psi\rangle$ in the spin 1/2 system.

To generalize this to other possible spin systems, we need to introduce new labels. We use the label *s* to denote the spin of the system, such as spin 1/2, spin 1, spin 3/2. In each of these cases, a measurement of a spin projection along any axis yields results ranging from a maximum value of $s\hbar$ to a minimum value of $-s\hbar$, in unit steps of the value $\hbar$. We denote the possible values of the spin projection along the z-axis by *m*, the integer or half-integer multiplying $\hbar$. A quantum state with specific values of *s* and *m* is denoted as $|s,m\rangle$, yielding the eigenvalue equations

$$\begin{aligned}\mathbf{S}^2|s,m\rangle &= s(s+1)\hbar^2|s,m\rangle\\ S_z|s,m\rangle &= m\hbar|s,m\rangle\end{aligned}. \qquad (2.86)$$

The connection between this new notation and our previous spin 1/2 notation is

$$\begin{aligned}\left|\tfrac{1}{2},\tfrac{1}{2}\right\rangle &= |+\rangle\\ \left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle &= |-\rangle\end{aligned}. \qquad (2.87)$$

We will continue to use the previous notation. We will discuss the spin 1 case below.

For the case of spin 1/2, note that the expectation value of the operator $\mathbf{S}^2$ is

$$\left\langle \mathbf{S}^2 \right\rangle = \tfrac{3}{4}\hbar^2, \qquad (2.88)$$

which would imply that the "length" of the spin vector is

$$\sqrt{\mathbf{S}^2} = \sqrt{3}\frac{\hbar}{2}. \qquad (2.89)$$

This is appreciably longer than the measured projection of $\hbar/2$, implying that the spin vector can never be fully aligned along any axis.

The spin 1/2 system we have discussed up to this point is the simplest quantum mechanical system. There are other two-level systems based upon other physical

systems, but they are mathematically equivalent to the two-state spin 1/2 system we have explored. The next simplest system is the spin 1 or three-level system.

2.6 Spin 1 System

The Stern-Gerlach experiment depicted in Fig. 1.1 can be performed on a variety of atoms or particles. Such experiments always result in a finite number of discrete beams exiting the analyzer magnet. For the case of three output beams, the deflections are consistent with magnetic moments arising from spin angular momentum projections of $\hbar$, 0, and $-\hbar$. These results imply that the spin projection quantum numbers m are 1, 0 and –1 and the spin quantum number $s = 1$. This is what we call a spin 1 system. A schematic diagram of the experiment is shown in Fig. 2.5. Using our notation from above, the quantum state corresponding to the state with spin projection $\hbar$ along the z-axis is labeled

$$|s=1, m=1\rangle = |1,1\rangle. \tag{2.90}$$

When dealing with a spin 1 system by itself, we will simply denote this state as $|1\rangle$ since the value s is understood. Thus we are lead to the three eigenvalue equations for the operator S_z

$$\begin{aligned} S_z|1\rangle &= \hbar|1\rangle \\ S_z|0\rangle &= 0 \\ S_z|-1\rangle &= -\hbar|-1\rangle \end{aligned} \quad . \tag{2.91}$$

As we did in the spin 1/2 case, we commonly choose the S_z basis as the standard basis in which to express operators and kets using matrix representation. We thus have

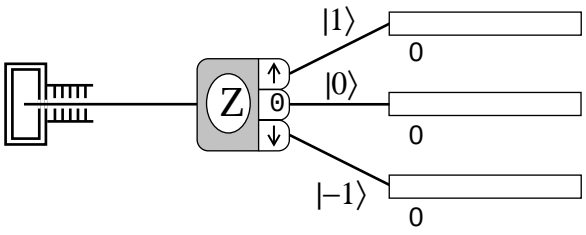

Figure 2.5. Spin 1 Stern-Gerlach experiment.

$$S_z \doteq \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{2.92}$$

where we again use the convention that the ordering of the rows and columns follows the eigenvalues in descending order. The eigenvectors of the S_z operator are

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad |0\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad |-1\rangle \doteq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.93}$$

The same four experiments performed on the spin 1/2 system can be performed on a spin 1 system. Conceptually the results are the same. One important difference occurs in experiment 2, where a measurement of S_z is first performed to prepare a particular state, and then a subsequent measurement of S_x (or S_y) is performed. Based upon the results of the spin 1/2 experiment, one might expect each of the possible projections to have 1/3 probability. Such is not the case. Rather one set of results is

$$\begin{aligned} |{}_x\langle 1|1\rangle|^2 &= \tfrac{1}{4} \\ |{}_x\langle 0|1\rangle|^2 &= \tfrac{1}{2} \\ |{}_x\langle -1|1\rangle|^2 &= \tfrac{1}{4} \end{aligned}, \tag{2.94}$$

as illustrated in Fig. 2.6. This results can be used to determine the S_x eigenstates in terms of the S_z basis:

$$\begin{aligned} |1\rangle_x &= \frac{1}{2}|1\rangle + \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|-1\rangle \\ |0\rangle_x &= \frac{1}{\sqrt{2}}|1\rangle - \frac{1}{\sqrt{2}}|-1\rangle \\ |-1\rangle_x &= \frac{1}{2}|1\rangle - \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{2}|-1\rangle \end{aligned}. \tag{2.95}$$

Likewise, we can find the S_y eigenstates:

$$\begin{aligned} |1\rangle_y &= \frac{1}{2}|1\rangle + \frac{i}{\sqrt{2}}|0\rangle - \frac{1}{2}|-1\rangle \\ |0\rangle_y &= \frac{1}{\sqrt{2}}|1\rangle + \frac{1}{\sqrt{2}}|-1\rangle \\ |-1\rangle_y &= \frac{1}{2}|1\rangle - \frac{i}{\sqrt{2}}|0\rangle - \frac{1}{2}|-1\rangle \end{aligned}. \tag{2.96}$$

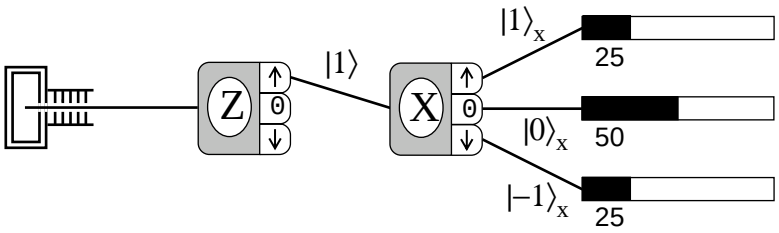

Figure 2.6. Experiment 2 in the spin 1 case.

Problems

2.1. Given the following information:

$$S_x|\pm\rangle_x = \pm\frac{\hbar}{2}|\pm\rangle_x \qquad S_y|\pm\rangle_y = \pm\frac{\hbar}{2}|\pm\rangle_y$$

$$|\pm\rangle_x = \frac{1}{\sqrt{2}}\left[|+\rangle \pm |-\rangle\right] \qquad |\pm\rangle_y = \frac{1}{\sqrt{2}}\left[|+\rangle \pm i|-\rangle\right]$$

find the matrix representations of S_x and S_y in the S_z basis.

2.2. Find the matrix representation of S_z in the S_x basis, for spin 1/2. Diagonalize this matrix to find the eigenvalues and the eigenvectors in this basis.

2.3. Calculate the commutators of the spin 1/2 operators S_x, S_y, and S_z.

2.4. In part (1) of SPINS Lab #2, you measured the probabilities that each of the unknown initial states $|\psi_i\rangle$ (i = 1, 2, 3, 4) was measured to be in one of the six spin eigenstates $|\pm\rangle$, $|\pm\rangle_x$, and $|\pm\rangle_y$. Using your measured values, deduce the unknown initial states (write each of them using the $|\pm\rangle$ basis). In each case, use your result to calculate the theoretical values of the probabilities for each projection measurement and compare with your experimental results. Discuss the experiments you did to confirm your results for $|\psi_3\rangle$ and $|\psi_4\rangle$.

2.5. Consider a three dimensional ket space. In the basis defined by three orthogonal kets $|1\rangle$, $|2\rangle$, and $|3\rangle$, the operators A and B are represented by

$$A \doteq \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \qquad B \doteq \begin{pmatrix} b_1 & 0 & 0 \\ 0 & 0 & b_2 \\ 0 & b_2 & 0 \end{pmatrix},$$

where all the quantities are real.

a) Do the operators A and B commute?

b) Find the eigenvalues and normalized eigenvectors of both operators.

c) Assume the system is initially in the state $|2\rangle$. Then the observable corresponding to the operator B is measured. What are the possible results of this measurement and the probabilities of each result? After this measurement, the observable corresponding to the operator A is measured. What are the possible results of this measurement and the probabilities of each result?

d) How are questions (a) and (c) above related?

2.6. (Townsend 3.17) A spin-1 particle is in the state

$$|\psi\rangle \doteq \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3i \end{pmatrix}$$

a) What are the probabilities that a measurement of S_z will yield the values $\hbar$, 0, $-\hbar$ for this state? What is $\langle S_z \rangle$?

b) What is $\langle S_x \rangle$ for this state? *Suggestion*: Use matrix mechanics to evaluate the expectation value.

c) What is the probability that a measurement of S_x will yield the value $\hbar$ for this state?

2.7. In part (2) of SPINS Lab #3, you measured the spin projections of the unknown (spin 1) initial states $|\psi_i\rangle$ (i = 1, 2, 3, 4) onto the nine spin eigenstates along the three axes. Using your measured values, deduce the unknown initial states.

Chapter 3 SCHRÖDINGER TIME EVOLUTION

3.1 Schrödinger Equation

The 6th postulate of quantum mechanics says that the time evolution of a quantum system is governed by the equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle, \tag{3.1}$$

where the operator H corresponds to the total energy of the system and is called the Hamiltonian operator of the system since it is derived from the classical Hamiltonian. This equation is known as the **Schrödinger equation**.

> Postulate 6
> The time evolution of a quantum system is determined by the Hamiltonian or total energy operator $H(t)$ through the Schrödinger equation
>
> $$i\hbar\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle.$$

The eigenstates of the Hamiltonian are called the energy eigenstates of the system and the eigenvalues of H are the allowed energies of the quantum system. If we label the allowed energies as E_i, then we can write the eigenvalue equation as

$$H|E_i\rangle = E_i|E_i\rangle. \tag{3.2}$$

Since H is an observable, it is an Hermitian operator and its eigenvectors form a complete basis. Since H is the only operator appearing in the Schrödinger equation, it would seem reasonable (and will prove invaluable) to consider the energy eigenstates as the basis of choice for expanding general state vectors:

$$|\psi\rangle = \sum_i c_i|E_i\rangle. \tag{3.3}$$

For now assume that the Hamiltonian is time independent (we will do $H(t)$ later). Then the eigenvectors of H must themselves be time independent. Thus if a general state $|\psi\rangle$ is to be time dependent, as the Schrödinger equation implies, then the time dependence must be in the expansion coefficients c_i, which we write as $c_i(t)$. A general time dependent state is then written as

$$|\psi(t)\rangle = \sum_i c_i(t)|E_i\rangle. \tag{3.4}$$

Substitute this state into the Schrödinger equation (Eq. (3.1)) to get

$$\begin{aligned} i\hbar\frac{d}{dt}\sum_i c_i(t)\left|E_i\right\rangle &= H\sum_i c_i(t)\left|E_i\right\rangle \\ i\hbar\sum_i \frac{dc_i(t)}{dt}\left|E_i\right\rangle &= \sum_i c_i(t)E_i\left|E_i\right\rangle \end{aligned} \quad . \tag{3.5}$$

Each side of this equation is a sum over all the states of the system. To simplify this equation, take the inner product of the ket on each side with the ket $\left|E_j\right\rangle$. The orthonormality condition $\left\langle E_j\middle|E_i\right\rangle = \delta_{ij}$ can then be used to collapse the sums:

$$\begin{aligned} \left\langle E_j\middle|i\hbar\sum_i \frac{dc_i(t)}{dt}\middle|E_i\right\rangle &= \left\langle E_j\middle|\sum_i c_i(t)E_i\middle|E_i\right\rangle \\ i\hbar\sum_i \frac{dc_i(t)}{dt}\left\langle E_j\middle|E_i\right\rangle &= \sum_i c_i(t)E_i\left\langle E_j\middle|E_i\right\rangle \\ i\hbar\sum_i \frac{dc_i(t)}{dt}\delta_{ij} &= \sum_i c_i(t)E_i\delta_{ij} \\ i\hbar\frac{dc_j(t)}{dt} &= c_j(t)E_j \end{aligned} \quad . \tag{3.6}$$

This first-order differential equation can be written as

$$\frac{dc_j(t)}{dt} = -i\frac{E_j}{\hbar}c_j(t) \tag{3.7}$$

and solved to obtain

$$c_j(t) = c_j(0)e^{-iE_jt/\hbar}. \tag{3.8}$$

In this equation, we have denoted the initial condition as $c_j(0)$, but will simply denote it as c_j hereafter. This solution can be summarized by saying that if the initial state of a system at time $t = 0$ is given by

$$\left|\psi(0)\right\rangle = \sum_i c_i\left|E_i\right\rangle, \tag{3.9}$$

then the time evolution of this state under the action of the time-independent Hamiltonian H is given by

$$\left|\psi(t)\right\rangle = \sum_i c_i e^{-iE_it/\hbar}\left|E_i\right\rangle. \tag{3.10}$$

So the time dependence of the original state vector is found by simply multiplying each energy eigenstate coefficient by a phase factor dependent on that energy. Note that the factor $E/\hbar$ is an angular frequency, so that the time dependence is of the form $e^{-i\omega t}$ that is common throughout many areas of physics. It is important to remember that one must use the energy eigenstates in order to make this simple statement. This accounts for

the importance of the Hamiltonian operator and the common practice of using the energy eigenstates as the basis of choice.

A few simple examples can help to illustrate some of the important consequences of this time evolution of the quantum mechanical state vector. First consider the simplest possible situation where the system is initially in one particular energy eigenstate:

$$|\psi(0)\rangle = |E_1\rangle, \tag{3.11}$$

for example. The prescription for time evolution tells us that after some time t the system will be in the state

$$|\psi(t)\rangle = e^{-iE_1 t/\hbar}|E_1\rangle. \tag{3.12}$$

But this state differs from the original state only by a phase factor, which we have said before does not affect any measurements. For example, if we measure an observable A, then the probability of measuring an eigenvalue a_j is given by

$$\begin{aligned} \mathcal{P}(a_j) &= \left|\langle a_j|\psi(t)\rangle\right|^2 \\ &= \left|\langle a_j|e^{-iE_1 t/\hbar}|E_1\rangle\right|^2 \\ &= \left|\langle a_j|E_1\rangle\right|^2 \end{aligned} \tag{3.13}$$

This probability is time-independent and is equal to the probability at the initial time. Thus we conclude that there is no measureable time evolution for this state. Hence the energy eigenstates are generally called **stationary states**. If a system begins in an energy eigenstate, then it will remain in that state.

Now consider the next simplest situation, wherein the initial state is a superposition of two energy eigenstates:

$$|\psi(0)\rangle = c_1|E_1\rangle + c_2|E_2\rangle. \tag{3.14}$$

In this case, time evolution takes the initial state to the later state

$$|\psi(t)\rangle = c_1 e^{-iE_1 t/\hbar}|E_1\rangle + c_2 e^{-iE_2 t/\hbar}|E_2\rangle. \tag{3.15}$$

A measurement of the system energy at the time t would yield the value E_1 with a probability

$$\begin{aligned} \mathcal{P}(E_1) &= \left|\langle E_1|\psi(t)\rangle\right|^2 \\ &= \left|\langle E_1|\left[c_1 e^{-iE_1 t/\hbar}|E_1\rangle + c_2 e^{-iE_2 t/\hbar}|E_2\rangle\right]\right|^2 \\ &= |c_1|^2 \end{aligned} \tag{3.16}$$

which is independent of time. The same is true for the probability of measuring the energy E_2. Thus the probabilities of measuring the energies are stationary, as they were in the first example.

However, now consider what happens if another observable is measured on this system in this superposition state. There are two distinct situations: (1) If the other observable *A* commutes with the Hamiltonian *H*, then *A* and *H* have common eigenstates. In this case, measuring *A* is equivalent to measuring *H* since the inner products used to calculate the probabilities use the same eigenstates. (2) If *A* and *H* do not commute, then they do not share common eigenstates. In this case, the eigenstates of *A* in general consist of superpositions of energy eigenstates. For example, suppose that the eigenstate of *A* corresponding to the eigenvalue a_1 were

$$|a_1\rangle = \alpha_1|E_1\rangle + \alpha_2|E_2\rangle. \tag{3.17}$$

Then the probability of measuring the eigenvalue a_1 would be

$$\begin{aligned}\mathcal{P}(a_1) &= \left|\langle a_1|\psi(t)\rangle\right|^2 \\ &= \left|\left[\alpha_1^*\langle E_1| + \alpha_2^*\langle E_2|\right]\left[c_1 e^{-iE_1 t/\hbar}|E_1\rangle + c_2 e^{-iE_2 t/\hbar}|E_2\rangle\right]\right|^2 \\ &= \left|\alpha_1^* c_1 e^{-iE_1 t/\hbar} + \alpha_2^* c_2 e^{-iE_2 t/\hbar}\right|^2 \\ &= \left|\alpha_1^* c_1 + \alpha_2^* c_2 e^{-i(E_2-E_1)t/\hbar}\right|^2\end{aligned} \tag{3.18}$$

The differing phases of the two components of $|\psi(t)\rangle$ lead to a time dependence in the probability. Since only the relative phase is important in the probability calculation, the time dependence is determined by the difference of the energies of the two states involved in the superposition. The corresponding frequency

$$\omega_{21} = \frac{E_2 - E_1}{\hbar} \tag{3.19}$$

is often called the **Bohr frequency**.

To summarize, we list below a recipe for solving a standard time-dependent quantum mechanics problem with a time-independent Hamiltonian.

Given a Hamiltonian *H* and an initial state $|\psi(0)\rangle$, what is the probability that a_n is measured at time *t*?

1. Diagonalize *H* (find eigenvalues E_i and eigenvectors $|E_i\rangle$)
2. Write $|\psi(0)\rangle$ in terms of energy eigenstates $|E_i\rangle$
3. Multiply each eigenstate coefficient by $e^{-i\frac{E_i}{\hbar}t}$ to get $|\psi(t)\rangle$
4. Calculate probability $P(a_n) = \left|\langle a_n|\psi(t)\rangle\right|^2$

3.2 Spin Precession

Now apply this new concept of Schrödinger time evolution to the case of a spin 1/2 system. The Hamiltonian represents the total energy of the system, but since only energy differences are important in time dependent solutions (and since we can define the zero of potential energy as we wish), we need only consider energy terms that differentiate between the two possible spin states in the system. Our experience with the Stern-Gerlach apparatus tells us that the potential energy of the magnetic dipole differs for the two possible spin projection states. So to begin, we will consider the potential energy of a magnetic dipole in a uniform magnetic field as the sole term in the Hamiltonian. Recalling that the magnetic dipole is given by

$$\boldsymbol{\mu} = g\frac{q}{2m_e c}\mathbf{S}, \tag{3.20}$$

the Hamiltonian is

$$\begin{aligned} H &= -\boldsymbol{\mu}\bullet\mathbf{B} \\ &= -g\frac{q}{2m_e c}\mathbf{S}\bullet\mathbf{B}, \\ &= \frac{e}{m_e c}\mathbf{S}\bullet\mathbf{B} \end{aligned} \tag{3.21}$$

where $q = -e$ and $g = 2$ have been used in the last line. The gyromagnetic ratio, g, is slightly different from 2, but we ignore that for now.

3.2.1 Magnetic Field in z-direction

For our first example we assume that the magnetic field is uniform and directed along the z-axis. Writing the magnetic field as

$$\mathbf{B} = B_0\hat{\mathbf{z}}, \tag{3.22}$$

allows the Hamiltonian to be simplified to

$$\begin{aligned} H &= \frac{eB_0}{m_e c}S_z, \\ &= \omega_0 S_z \end{aligned} \tag{3.23}$$

where we have introduced the definition

$$\omega_0 \equiv \frac{eB_0}{m_e c}. \tag{3.24}$$

This definition of an angular frequency simplifies the notation now and will lead to an obvious interpretation at the end.

Since the Hamiltonian is proportional to the S_z operator, H and S_z commute and share common eigenstates. This is clear if we write the Hamiltonian as a matrix in the S_z representation:

$$H \doteq \frac{\hbar\omega_0}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \tag{3.25}$$

Since H is diagonal, its eigenstates must be the basis states of the representation, while its eigenvalues are the diagonal elements of the matrix. The eigenvalue equations for the Hamiltonian are then

$$\begin{aligned} H|+\rangle &= \omega_0 S_z|+\rangle = \frac{\hbar\omega_0}{2}|+\rangle = E_+|+\rangle \\ H|-\rangle &= \omega_0 S_z|-\rangle = -\frac{\hbar\omega_0}{2}|+\rangle = E_-|-\rangle \end{aligned} \tag{3.26}$$

with eigenvalues and eigenvectors given by

$$\begin{aligned} E_+ &= \frac{\hbar\omega_0}{2} & E_- &= -\frac{\hbar\omega_0}{2} \\ |E_+\rangle &= |+\rangle & |E_-\rangle &= |-\rangle \end{aligned} \tag{3.27}$$

Now consider a few examples to illustrate the key features of the behavior of a spin 1/2 system in a uniform magnetic field. First consider the case where the initial state is spin up along z-axis:

$$|\psi(0)\rangle = |+\rangle. \tag{3.28}$$

The Schrödinger equation time evolution takes this to

$$\begin{aligned} |\psi(t)\rangle &= e^{-iE_+t/\hbar}|+\rangle \\ &= e^{-i\omega_0 t/2}|+\rangle \end{aligned}. \tag{3.29}$$

As we saw before, since the initial state is an energy eigenstate, the time evolved state simply has a phase factor in front, which does not represent a physical change of the state. The probability for measuring the spin to be up along the z-axis is

$$\begin{aligned} \mathcal{P}(+) &= \left|\langle +|\psi(t)\rangle\right|^2 \\ &= \left|\langle +|e^{-i\omega_0 t/2}|+\rangle\right|^2 \\ &= 1 \end{aligned}. \tag{3.30}$$

As expected, this is not time dependent. As before, we thus refer to $|+\rangle$ as a stationary state for this system.

Next consider the most general initial state, which we saw earlier corresponds to spin up along an arbitrary direction. Thus we will write the initial state as

$$|\psi(0)\rangle = |+\rangle_n = \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{i\phi}|-\rangle \tag{3.31}$$

or using matrix notation:

$$|\psi(0)\rangle \doteq \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\sin(\theta/2) \end{pmatrix}. \tag{3.32}$$

The time evolution simply introduces a time dependent phase term in each component, giving

$$\begin{aligned} |\psi(t)\rangle &\doteq \begin{pmatrix} e^{-iE_+t/\hbar}\cos(\theta/2) \\ e^{-iE_-t/\hbar}e^{i\phi}\sin(\theta/2) \end{pmatrix} \\ &\doteq \begin{pmatrix} e^{-i\omega_0 t/2}\cos(\theta/2) \\ e^{i\omega_0 t/2}e^{i\phi}\sin(\theta/2) \end{pmatrix} \\ &\doteq e^{-i\omega_0 t/2}\begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)}\sin(\theta/2) \end{pmatrix} \end{aligned} \quad . \tag{3.33}$$

If we again note that an overall phase does not have measureable effects, then the evolved state is simply a spin up eigenstate along a direction that has the same polar angle θ and a new azimuthal angle $\phi + \omega_0 t$. The state appears to have simply rotated around the z-axis, the axis of the magnetic field, by the angle $\omega_0 t$. Of course, we have to limit our discussion to results of measurements, so let's calculate the probability for measuring the spin projection along the z-axis first:

$$\begin{aligned} \mathcal{P}(+) &= \left|\langle + |\psi(t)\rangle\right|^2 \\ &= \left|\begin{pmatrix} 1 & 0 \end{pmatrix} e^{-i\omega_0 t/2}\begin{pmatrix} \cos(\theta/2) \\ e^{i(\phi+\omega_0 t)}\sin(\theta/2) \end{pmatrix}\right|^2 \\ &= \left|e^{-i\omega_0 t/2}\cos(\theta/2)\right|^2 \\ &= \cos^2(\theta/2) \end{aligned} \quad . \tag{3.34}$$

This is time independent since the S_z eigenstates are also energy eigenstates for this problem, and it is consistent with the interpretation that the angle θ that the spin vector makes with the z-axis does not change. The probability for measuring spin up along the x-axis is

$$\begin{aligned}
\mathcal{P}(+_x) &= \left|{}_x\langle +|\psi(t)\rangle\right|^2 \\
&= \left|\tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}e^{-i\omega_0 t/2}\begin{pmatrix}\cos(\theta/2)\\ e^{i(\phi+\omega_0 t)}\sin(\theta/2)\end{pmatrix}\right|^2 \\
&= \tfrac{1}{2}\left|\cos(\theta/2)+e^{i(\phi+\omega_0 t)}\sin(\theta/2)\right|^2 \\
&= \tfrac{1}{2}\left[\cos^2(\theta/2)+\cos(\theta/2)\sin(\theta/2)\left(e^{i(\phi+\omega_0 t)}+e^{-i(\phi+\omega_0 t)}\right)+\sin^2(\theta/2)\right] \\
&= \tfrac{1}{2}\left[1+\sin\theta\cos(\phi+\omega_0 t)\right]
\end{aligned} \quad .(3.35)$$

This is time dependent since the S_x eigenstates are not stationary states. The time dependence is consistent with the spin precessing around the z-axis. To further demonstrate this it is useful to calculate the expectation values for each of the spin components. For S_z, we have

$$\begin{aligned}
\langle S_z\rangle &= \langle\psi(t)|S_z|\psi(t)\rangle \\
&= e^{i\omega_0 t/2}\left(\cos(\theta/2)\quad e^{-i(\phi+\omega_0 t)}\sin(\theta/2)\right)\frac{\hbar}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}e^{-i\omega_0 t/2}\begin{pmatrix}\cos(\theta/2)\\ e^{i(\phi+\omega_0 t)}\sin(\theta/2)\end{pmatrix} \\
&= \frac{\hbar}{2}\left[\cos^2(\theta/2)-\sin^2(\theta/2)\right] \\
&= \frac{\hbar}{2}\cos\theta
\end{aligned} \quad , \quad (3.36)$$

while the other components are

$$\begin{aligned}
\langle S_x\rangle &= \langle\psi(t)|S_x|\psi(t)\rangle \\
&= e^{i\omega_0 t/2}\left(\cos(\theta/2)\quad e^{-i(\phi+\omega_0 t)}\sin(\theta/2)\right)\frac{\hbar}{2}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}e^{-i\omega_0 t/2}\begin{pmatrix}\cos(\theta/2)\\ e^{i(\phi+\omega_0 t)}\sin(\theta/2)\end{pmatrix} \\
&= \frac{\hbar}{2}\cos(\theta/2)\sin(\theta/2)\left[e^{i(\phi+\omega_0 t)}+e^{-i(\phi+\omega_0 t)}\right] \\
&= \frac{\hbar}{2}\sin\theta\cos(\phi+\omega_0 t)
\end{aligned} \quad , \quad (3.37)$$

and

$$\begin{aligned}\langle S_y\rangle &= \langle\psi(t)|S_y|\psi(t)\rangle \\ &= e^{i\omega_0 t/2}\begin{pmatrix}\cos(\theta/2) & e^{-i(\phi+\omega_0 t)}\sin(\theta/2)\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}e^{-i\omega_0 t/2}\begin{pmatrix}\cos(\theta/2)\\ e^{i(\phi+\omega_0 t)}\sin(\theta/2)\end{pmatrix} \\ &= \frac{\hbar}{2}\cos(\theta/2)\sin(\theta/2)\left[-ie^{i(\phi+\omega_0 t)}+ie^{-i(\phi+\omega_0 t)}\right] \\ &= \frac{\hbar}{2}\sin\theta\sin(\phi+\omega_0 t)\end{aligned} \quad . \tag{3.38}$$

The expectation value of the total spin vector $\langle\mathbf{S}\rangle$ is shown in Fig 3.1, where it is seen to precess around the magnetic field direction with an angular frequency ω_0. This frequency of precession is known as the **Larmor frequency**. Classically this is what one expects when a magnetic moment is placed in a uniform magnetic field.

A classical magnetic moment $\boldsymbol{\mu}$ experiences a torque $\boldsymbol{\mu}\times\mathbf{B}$ when placed in a magnetic field. If the magnetic moment is aligned with an angular momentum $\mathbf{L}$, then we can write

$$\boldsymbol{\mu} = \gamma\mathbf{L}, \tag{3.39}$$

where γ is the gyromagnetic ratio of the system. The equation of motion for the angular momentum

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\mu}\times\mathbf{B} \tag{3.40}$$

then results in

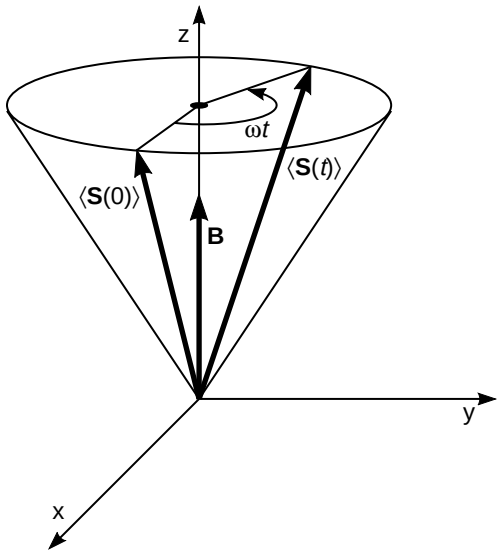

Figure 3.1. Expectation value of spin in magnetic field.

$$\frac{d\boldsymbol{\mu}}{dt} = \gamma\boldsymbol{\mu} \times \mathbf{B}, \tag{3.41}$$

Since the torque is perpendicular to the angular momentum, it causes the magnetic moment to precess about the field with an angular frequency $\omega_0 = \gamma B$.

In the quantum mechanical example we are considering, the gyromagnetic ratio is negative (meaning the spin and magnetic moment are anti-parallel), so the precession is counterclockwise around the field. A positive gyromagnetic ratio would result in clockwise precession. This precession makes it clear that the system has angular momentum, as opposed to simply a magnetic dipole.

Precession experiments like this are of great practical value. For example, if we measure the magnetic field strength and the precession frequency, then the gyromagnetic ratio can be determined. This spin precession problem is also of considerable theoretical utility since it is mathematically equivalent to many other quantum systems that can be modeled as two-state systems. This utility is broader than you might guess at first glance since many multi-state quantum systems can be reduced to two-state systems if the experiment is designed to only interact with two of the many levels of the system.

3.2.2 Magnetic field in general direction

For our second example, consider a more general direction for the magnetic field by adding a magnetic field component along the x-axis to the already existing field along the z-axis. The simplest approach to solving this new problem would be to redefine the coordinate system so the z-axis pointed along the direction of the new total magnetic field. Then the solution would be the same as was obtained above, with a new value for the magnitude of the magnetic field being the only change. This approach would be considered astute in many circumstances, but we will not take it since we want to get practice solving this new type of problem and we want to address some issues that are best posed in the original coordinate system. Thus we define a new magnetic field as

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_1\hat{\mathbf{x}}. \tag{3.42}$$

This field is oriented in the xz-plane at an angle θ with respect to the z-axis, as shown in Fig. 3.2. In light of the solution above, it it useful to define Larmor frequencies

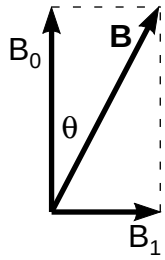

Figure 3.2. Magnetic field in general direction.

associated with each of the field components:

$$\omega_0 \equiv \frac{geB_0}{2m_e c} \qquad\qquad \omega_1 \equiv \frac{geB_1}{2m_e c}. \tag{3.43}$$

Using these definitions, the Hamiltonian becomes

$$\begin{aligned} H &= -\boldsymbol{\mu} \bullet \mathbf{B} \\ &= \omega_0 S_z + \omega_1 S_x, \end{aligned} \tag{3.44}$$

or in matrix representation

$$H \doteq \frac{\hbar}{2}\begin{pmatrix} \omega_0 & \omega_1 \\ \omega_1 & -\omega_0 \end{pmatrix}. \tag{3.45}$$

This Hamiltonian is not diagonal, so its eigenstates are not the same as the eigenstates of S_z. Rather we must use the diagonalization procedure to find the new eigenvalues and eigenvectors. The characteristic equation determining the eigenvalues is

$$\begin{vmatrix} \frac{\hbar}{2}\omega_0 - \lambda & \frac{\hbar}{2}\omega_1 \\ \frac{\hbar}{2}\omega_1 & -\frac{\hbar}{2}\omega_0 - \lambda \end{vmatrix} = 0 \qquad , \tag{3.46}$$

$$-\left(\frac{\hbar}{2}\omega_0\right)^2 + \lambda^2 - \left(\frac{\hbar}{2}\omega_1\right)^2 = 0$$

with solution

$$\lambda = \pm\frac{\hbar}{2}\sqrt{\omega_0^2 + \omega_1^2}\,. \tag{3.47}$$

Note that the eigenvalues are $\pm(\hbar\omega_0/2)$ when $\omega_1 = 0$, which they must be given our previous solution. Rather than solve directly for the eigenvectors, let's make them obvious by rewriting the Hamiltonian. From Fig. 3.2 it is clear that the angle θ is determined by the equation

$$\tan\theta = \frac{B_1}{B_0} = \frac{\omega_1}{\omega_0}. \tag{3.48}$$

Using this, the Hamiltonian can be written as

$$H \doteq \frac{\hbar}{2}\sqrt{\omega_0^2 + \omega_1^2}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}. \tag{3.49}$$

If we let $\hat{\mathbf{n}}$ be the unit vector in the direction of the total magnetic field, then the Hamiltonian is simply

$$H = \frac{\hbar}{2}\sqrt{\omega_0^2 + \omega_1^2}\; S_n.$$

This is what we expected at the beginning: that the problem could be solved by using the field direction to define a coordinate system. Thus the eigenvalues are as we found above and the eigenstates are the spin up and down states along the direction $\hat{\mathbf{n}}$, which are

$$\begin{aligned} |+\rangle_n &= \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}|-\rangle \\ |-\rangle_n &= \sin\frac{\theta}{2}|+\rangle - \cos\frac{\theta}{2}|-\rangle \end{aligned} \tag{3.50}$$

for this case since the azimuthal angle ϕ is zero. These are the same states you would find by directly solving for the eigenstates of the Hamiltonian. Since we have already done that for the S_n case earlier, we do not repeat it here.

Now consider performing the following experiment: begin with the system in the spin up state along the z-axis, and measure the spin projection along the z-axis after the system has evolved in this magnetic field for some time. Let's specifically calculate the probability that the initial $|+\rangle$ is later found to have evolved to the $|-\rangle$ state. This is commonly known as a **spin flip**. According to our time evolution prescription, we must first write the initial state in terms of the energy eigenstates of the system. In the examples above, this was trivial since the energy eigenstates were the $|\pm\rangle$ states that we used to express all general states. Since this new problem is more difficult, we will proceed more slowly. The initial state

$$|\psi(0)\rangle = |+\rangle \tag{3.51}$$

must be written in the $|\pm\rangle_n$ basis. Since this basis is complete, we can apply the closure relation to get

$$\begin{aligned} |\psi(0)\rangle &= \left[|+\rangle_{n\,n}\langle+| + |-\rangle_{n\,n}\langle-|\right]|+\rangle \\ &= |+\rangle_{n\,n}\langle+|+\rangle + |-\rangle_{n\,n}\langle-|+\rangle \\ &= {}_n\langle+|+\rangle|+\rangle_n + {}_n\langle-|+\rangle|-\rangle_n \\ &= \cos\frac{\theta}{2}|+\rangle_n + \sin\frac{\theta}{2}|-\rangle_n \end{aligned} \tag{3.52}$$

The time evolved state is found by multiplying each coefficient by a phase factor dependent on the energy of that eigenstate:

$$|\psi(t)\rangle = e^{-iE_+t/\hbar}\cos\frac{\theta}{2}|+\rangle_n + e^{-iE_-t/\hbar}\sin\frac{\theta}{2}|-\rangle_n. \tag{3.53}$$

We will leave it in this form and substitute the energy eigenvalues

$$E_\pm = \pm\frac{\hbar}{2}\sqrt{\omega_0^2+\omega_1^2} \tag{3.54}$$

at the end of the example.

The probability of a spin flip is

$$
\begin{aligned}
\mathcal{P}(+\to -) &= \left|\langle -|\psi(t)\rangle\right|^2 \\
&= \left|\langle -|\left[e^{-iE_+t/\hbar}\cos\frac{\theta}{2}|+\rangle_n + e^{-iE_-t/\hbar}\sin\frac{\theta}{2}|-\rangle_n\right]\right|^2 \\
&= \left|e^{-iE_+t/\hbar}\cos\frac{\theta}{2}\langle -|+\rangle_n + e^{-iE_-t/\hbar}\sin\frac{\theta}{2}\langle -|-\rangle_n\right|^2 \\
&= \left|e^{-iE_+t/\hbar}\cos\frac{\theta}{2}\sin\frac{\theta}{2} + e^{-iE_-t/\hbar}\sin\frac{\theta}{2}\left(-\cos\frac{\theta}{2}\right)\right|^2 . \qquad (3.55)\\
&= \cos^2\frac{\theta}{2}\sin^2\frac{\theta}{2}\left|1 - e^{i(E_+-E_-)t/\hbar}\right|^2 \\
&= \tfrac{1}{2}\sin^2\theta\left(1+\cos\frac{(E_+-E_-)t}{\hbar}\right) \\
&= \sin^2\theta\sin^2\left(\frac{(E_+-E_-)t}{2\hbar}\right)
\end{aligned}
$$

Again, this makes it clear that the probability oscillates at the frequency determined by the difference in energy of the eigenstates. This formula has important applications in many problems and is often called **Rabi's formula**. In the example at hand, the probability becomes

$$\mathcal{P}(+\to -) = \frac{\omega_1^2}{\omega_0^2+\omega_1^2}\sin^2\left(\frac{\sqrt{\omega_0^2+\omega_1^2}}{2}t\right). \qquad (3.56)$$

To gain insight into this formula, consider first the simple cases where the solution is obvious from our solution of the example in Sec. 3.2.1. For example, if there is no added field in the x-direction, then $\omega_1 = 0$ and $\mathcal{P}(+\to -) = 0$ since the initial state is a stationary state. In the other extreme, if there is no field component in the z-direction, then $\omega_0 = 0$ and $\mathcal{P}(+\to -)$ oscillates between 0 and 1 at the frequency ω_1, as shown in Fig. 3.3 (a). This corresponds to spin precession around the field in the x-direction, with a complete spin flip from $|+\rangle$ to $|-\rangle$ and back again occurring at the precession frequency ω_1. In the general case where both magnetic field components are present, the probability does not reach one and so there is no time at which the spin is certain to flip over. If the x–component of the field is small compared to the z-component, then $\omega_1 << \omega_0$ and $\mathcal{P}(+\to -)$ oscillates between 0 and a value much less than one at a frequency approximately equal to ω_0, as shown in Fig. 3.3 (b).

Now consider how we can relate this solution to other two-level systems to show the power and utility of this simple problem. We are interested in any two-state system with a Hamiltonian that, like Eq. (3.45), is not diagonal. To keep the problem conceptually similar to the spin problem, consider a system that starts in one of the basis kets and then ask for the probability that the other ket is measured after some time *t*.

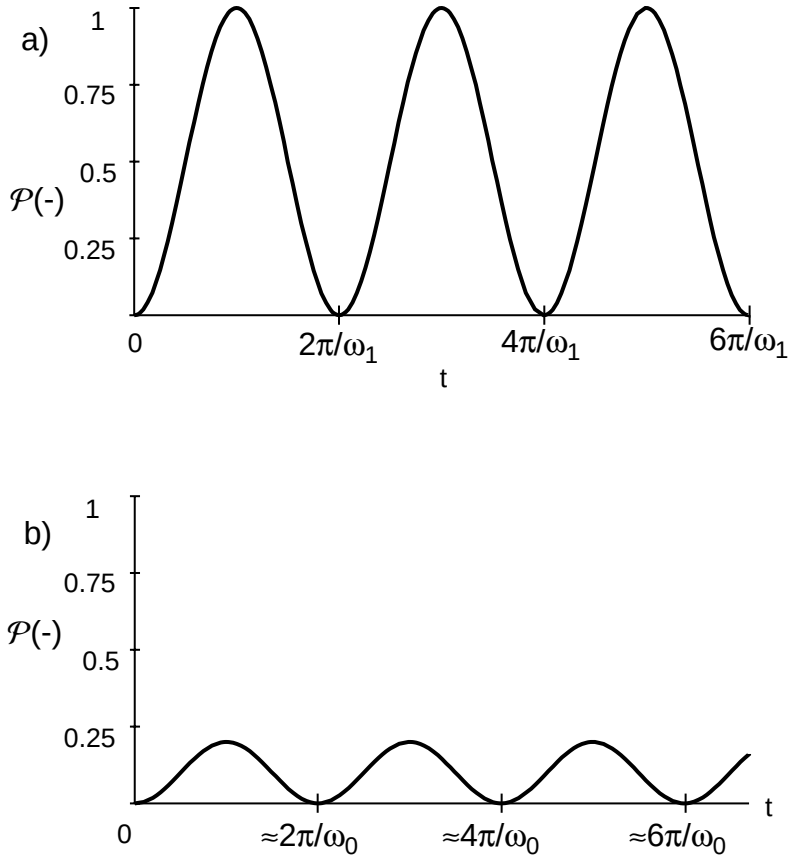

Figure 3.3. Spin flip probability for (a) field only in x-direction and (b) field with x- and z-components.

Since the Hamiltonian is not diagonal, these kets are not the same as the energy eigenstates. Hence the probability of measuring one of these states will oscillate.

3.3 Neutrino Oscillations

A modern example of this type of problem is provided by the phenomenon of neutrino oscillations. Neutrinos are uncharged, relativistic particles that are produced in weak interaction processes. In nuclear beta decay, neutrinos are produced in processes such as

$$\begin{aligned} n &\rightarrow p + e^- + \bar{\nu}_e \\ p &\rightarrow n + e^+ + \nu_e \end{aligned}, \tag{3.57}$$

where the subscript labels the neutrino as an electron neutrino and the bar labels $\bar{\nu}_e$ as an anti-neutrino. These neutrinos are produced at relativistic speeds and have little if any rest mass. They are often considered to be massless, like photons, but there is no compelling theoretical reason for them to be massless (as there is for photons). The mass

of a neutrino is thus an open experimental question, and has some interesting consequences. One of the reasons we don't yet know the neutrino mass is that neutrinos are very hard to detect. They can undergo reactions of the type

$$\begin{aligned} \nu_e + n &\rightarrow p + e^- \\ \bar{\nu}_e + p &\rightarrow n + e^+ \end{aligned}, \tag{3.58}$$

which means that a neutrino colliding with a nucleus will change the nucleus and produce an electron or positron. Unfortunately, the probabilities or cross-sections for these events are very small. As a result, we have large numbers of neutrinos passing through our bodies all the time with virtually no effect.

There are other types of neutrinos associated with other reactions, such as

$$\begin{aligned} \pi^+ &\rightarrow \mu^+ + \nu_\mu \\ \mu^- &\rightarrow e^- + \nu_\mu + \bar{\nu}_e \end{aligned}, \tag{3.59}$$

which represent the decay of a pion (π) to a muon (μ) and the decay of a muon to an electron, respectively. A muon behaves exactly like an electron but has a larger mass. Electrons, muons, and a third particle (tau) and their associated neutrinos are collectively called Leptons. In reactions involving these particles it is convenient to define a lepton flavor quantum number, with the assigned values $L_e = 1$ for the electron e^- and its associated neutrino ν_e, $L_e = -1$ for the positron e^+ and the antineutrino $\bar{\nu}_e$, $L_\mu = 1$ for the muon μ^- and its associated neutrino ν_μ, and $L_\mu = -1$ for the μ^+ and $\bar{\nu}_\mu$. With these assignments, the individual electron and muon flavor numbers are conserved in the processes shown above. However, there is no theoretical basis for this conservation, and so we allow for the possibility that these quantum numbers are only approximately conserved. This possibility then allows for reactions of the type

$$\nu_e \leftrightarrow \nu_\mu, \tag{3.60}$$

where an electron neutrino changes its flavor and becomes a muon neutrino, or the reverse. Such changes are often called **neutrino mixing** or **neutrino oscillations**.

The labeling of neutrinos according to their association with electrons or muons arises from their behavior in the weak processes described above. In other words, the quantum states $|\nu_e\rangle$ and $|\nu_\mu\rangle$ are eigenstates of the Hamiltonian describing the weak interaction. However, when neutrinos propagate in free space, the weak interaction is not relevant and the only Hamiltonian of relevance is that due to the relavisitic energy of the particles, which includes their rest mass and momenta. The eigenstates of this Hamiltonian are generally referred to as the mass eigenstates. If the masses of the two types of neutrinos (electron and muon) are different, then, in general, the mass eigenstates will not coincide with the weak interaction eigenstates. This distinction between sets of eigenstates allows for flavor changing processes.

To see why this is so, let the mass eigenstates be labeled $|\nu_1\rangle$ and $|\nu_2\rangle$. Either one of the two bases (mass or weak eigenstates) can be used as a complete basis upon

which to expand any general state in this system. Let's assume that the relation between the bases can be written as

$$\begin{aligned} |\nu_e\rangle &= \cos\frac{\theta}{2}|\nu_1\rangle + \sin\frac{\theta}{2}|\nu_2\rangle \\ |\nu_\mu\rangle &= \sin\frac{\theta}{2}|\nu_1\rangle - \cos\frac{\theta}{2}|\nu_2\rangle \end{aligned} . \tag{3.61}$$

This form allows us to use our results from above while still being general enough for the problem at hand. The angle θ/2 is generally referred to as the **mixing angle** (some treatments drop the factor 1/2, but we retain it to be consistent with the previous spin 1/2 discussion). If the mixing angle is small then the relations become

$$\begin{aligned} |\nu_e\rangle &\approx |\nu_1\rangle \\ |\nu_\mu\rangle &\approx |\nu_2\rangle \end{aligned} . \tag{3.62}$$

Assume that an electron neutrino is created in some weak interaction process and then propagates through free space to a detector. We wish to know the probability that a muon neutrino is detected, which is the signature of neutrino flavor mixing. Thus the initial state vector is

$$\begin{aligned} |\psi(0)\rangle &= |\nu_e\rangle \\ &= \cos\frac{\theta}{2}|\nu_1\rangle + \sin\frac{\theta}{2}|\nu_2\rangle \end{aligned} . \tag{3.63}$$

During the free-space propagation, the energy eigenstates of the system are the mass eigenstates since there is no weak interaction present. Thus the Schrödinger time evolution for this state is

$$|\psi(t)\rangle = \cos\frac{\theta}{2}e^{-iE_1t/\hbar}|\nu_1\rangle + \sin\frac{\theta}{2}e^{-iE_2t/\hbar}|\nu_2\rangle. \tag{3.64}$$

The energy eigenvalues are simply the relativistic energies, which are determined by the rest masses and the momenta:

$$E_i = \sqrt{(pc)^2 + \left(m_ic^2\right)^2} \quad , \qquad i = 1,2\,. \tag{3.65}$$

Assuming that the neutrinos are highly relativistic ($mc^2 << pc$), we find

$$\begin{aligned} E_i &= pc\left[1+\left(\frac{m_i c^2}{pc}\right)^2\right]^{1/2} \\ &\cong pc\left[1+\frac{1}{2}\left(\frac{m_i c^2}{pc}\right)^2\right]. \\ &\cong pc+\frac{\left(m_i c^2\right)^2}{2pc} \end{aligned} \tag{3.66}$$

The beauty of studying two-level systems such as spin 1/2 and neutrino oscillations is that they are formally identical. In the spin 1/2 case we phrased the problem in terms of finding the probability of a spin flip, whereas here we are looking for a change in the flavor of the neutrino. In both cases, the initial and final states are not energy eigenstates, but rather orthogonal states in a different basis. Since the problems are mathematically identical, the probability of a transition between the orthogonal states takes the same form. The probability of a neutrino oscillation is thus given by the same equation as the spin flip probability (Eq. (3.55))

$$\begin{aligned} \mathcal{P}(\nu_e \to \nu_\mu) &= \left|\left\langle \nu_\mu \middle| \psi(t)\right\rangle\right|^2 \\ &= \sin^2\theta \sin^2\left(\frac{(E_1-E_2)t}{2\hbar}\right), \end{aligned} \tag{3.67}$$

where the parameter θ has been defined the same in both problems and the energy difference $E_+ - E_-$ has been changed to the energy difference $E_1 - E_2$. This energy difference can be written as

$$\begin{aligned} E_1 - E_2 &= \frac{\left(m_1 c^2\right)^2}{2pc} - \frac{\left(m_2 c^2\right)^2}{2pc} \\ &= \frac{c^3}{2p}\left(m_1^2 - m_2^2\right) \end{aligned}. \tag{3.68}$$

Since the neutrinos move at nearly the speed of light *c*, we can approximate the time from the creation of the electron neutrino to the detection of the muon neutrino as $t \cong L/c$, where *L* is the distance from the source to the detector. This gives a probability for neutrino flavor change of

$$\mathcal{P}(\nu_e \to \nu_\mu) = \sin^2\theta \sin^2\left(\frac{\left(m_1^2 - m_2^2\right)Lc^2}{2p\hbar}\right). \tag{3.69}$$

As a function of the distance *L*, the probability oscillates from 0 to a maximum value of $\sin^2\theta$ - hence the term neutrino oscillations. By measuring the fractions of different neutrino flavors at a distance from a neutrino source (e.g., reactor, sun, supernova) and

comparing to a model for the expected fractions, experimenters thus hope to be able to infer the masses, or at least the mass differences, of the different neutrinos.

3.4 Magnetic Resonance

In the spin precession example in Sec. 3.2.2, we concluded that a complete spin flip required a large field in the x-direction. This represents a large change or perturbation compared to the initial situation of a field in the z-direction. Now consider whether we can induce a complete spin flip without such a large perturbation. That is, what small magnetic field can we add to the system that will cause a $|+\rangle$ state to oscillate to a $|-\rangle$ state? To answer this, it is instructive to consider the classical problem first.

A classical magnetic moment aligned with an angular momentum precesses around the direction of an applied magnetic field. Now imagine going to a reference frame that rotates about the field (assume z-direction) with the same frequency as the precession. An observer in that frame would see the magnetic moment stationary and so would conclude that there is no magnetic field in that frame. If that rotating observer were asked to flip the magnetic moment from up to down along the z-axis, she would answer "Simple, just impose a small magnetic field perpendicular to the z-axis, which will cause the spin to precess around that direction." Since that field is the only field acting in that frame, it can be as small as one likes. The magnitude simply determines the time for the spin to flip.

In this situation the transverse applied field is stationary in the rotating frame, so it will appear to be rotating at the precessional frequency in the original frame. Thus we could write it as

$$\mathbf{B} = B_1 \cos(\omega t)\hat{\mathbf{x}} + B_1 \sin(\omega t)\hat{\mathbf{y}}, \tag{3.70}$$

where we allow the frequency ω to differ from the precessional frequency ω_0 in order to solve the problem more generally. In that case, there would be some residual precession in the rotating frame and so the rotating observer would conclude that there is some residual field in the z-direction. Hence, we expect that the added transverse field would not cause a complete flipping of the magnetic moment from up to down in this general case.

Let's now apply this reasoning to the quantum mechanical case. Assume a magnetic field of the form

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_1\left[\cos(\omega t)\hat{\mathbf{x}} + \sin(\omega t)\hat{\mathbf{y}}\right], \tag{3.71}$$

which results in a Hamiltonian

$$\begin{aligned} H &= -\boldsymbol{\mu} \bullet \mathbf{B} \\ &= \omega_0 S_z + \omega_1\left[\cos(\omega t)S_x + \sin(\omega t)S_y\right]. \end{aligned} \tag{3.72}$$

We again define the Larmor frequencies corresponding to the two magnetic field components,

$$\omega_0 \equiv \frac{geB_0}{2m_e c} \qquad\qquad \omega_1 \equiv \frac{geB_1}{2m_e c}. \tag{3.73}$$

The matrix representation of the Hamiltonian becomes

$$H \doteq \frac{\hbar}{2}\begin{pmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{pmatrix}. \tag{3.74}$$

This Hamiltonian is time dependent, so we can no longer use our simple prescription for Schrödinger time evolution. Rather, we must return to the Schrödinger equation and solve it with these new time-dependent terms. If we write the state vector as

$$|\psi(t)\rangle = c_+(t)|+\rangle + c_-(t)|-\rangle \doteq \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix}, \tag{3.75}$$

then Schrödinger's equation

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle \tag{3.76}$$

leads to

$$\begin{aligned} i\hbar \dot{c}_+(t) &= \frac{\hbar\omega_0}{2} c_+(t) + \frac{\hbar\omega_1}{2} e^{-i\omega t} c_-(t) \\ i\hbar \dot{c}_-(t) &= \frac{\hbar\omega_1}{2} e^{i\omega t} c_+(t) - \frac{\hbar\omega_0}{2} c_-(t) \end{aligned}, \tag{3.77}$$

where $\dot{c}_+(t)$ denotes a time derivative. To solve these time-dependent coupled differential equations, it is useful to follow the lead of the classical discussion and consider the problem from the rotating frame. Though we don't yet have the complete tools to know how to effect this transformation, we will take it on faith that after a frame transformation the state vector can be written as

$$|\tilde{\psi}(t)\rangle = c_+(t)e^{i\omega t/2}|+\rangle + c_-(t)e^{-i\omega t/2}|-\rangle \doteq \begin{pmatrix} c_+(t)e^{i\omega t/2} \\ c_-(t)e^{-i\omega t/2} \end{pmatrix}, \tag{3.78}$$

where $|\tilde{\psi}(t)\rangle$ is the state vector as viewed from the rotating frame. If we call the coefficients of this vector $\alpha_\pm(t)$, then we can write

$$|\tilde{\psi}(t)\rangle = \alpha_+(t)|+\rangle + \alpha_-(t)|-\rangle \doteq \begin{pmatrix} \alpha_+(t) \\ \alpha_-(t) \end{pmatrix}, \tag{3.79}$$

where the relations between the sets of coefficients are

$$\begin{aligned} c_+(t) &= e^{-i\omega t/2}\alpha_+(t) \\ c_-(t) &= e^{i\omega t/2}\alpha_-(t) \end{aligned}. \tag{3.80}$$

The state vector in the non-rotating frame can thus be written as

$$|\psi(t)\rangle = \alpha_+(t)e^{-i\omega t/2}|+\rangle + \alpha_-(t)e^{i\omega t/2}|-\rangle \doteq \begin{pmatrix} \alpha_+(t)e^{-i\omega t/2} \\ \alpha_-(t)e^{i\omega t/2} \end{pmatrix}. \tag{3.81}$$

Another way of viewing this transformation is to say that based upon earlier solutions of similar problems (Eq. (3.33)), we expect the coefficients $c_\pm(t)$ to have time dependence of the form $e^{\mp i\omega t/2}$ and so we have extracted that part of the solution and now need to solve for the remaining time dependence in the coefficients $\alpha_\pm(t)$. In this view, we have simply performed a mathematical trick to make the solution easier.

If we now substitute the expressions for $c_\pm(t)$ in terms of $\alpha_\pm(t)$ into the differential equations, we get

$$\begin{aligned} i\hbar\dot{\alpha}_+(t) &= -\frac{\hbar\Delta\omega}{2}\alpha_+(t) + \frac{\hbar\omega_1}{2}\alpha_-(t) \\ i\hbar\dot{\alpha}_-(t) &= \frac{\hbar\omega_1}{2}\alpha_+(t) + \frac{\hbar\Delta\omega}{2}\alpha_-(t) \end{aligned}, \tag{3.82}$$

where we have defined a new term

$$\Delta\omega \equiv \omega - \omega_0, \tag{3.83}$$

which is the difference between the angular frequencies of the rotating field and the Larmor precession due to the z-component of the magnetic field. Since $\alpha_\pm(t)$ are the coefficients of the transformed state vector $|\tilde{\psi}(t)\rangle$, these differential equations can be considered as comprising the Schrödinger equation

$$i\hbar\frac{d}{dt}|\tilde{\psi}(t)\rangle = \tilde{H}|\tilde{\psi}(t)\rangle, \tag{3.84}$$

where the new Hamiltonian $\tilde{H}$ has the matrix representation

$$\tilde{H} \doteq \frac{\hbar}{2}\begin{pmatrix} -\Delta\omega & \omega_1 \\ \omega_1 & \Delta\omega \end{pmatrix}. \tag{3.85}$$

Thus we have transformed (by rotation or mathematical sleight of hand) the original problem into a new problem that has a time-independent Hamiltonian. Once we solve the new problem, we can use the transformation equations to find the solution to the original problem. However, since the new Hamiltonian $\tilde{H}$ is time-independent, we already know the solution. That is, this new problem has the same Hamiltonian as the spin precession problem in Sec. 3.2.2, except that the term ω_0 is replaced by the new term -$\Delta\omega$. We are interested in finding the same probability $\mathcal{P}(+\to-)$ that an initial $|+\rangle$ state is later found to have evolved to the $|-\rangle$ state. The rotational transformation does not alter the $|\pm\rangle$ basis states so if

$$|\psi(0)\rangle = |+\rangle, \tag{3.86}$$

then

$$|\tilde{\psi}(0)\rangle = |+\rangle. \tag{3.87}$$

The probability for a spin flip is given by

$$\begin{aligned}\mathcal{P}(+\rightarrow -) &= \left|\langle -|\psi(t)\rangle\right|^2 \\ &= \left|c_-(t)\right|^2\end{aligned}. \tag{3.88}$$

The equations relating the coefficients can be used to write

$$\begin{aligned}\left|c_-(t)\right|^2 &= \left|e^{-i\omega t/2}\alpha_-(t)\right|^2 \\ &= \left|\alpha_-(t)\right|^2 \\ &= \left|\langle -|\tilde{\psi}(t)\rangle\right|^2\end{aligned}, \tag{3.89}$$

which means that the probability we desire is

$$\mathcal{P}(+\rightarrow -) = \left|\langle -|\tilde{\psi}(t)\rangle\right|^2. \tag{3.90}$$

Thus we can use the spin precession solution we found above (Eq. (3.56)), with the change ω_0 –> $-\Delta\omega$, resulting in

$$\begin{aligned}\mathcal{P}(+\rightarrow -) &= \frac{\omega_1^2}{\Delta\omega^2+\omega_1^2}\sin^2\left(\frac{\sqrt{\Delta\omega^2+\omega_1^2}}{2}t\right) \\ &= \frac{\omega_1^2}{(\omega-\omega_0)^2+\omega_1^2}\sin^2\left(\frac{\sqrt{(\omega-\omega_0)^2+\omega_1^2}}{2}t\right)\end{aligned}. \tag{3.91}$$

This is the famous Rabi flopping equation and has important applications all throughout physics.

The probability of a spin flip oscillates with an angular frequency given by

$$\Omega = \sqrt{(\omega-\omega_0)^2+\omega_1^2}, \tag{3.92}$$

that is typically referred to as the **generalized Rabi frequency**. The term **Rabi frequency** generally refers to the frequency ω_1, which is the value of the generalized Rabi frequency when the frequency ω of the rotating field is we set equal to the Larmor precession frequency ω_0 of the system in the presence of the magnetic field B_0 alone. For this choice of ω, the probability of a spin flip becomes

$$\mathcal{P}(+\rightarrow -) = \sin^2\left(\frac{\omega_1}{2}t\right), \tag{3.93}$$

which implies that the spin is flipped with 100% probability at an angular frequency ω_1. For other choices of the frequency ω, the probability of a spin flip will oscillate with an amplitude smaller than one. The amplitude of the spin flip oscillation, as a function of the frequency ω of the rotating field, is plotted in Fig. 3.4. This curve is generally referred to as a **Lorentzian** curve and clearly exhibits the important **resonant** behavior of

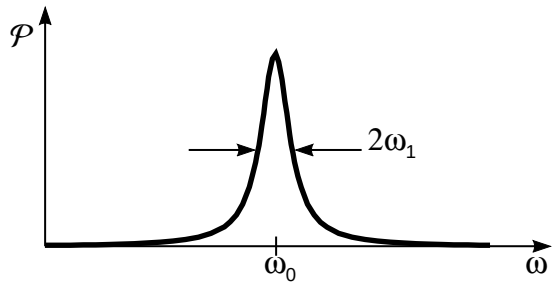

Figure 3.4. Magnetic resonance curve.

the spin flip probability. The full width at half maximum (**FWHM**) of the resonance curve is $2\omega_1$.

For the resonance condition $\omega = \omega_0$, the probability of a spin flip as a function of time is plotted in Fig. 3.5. Since the frequency ω_1 is proportional to the applied field B_1, the rate of spin flipping increases with increasing field. However, it is important to note that there is still 100% probability of a spin flip for very small fields. This is the property we were looking for at the beginning of the problem – a way to flip the spin without perturbing the system appreciably. After a time t given by $\omega_1 t = \pi$, the probability for a spin flip is 100%. We have assumed that the applied field is on continuously, but this spin flip can also be produced by a pulsed field with a magnitude and duration that satisfy $\omega_1 t = \pi$. Such a pulse is often called a π-pulse and is used to flip a spin, or more generally to make a transition from one energy state to another with 100% certainty. The diagram on the right of Fig. 3.5 illustrates the energy levels of the spin in the magnetic field and how transitions between the levels can be associated with the spin flip

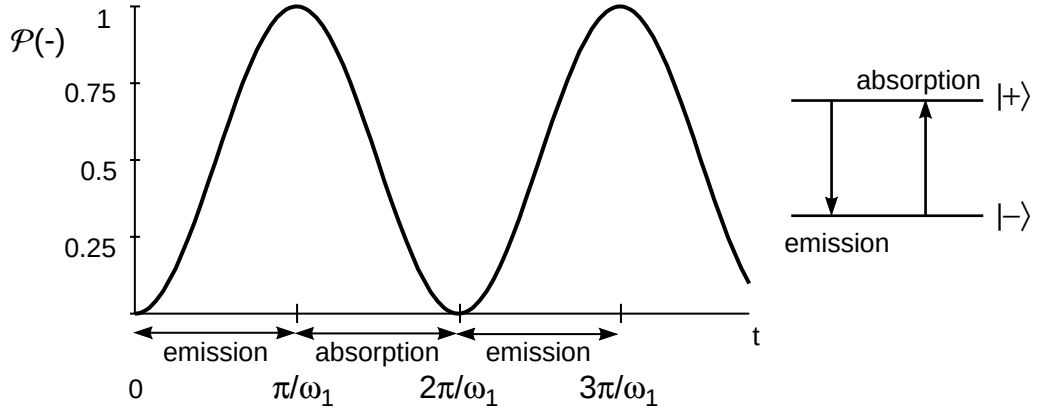

Figure 3.5. Rabi oscillations for resonance condition.

oscillation. A transition from an upper level to a lower level takes energy from the atom and gives it to the magnetic field and is known as **emission**, while the opposite process takes energy from the field and is known as **absorption**. This same picture of the interaction between a two-level system and a field is used to explain how atoms absorb and emit light.

Problems

3.1 Consider a simple two-state quantum system with a Hamiltonian

$$H \doteq \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}.$$

Another physical observable A is described by the operator

$$A \doteq \begin{pmatrix} 0 & a \\ a & 0 \end{pmatrix},$$

where a is real and positive. Let the initial state of the system be $|\psi(0)\rangle = |a_1\rangle$, where $|a_1\rangle$ is the eigenstate corresponding to the larger of the two possible eigenvalues of A. What is the frequency of oscillation (*i.e.*, the Bohr frequency) of the expectation value of A?

3.2 (Townsend 4.5) A beam of identical neutral particles with spin 1/2 is prepared in the $|+\rangle$ state. The beam enters a uniform magnetic field B_o, which is in the x-z plane oriented at an angle θ with respect to the z-axis. At a time T later, the beam enters a Stern-Gerlach analyzer oriented along the y-axis. What is the probability that particles will be detected with $S_y = \hbar/2$? Check your result by evaluating the special cases $\theta = 0$ and $\theta = \pi/2$.

3.3 (Goswami 16.6) Consider a spin 1/2 particle. At time $t = 0$, the particle is in the state $|+\rangle$.

a) If S_x is measured at $t = 0$, what is the probability of getting a value $\hbar/2$?

b) Suppose instead of performing the above measurement, the system is allowed to evolve in a magnetic field $\mathbf{B} = B_0\hat{\mathbf{y}}$. Using the S_z basis, calculate the state of the system after a time t.

c) At time t, suppose we measure S_x; what is the probability that a value $\hbar/2$ will be found?

3.4 (Goswami 16.7) Consider a spin 1/2 particle. At time $t = 0$, the particle is in the state $|+\rangle$.

a) At time $t = 0$, we measure S_x and find a value $\hbar/2$. What is the state vector immediately after the measurement?

b) At the same instant of the measurement, we apply a magnetic field $\vec{\mathbf{B}} = B_0\hat{\mathbf{z}}$ on the particle and allow the particle to precess for a time T. What is the state of the system at $t = T$?

c) At $t = T$, the magnetic field is very rapidly rotated so that it is now $\vec{\mathbf{B}} = B_0\hat{\mathbf{y}}$. After another time interval T, a measurement of S_x is carried out once more. What is the probability that a value $\hbar/2$ is found?

3.5 (Townsend 4.4) A beam of identical neutral particles with spin 1/2 travels along the y-axis. The beam passes through a series of two Stern-Gerlach (SG) spin analyzing magnets, each of which is designed to analyze the spin projection along the z-axis. The first Stern-Gerlach analyzer only allows particles with spin **up** (along the z-axis) to pass through. The second Stern-Gerlach analyzer only allows particles with spin **down** (along the z-axis) to pass through. The particles travel at speed v_0 between the two analyzers, which are separated by a region of length ℓ_0 in which there is a uniform magnetic field B_0 pointing in the x-direction. Determine the smallest value of ℓ_0 such that only 25% of the particles transmitted by the first analyzer are transmitted by the second analyzer.

3.6 (Townsend 4.13) Let the matrix representation of the Hamiltonian of a three-state system be

$$H \doteq \begin{pmatrix} E_0 & 0 & A \\ 0 & E_1 & 0 \\ A & 0 & E_0 \end{pmatrix}$$

using the basis states $|1\rangle$, $|2\rangle$, and $|3\rangle$.

a) If the state of the system at time $t = 0$ is $|\psi(0)\rangle = |2\rangle$, what is the probability that the system is in state $|2\rangle$ at time t?

b) If, instead, the state of the system at time $t = 0$ is $|\psi(0)\rangle = |3\rangle$, what is the probability that the system is in state $|3\rangle$ at time t?

3.7 In part 3 of Spins Lab #3 you built a spin-1 interferometer and measured the relative probabilities after the final SG device for the seven possible cases where one beam, a pair of beams, or all three beams from the middle SG device were used. Show how you used the projection postulate to calculate the theoretical probabilities.

BIBLIOGRAPHY

R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics, Volume 3, Quantum Mechanics*, Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1965.

J. J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley Publishing Company, Inc., Redwood City, California, 1985.

C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*, John Wiley & Sons, New York, 1977.

J. S. Townsend, *A Modern Approach to Quantum Mechanics*, McGraw Hill, Inc., New York, 1992.