I. Introduction

1 Scope and Nature of Classical Physics

The ultimate aim of classical mechanics is to predict the trajectories of objects given that the nature of the interactions between them and their environments is known and that the initial states of the objects are provided. In introductory and intermediate-level physics courses, one often focuses on one or two objects and attempts to find their positions as functions of time by using Newton's laws.

Example: Consider a block with known mass that is free to move across a frictionless horizontal surface. The block is connected to one end of a spring of known stiffness and the other end of the spring is attached to an immovable wall.

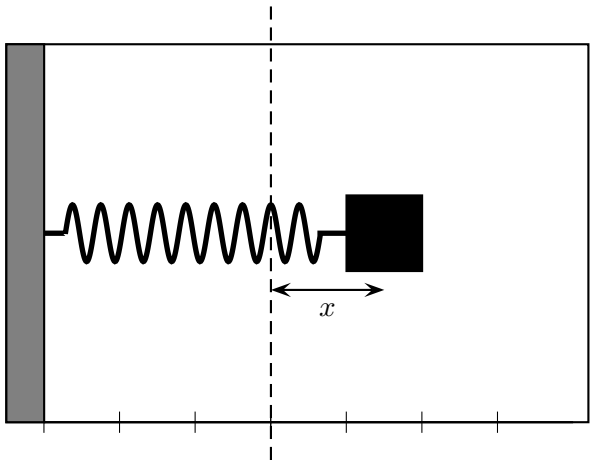


Figure I.1.1: Spring and mass system on a horizontal surface viewed from above. The dashed line indicates the equilibrium position of the center of the block.

Suppose that the mass of the block is m, the spring constant is k and x denotes the displacement of the center of the block from equilibrium. Newton's second law applied to an object that moves in one dimension gives

$$\mathbf{F}_{\text{net}} = m\mathbf{a} \tag{I.1.1}$$

which gives

$$-kx = \frac{d^2x}{dt^2} \tag{I.1.2}$$

where t is time. Via this route, the problem of determining the trajectory (i.e. the dependence of x on t) of a physical system has been reduced to the mathematical problem of solving a second order differential equation. Fortunately the type of differential equation that appears in Eq. (I.1.2) can be solved fairly easily. The general solution is

$$x(t) = x(0)\,\cos\omega t + \frac{v(0)}{\omega}\,\sin\omega t \tag{I.1.3}$$

where

$$\omega = \sqrt{\frac{k}{m}} \tag{I.1.4}$$

and the two constants $x(0)$ and $v(0)$ have the physical interpretation of the initial position and velocity respectively. The graph of position as a function of time is sometimes called a trajectory. Fig. I.1.2 illustrates an example of this for the block connected to a spring.

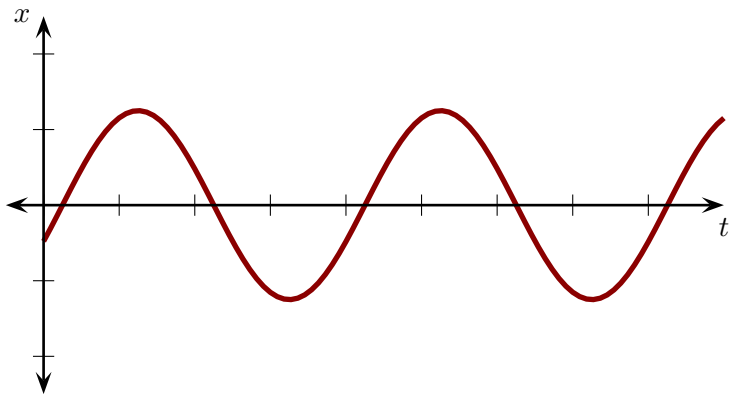


Figure I.1.2: Position vs. time graph for a block connected to a spring.

With a little insight it is clear that such a plot of position vs. time indicates oscillatory motion and, in this sense, the trajectory offers a fairly concrete description of the block's motion.

In the example of the spring and block system, we see that Newton's laws eventually give an expression for the trajectory of the block. All other physically interesting quantities, such as velocity, momentum and energy, that pertain to the block can be derived from the equation for its trajectory.

Exercise: Show that the velocity of the block, as a function of time, is given by

$$v(t) = -\omega x(0)\,\sin\omega t + v(0)\,\cos\omega t$$

and the total mechanical energy is given by

$$E = \frac{1}{2}\,m\omega^2[x(0)]^2 + \frac{1}{2}\,m[v(0)]^2.$$

The general scheme of classical mechanics is:

> Given the forces acting on an object of known mass, Newton's second law can be applied to eventually determine a complete expression for the position of the object as a function of time.

Needless to say, this is frequently only true in principle. In practice the equations of motion, analogous to Eq. (I.1.2), for most systems are too complicated to admit known solutions. In many cases involving just one object, it is possible to make some progress in understanding the motion of a system by abandoning the quest for complete knowledge of the trajectory of an object and rather considering speeds and energies. For situations involving many objects, such as a gas, it is uninteresting to focus on any individual object; the detailed description promised by Newton's laws must be abandoned for descriptions involving averages that appear in thermodynamics and statistical mechanics. Nevertheless, in principle Newton's laws could be applied to determine the complete details of all the objects' motions provided that the interactions between them are well known.

Newton's mechanics and the theories derived from it are tremendously successful in describing a vast array of physical phenomena. However, by the early 1900's careful application of Newton's laws contradicted the observed phenomena in a variety of physical situations, typically dealing with physics at the atomic and molecular level. Some of these such as blackbody radiation, the photoelectric effect, the Compton effect, the stability of an atom, atomic spectra, and the Frank-Hertz effect are often described in detail in sophomore-level modern physics courses. A coherent description, which entailed radical departures from Newton's mechanics, of these and other physical phenomena emerged in the period from 1900 to the 1920's. By the late 1920's a new theory, capable of describing all of these and more very accurately had emerged; this became known as *quantum mechanics* or *quantum theory.*

2 Scope and Nature of Quantum Theory

Quantum theory is a general purpose theory which claims to provide a framework for understanding all physical phenomena, in the same way that, prior to 1900, Newton's laws were thought to describe the physical world. That is, it provides a set of mathematical rules which ultimately allow for a description of any physical system.

Example: An ion, of charge $+q$, can be placed in a one dimensional electrostatic potential which is described, to a good approximation, by

$$V = \frac{1}{2}\frac{k}{q}x^2 \tag{I.2.5}$$

where k is a constant with units of $\mathrm{V/m^2}$. In classical mechanics, the components of the force exerted on the ion as a result of this are determined via

$$F_x = -q\frac{\partial V}{\partial x} \tag{I.2.6}$$

and it is easily seen that, in this case the x component of the force is

$$F_x = -kx. \tag{I.2.7}$$

This is mathematically identical to the situation of the block and spring and is merely another example of a simple harmonic oscillator. In classical mechanics, the task would be to determine the trajectory of the ion. In quantum mechanics, it will emerge that there is no sense in trying to do this. In fact, the concept of force is inapplicable and so is Eq. (I.2.6) and everything that follows from it.

The key mathematical entity in the quantum mechanical description of this type of particle is the wavefunction, $\Psi(x,t)$, which can produce complex values. One example of a wavefunction for the ion in the above potential, or any harmonic oscillator, turns out to be as illustrated in Fig. I.2.3.

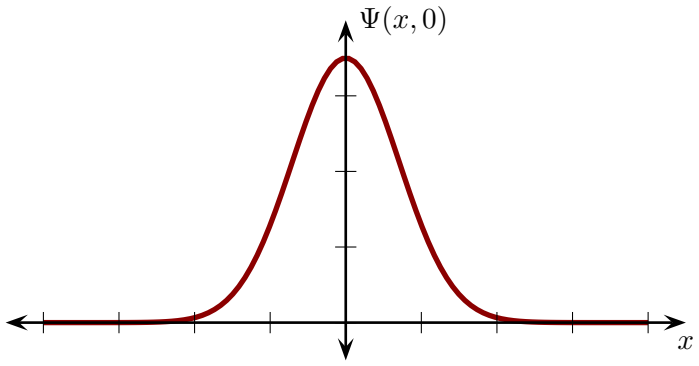


Figure I.2.3: One possible wavefunction for a harmonic oscillator at $t = 0$.

The means by which one can arrive at a sensible wavefunction for a simple harmonic oscillator are typically described in a sophomore-level modern physics course. What is important to note at this stage, is that *the wavefunction does not represent a trajectory of a particle as time passes.* In fact the wavefunction as illustrated in Fig. I.2.3 is that at *one particular instant.* It therefore cannot resemble any kind of evolution of the ion's physical state. What then, is its use in physical terms?

In general the state of a particle that is free to move in one dimension is described in quantum theory by a wavefunction $\Psi(x,t)$. This is useful for determining the outcomes and the associated statistics of measurements. For example,

$$\text{Prob(position measurement gives outcome from } x_0 \text{ to } x_0 + \mathrm{d}x) = |\Psi(x_0,t)|^2\,\mathrm{d}x \tag{I.2.8}$$

and there are a variety of mathematical rules for predicting the outcomes of measurements of other physical quantities statistically. This is really the only interpretation of the wavefunction. Quantum theory provides the mathematical rules which enable one to determine, in such cases, suitable wavefunctions and to use these to determine the outcomes

of measurements. The wavefunction is then more of a mathematical tool than a means for visualizing the physical evolution that the trajectory offers. In this sense quantum theory is more abstract than classical physics.

Quantum theory has been enormously successful in describing the physical world, particularly at small scales. Quantum theory features prominently in the description of, amongst others,

1. atoms and molecules and their interactions with light,
2. subatomic particles (high-energy physics, particle physics),
3. light at the low photon levels (quantum optics),
4. the properties of solids, including thermal and electrical properties (solid state physics),
5. superconductivity, superfluidity, and
6. quantum information.

II. Quantum Mechanics of Spin-1/2

The fundamental features of quantum theory can all be illustrated in the context of a physical system known as a spin-1/2 system. The fundamental measurable quantities for such systems are components of spin angular momentum. For spin-1/2 systems the usual classical variables such as position and momentum are irrelevant and such systems generally have no classical analogs. However, the mathematics for these is simpler than for particles with position degrees of freedom and it is easier to establish the connections between the fundamental mathematical building blocks of the theory and the physical properties of the systems.

There are many examples of spin-1/2 systems; these include electrons, protons, and various atomic nuclei. The entire understanding of nuclear magnetic resonance (NMR) is based on describing them via quantum theory. Also, there are many physical systems for which the mathematical description is similar to that of spin-1/2 systems; examples include photon polarization states, neutrino flavors, etc,....

1 Spin angular momentum and magnetic dipole moments

Spin angular momentum typically refers to the angular momentum of a rotating rigid object. The simplest example of this is a sphere of uniform mass density rotating about an axis. In classical mechanics, this can be regarded as a collection of particles, each orbiting with the same angular velocity about the same axis, as illustrated in Fig. II.1.1.

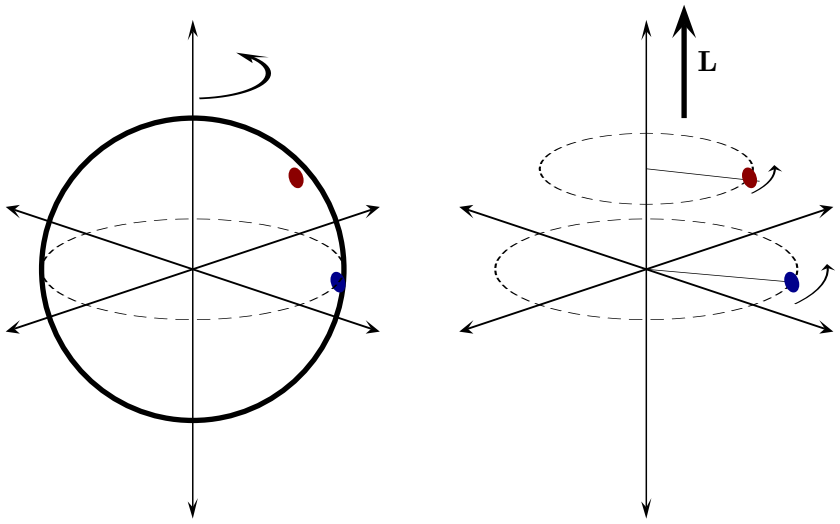


Figure II.1.1: Rigidly rotating sphere and orbiting particles on the surface of the sphere. The direction of the angular momentum is indicated via **L**.

The (orbital) angular momentum of a particle of mass m about the origin of the coordinate system is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

where **r** is the vector from the origin to the particle's location and **p** is its linear momentum. For a particle undergoing uniform circular motion with speed v, the magnitude of the angular momentum is

$$L = mvr \tag{II.1.1}$$

where r is the radius of orbit. The direction of **L** is along the axis of rotation. The total angular momentum of the entire solid object is obtained by adding or integrating over the angular momenta of its constituent particles.

Exercise: Show that Eq. (II.1.1) is valid.

One can envisage measuring the orbital angular momentum of an extended object by determining each of the quantities on the right hand side of Eq. (II.1.1) and calculating L accordingly. However, *if the particle is charged* there is an indirect method for determining **L** which is suitable for objects on the molecular or smaller scales. This uses the facts that a moving charged particle, in effect creates an electric current and this current produces a magnetic field. For particles which move in circular orbits, the magnetic field produced by the particle and the interactions with external magnetic fields are predominantly described in terms of the particle's *magnetic dipole moment*. For a loop of current of magnitude I that lies in a plane, the magnitude of the magnetic dipole moment is

$$\mu = IA \tag{II.1.2}$$

where A is the area of the loop. This is converted into a vector, $\boldsymbol{\mu}$, by stipulating that the direction of the magnetic dipole moment is perpendicular to the loop and oriented using the right hand rule. This is illustrated in Fig. II.1.2.

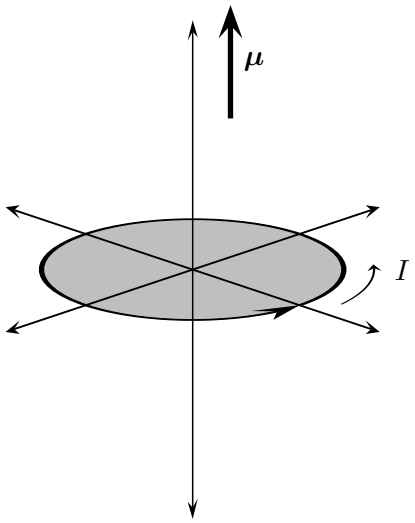


Figure II.1.2: Orbiting loop of positive charge and the current loop produced. The area of the shaded loop is A. The direction of the angular momentum is indicated via $\boldsymbol{\mu}$.

A single charged particle, with charge q, which orbits at a constant rate, effectively establishes a current of magnitude

$$I = \frac{|q|}{T}$$

where T is the period of orbit.

Exercise: Show that the current produced a particle with charge q and moving in a circle of radius r and at speed v has magnitude

$$I = \frac{|q|v}{2\pi r}$$

and show that this gives a magnetic dipole moment of magnitude

$$\mu = \frac{|q|vr}{2}. \tag{II.1.3}$$

Eqs. (II.1.3) and (II.1.1) yield

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

and taking directions into account for both positive and negative charges gives

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

and for an extended object, whose mass density is proportional to its charge density, superpositions of each side give

$$\boxed{\boldsymbol{\mu} = \frac{Q}{2M} \mathbf{L}} \tag{II.1.4}$$

where Q is the total charge and M the total mass of the object. For a rigid rotating object the total angular momentum is called the *spin angular momentum* and is denoted $\mathbf{S}$. If the mass and charge distributions are not proportional (e.g. the mass is uniformly distributed throughout the object but the charge just resides on the surface), the relationship between the magnetic dipole moment and the spin angular momentum becomes

$$\boxed{\boldsymbol{\mu} = g\frac{Q}{2M} \mathbf{S}} \tag{II.1.5}$$

where g is a dimensionless constant called the gyromagnetic or g-factor. This depends on the relative arrangement of the charge and mass distributions and its value can thus give some insight into the structure of the solid object. The same relationship extends to individual charge particles and with the mass and charge represented by m and q respectively, giving

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

The interaction between a magnetic dipole and an external magnetic field, $\mathbf{B}$, is determined via the potential energy of the dipole in the field,

$$V = -\boldsymbol{\mu} \cdot \mathbf{B} \tag{II.1.6}$$

and the general prescription for determining the force exerted on the dipole as a result of this potential energy,

$$\mathbf{F} = -\boldsymbol{\nabla} V \tag{II.1.7}$$

giving the following for components of the force exerted by the field on the magnetic dipole,

$$F_x = -\frac{\partial V}{\partial x}, \tag{II.1.8a}$$

$$F_y = -\frac{\partial V}{\partial y} \quad \text{and} \tag{II.1.8b}$$

$$F_z = -\frac{\partial V}{\partial z}. \tag{II.1.8c}$$

Since the dipole moment does not depend on position coordinates, the force exerted by a magnetic field only depends on the variation of the magnetic field with position. Thus, there will only be a net force on a dipole placed in a magnetic field if that field varies with position. However, the magnitude of the force will depend on the magnitude of the dipole moment. These facts are exploited in the Stern-Gerlach experiment.

2 Stern-Gerlach Experiment

2.1 Stern-Gerlach experiment setup

In the Stern-Gerlach experiment, first performed by Otto Stern and Walter Gerlach in 1922 in Frankfurt, a beam of particles is fired through a region which contains an inhomogeneous magnetic field (i.e. the field depends on position). In it's simplest conception, the field has, in a restricted region, the following form

$$\mathbf{B} = B(z)\hat{\boldsymbol{z}}.$$

The set up is illustrated conceptually in Fig. II.2.3.

While the particles are in the region of non-zero field the net force exerted on them is determined from Eq. (II.1.8) and gives

$$\mathbf{F}_{\text{net}} = \mu_z \frac{\partial B}{\partial z}\hat{\boldsymbol{z}}. \tag{II.2.1}$$

In the simplest realization $B(z) = B_0 z$ where B_0 is constant and, in this case, a classical analysis, using Eq. (II.2.1), predicts that the deflection of the particles d on the screen is proportional to μ_z. Thus measuring the deflection is equivalent to measuring μ_z. Despite the fact that the Stern-Gerlach experiment measures magnetic dipole moment, it is conventional to regard it as a measurement of a component of the particle's spin angular momentum. This follows since the spin angular momentum of the particle, $\mathbf{S}$, is related to the magnetic dipole moment via

$$\boldsymbol{\mu} = g\frac{q}{2m}\,\mathbf{S} \tag{II.2.2}$$

where q is the charge of the particle, m its mass and g the g-factor for the particle. The screen can then be calibrated so that the location at which a particle hits the screen is

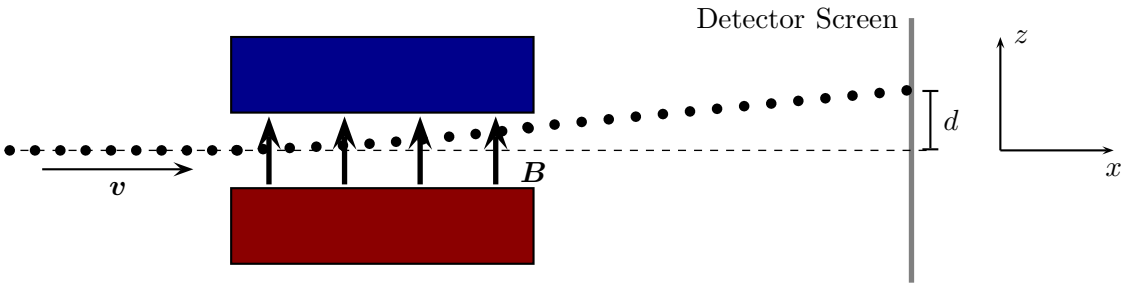


Figure II.2.3: Stern-Gerlach experimental setup. The two magnets produce an inhomogeneous magnetic field with constant direction in the gap between them and which is approximately zero beyond the gap. Particles fired into the gap with identical magnetic dipole moments are deflected by the same amount d.

represented by S_z (rather than d). It is typical to consider this entire experiment as nothing other than a measurement of S_z and the Stern-Gerlach apparatus as a merely a device for measuring S_z.

Typically the beam of particles which is fired through the gap is prepared in such a way that there is no inherently preferred direction for each particle's angular momentum prior to entering the magnetic field. This would imply that for classical particles the spin angular momenta, $\mathbf{S}$, are randomly oriented with all directions equally likely. For *classical particles,* it follows that the *continuous range* of values of S_z in the range $|\mathbf{S}| \leqslant S_z \leqslant |\mathbf{S}|$ are possible. Note that each particle will only arrive at one location.

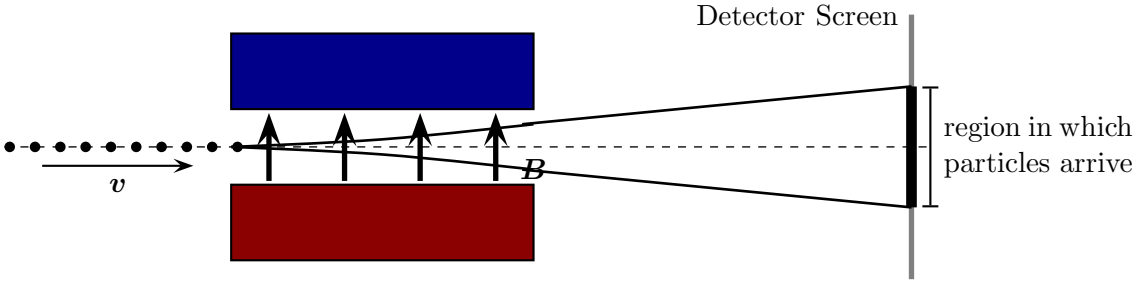


Figure II.2.4: Stern-Gerlach experimental setup with randomly prepared classical particles.

The Stern-Gerlach experiment indicates that, as illustrated in Fig. II.2.5, for spin-1/2 particles such as the electron, *only two deflections* are possible, corresponding to $S_z = +\hbar/2$ and $S_z = -\hbar/2$. For randomly prepared particles, each *only emerges at one location* and approximately half emerging in the upper beam, for which $S_z = +\hbar/2$, and half in the lower beam, for which $S_z = -\hbar/2$. Experiments indicate it is impossible to predict at which location any of these randomly prepared particles will arrive.

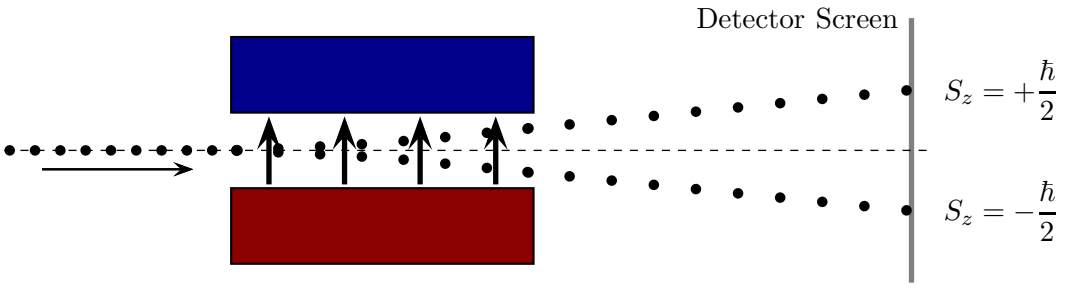


Figure II.2.5: Stern-Gerlach experiment for particles with randomly oriented magnetic dipole moments

The entire apparatus could be rotated through 90° about the x axis so that the magnetic field is now oriented along the y-axis. It follows that this would measure the y-component of the particle's spin, S_y. The possible outcomes of the experiment cannot be changed by a mere rotation of the apparatus and, for a spin-1/2 particle, must be $S_y = +\hbar/2$ or $S_y = -\hbar/2$. One can now conceive of a Stern-Gerlach experiment in which the magnetic field is oriented along an arbitrary direction, $\hat{\boldsymbol{n}}$, in which case the Stern-Gerlach device is an apparatus for measuring, S_n, the component of particle spin along $\hat{\boldsymbol{n}}$. Again the two possible outcomes are $S_n = +\hbar/2$ or $S_n = -\hbar/2$.

2.2 Schematic diagrams for the Stern-Gerlach experiment.

A schematic description of any of these Stern-Gerlach measurements requires a specification of the direction of the magnetic field, $\hat{\boldsymbol{n}}$ and the two possible locations in which the particle emerges, corresponding to $S_n = +\hbar/2$ or $S_n = -\hbar/2$. The entire arrangements of magnets can be represented by a box, labeled with $\hat{\boldsymbol{n}}$ and from which emerge two particle trajectories, each corresponding to one of the two outcomes. This is illustrated in Fig. II.2.6 and is called an SG $\hat{\boldsymbol{n}}$ measurement.

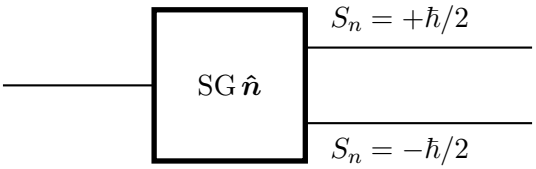


Figure II.2.6: Schematic diagram of an SG $\hat{\boldsymbol{n}}$ measurement. The horizontal line on the left indicates the trajectory of particles fired into the apparatus. Those on the right are the trajectories corresponding to each of the two outcomes.

2.3 Repeated Stern-Gerlach experiments

One can imagine a sequence of Stern-Gerlach experiments performed on the same particle. The particle is fired into one Stern-Gerlach apparatus, emerges in one of the two trajectories, and is then directed through another Stern-Gerlach apparatus. In such sequences it is possible to ask about the outcomes of later measurements given a knowledge of outcomes of earlier measurements in the sequence. For spin-1/2 particles this amounts to thought experiments; the actual experiments are either extremely difficult or have not been performed at all. However, one can show that the polarization states of individual photons can be described mathematically in a way completely analogous to the states of spin-1/2 particles; such photon states are relatively easy to subject to sequences of polarization measurements that are analogous to the sequences of Stern-Gerlach experiments to be considered.

2.3.1 Repeated measurements of the same type

Suppose that a particle is subjected two a succession of two SG $\hat{\boldsymbol{z}}$ measurements as illustrated in Fig. II.2.7. The experimental evidence indicates that any particle that emerges from the first apparatus in the upper trajectory, for which $S_z = +\hbar/2$, will emerge from the second apparatus in the upper trajectory, again giving $S_z = +\hbar/2$ with **certainty**. Similarly a particle which emerges from the first apparatus in the lower trajectory will always emerge from the second apparatus in the lower trajectory. This correlation between the two successive measurement outcomes is independent of the particle's state prior to entering the first apparatus. Note that there are states prior to the first apparatus such that the particle will never emerge in the lower (or alternatively the upper) trajectory; the correlation between the two measurement outcomes, however, is always preserved.

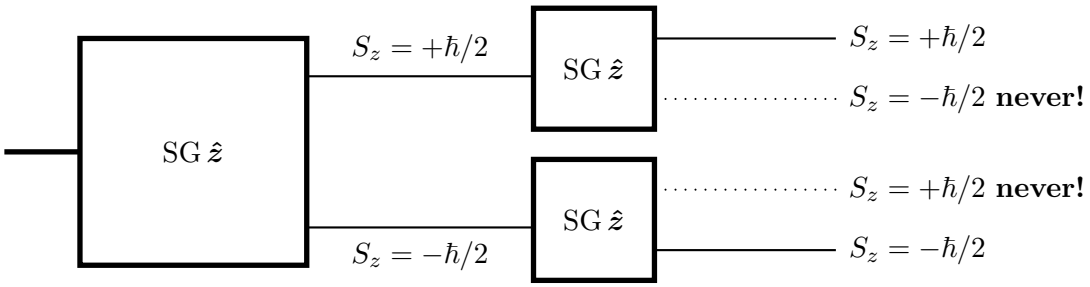


Figure II.2.7: Successive SG $\hat{\boldsymbol{z}}$ measurements.

It follows that if the first measurement yields $S_z = +\hbar/2$ then, *if the only subsequent actions on the particle are measurements of* S_z, these will always yield $S_z = +\hbar/2$. A similar statement applies to $S_z = -\hbar/2$. In this sense, one can say that after the first SG $\hat{\boldsymbol{z}}$ measurement the particle has a definite value of S_z. That is any subsequent measurement of S_z will yield just one of the values *with certainty*. In this sense, and only in this restricted sense, the particle is in the definite physical state for which $S_z = +\hbar/2$.

Definite states of physical systems such as those described above are denoted by a *"ket"*, an arrangement of brackets containing a label which indicates a certain outcome upon a particular measurement. These have the form

$$|\text{label}\rangle$$

where the contents of the ket are a label which is usually an abbreviated way of describing the particle's physical state. From the description of sequences of SG $\hat{\boldsymbol{z}}$ measurements, we use

$$\begin{aligned} |+\hat{\boldsymbol{z}}\rangle \quad &\Longleftrightarrow \quad \text{measuring } S_z \text{ yields } S_z = +\hbar/2 \text{ with certainty, and} \\ |-\hat{\boldsymbol{z}}\rangle \quad &\Longleftrightarrow \quad \text{measuring } S_z \text{ yields } S_z = -\hbar/2 \text{ with certainty.} \end{aligned} \tag{II.2.3}$$

Thus the particle which emerges from the first apparatus in the upper trajectory is in state $|+\hat{\boldsymbol{z}}\rangle$. However, neither of these states necessarily applies to particles entering the first apparatus, since these can emerge with either outcome for S_z. Note that the symbol within the brackets (here $+\hat{\boldsymbol{z}}$ or $-\hat{\boldsymbol{z}}$) is a *label* which indicates the relevant measurement and outcome.

This can be generalized to situations involving successive SG $\hat{\boldsymbol{n}}$ apparati, for some fixed direction $\hat{\boldsymbol{n}}$. It follows that the corresponding states of the spin-1/2 particle are:

$$\boxed{\begin{aligned} |+\hat{\boldsymbol{n}}\rangle \quad &\Longleftrightarrow \quad \text{measuring } S_n \text{ yields } S_n = +\hbar/2 \text{ with certainty} \\ |-\hat{\boldsymbol{n}}\rangle \quad &\Longleftrightarrow \quad \text{measuring } S_n \text{ yields } S_n = -\hbar/2 \text{ with certainty.} \end{aligned}} \tag{II.2.4}$$

Schematically the physical meaning of the general states $|+\hat{\boldsymbol{n}}\rangle$ and $|-\hat{\boldsymbol{n}}\rangle$ phrased in terms of future measurement outcomes is provided in Fig. II.2.8.

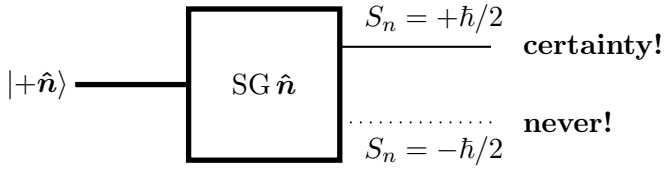


Figure II.2.8: The meaning of the state $|+\hat{\boldsymbol{n}}\rangle$ in terms of an input to SG $\hat{\boldsymbol{n}}$. The measurement outcome $S_n = +\hbar/2$ will occur with certainty.

A consequence of the relationships of Eq. (II.2.4) is that if a particle is subjected to an SG $\hat{\boldsymbol{n}}$ measurement and emerges with $S_n = +\hbar/2$ then it's state immediately after measurement is $|+\hat{\boldsymbol{n}}\rangle$. Similarly if it emerges immediately after measurement with $S_n = -\hbar/2$ then it's state immediately after measurement is $|-\hat{\boldsymbol{n}}\rangle$. An example is illustrated in Fig. II.2.9.

Example: Describe an experiment which is such that it prepares particles with certainty in the state $|+\hat{\boldsymbol{x}}\rangle$.

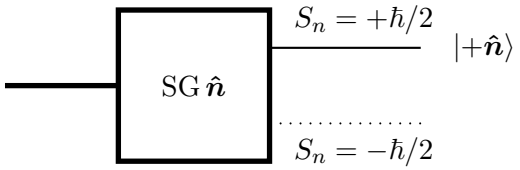


Figure II.2.9: The meaning of the state $|+\hat{\boldsymbol{n}}\rangle$ in terms of an outcome after SG $\hat{\boldsymbol{n}}$. A particle is supplied and if the measurement outcome is $S_n = +\hbar/2$ then subsequently the state is $|+\hat{\boldsymbol{n}}\rangle$.

Answer: Suppose that you had particles which are, in some sense, randomly prepared. Each is subjected to an SG $\hat{\boldsymbol{x}}$ apparatus. Those that emerge with $S_x = +\hbar/2$ are in the state $|+\hat{\boldsymbol{x}}\rangle$ and those that emerge with $S_x = -\hbar/2$ are in the state $|-\hat{\boldsymbol{x}}\rangle$. Keep all of the former and discard any of the latter. Note that this will not work if, by some chance all the particles are in the state $|-\hat{\boldsymbol{x}}\rangle$ prior to entering the SG $\hat{\boldsymbol{x}}$ apparatus. In this case more manipulations, to be explained later, are needed.

Thus there are infinitely many states available to a spin-1/2 particle corresponding to all possible directions $\hat{\boldsymbol{n}}$. In fact, it will emerge that every state available to a spin-1/2 is of the form $|+\hat{\boldsymbol{n}}\rangle$ for some unit vector $\hat{\boldsymbol{n}}$. Physically this means that for any particle there is some (possibly unknown to the experimentalist) direction $\hat{\boldsymbol{n}}$ such that if the particle is subject to an SG $\hat{\boldsymbol{n}}$ apparatus the measurement outcome will be $S_n = +\hbar/2$ *with certainty.*

2.3.2 Repeated measurements of the different types

To understand how the many possible states are related to each other consider successive Stern-Gerlach measurements of different types. For example suppose that the particle is first subjected to an SG $\hat{\boldsymbol{z}}$ apparatus and then to an SG $\hat{\boldsymbol{x}}$ apparatus as illustrated in Fig. II.2.10.

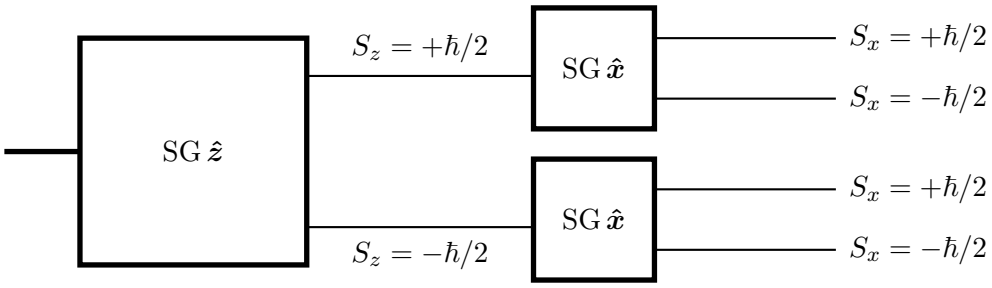


Figure II.2.10: Successive Stern-Gerlach measurements of different types.

Experiments show that if the particle emerges from the SG $\hat{\boldsymbol{z}}$ apparatus in the upper

trajectory then it can subsequently emerge in *either* (but not simultaneously both) of the trajectories after the SG $\hat{\boldsymbol{x}}$ apparatus. Unlike successive SG $\hat{\boldsymbol{z}}$ measurements no final trajectory is excluded. Furthermore the probability with which it emerges in the upper trajectory is 1/2 and the probability with which it emerges in the lower trajectory is found to be 1/2. In terms of states, one can say that

$$|+\hat{\boldsymbol{z}}\rangle \text{ into SG}\,\hat{\boldsymbol{x}} \quad \longrightarrow \quad \begin{cases} S_x = +\dfrac{\hbar}{2} & \text{with probability } \dfrac{1}{2} \\ S_x = -\dfrac{\hbar}{2} & \text{with probability } \dfrac{1}{2}. \end{cases} \tag{II.2.5}$$

In general a particle can be subjected to an SG $\hat{\boldsymbol{m}}$ apparatus followed by an SG $\hat{\boldsymbol{n}}$ apparatus where $\hat{\boldsymbol{m}}$ and $\hat{\boldsymbol{n}}$ are any two unit vectors. Suppose that the particle emerges from SG $\hat{\boldsymbol{m}}$ with $S_m = +\hbar/2$. Thus the state immediately after the first SG apparatus is $|+\hat{\boldsymbol{m}}\rangle$. The experimental evidence is that it will emerge from SG $\hat{\boldsymbol{n}}$ with $S_n = +\hbar/2$ with probability $(1 + \hat{\boldsymbol{m}} \cdot \hat{\boldsymbol{n}})/2$. Thus the general rules are:

$$\boxed{|+\hat{\boldsymbol{m}}\rangle \text{ into SG}\,\hat{\boldsymbol{n}} \quad \longrightarrow \quad \begin{cases} S_n = +\dfrac{\hbar}{2} & \text{with probability } \dfrac{1}{2}(1 + \hat{\boldsymbol{m}} \cdot \hat{\boldsymbol{n}}) \\ S_n = -\dfrac{\hbar}{2} & \text{with probability } \dfrac{1}{2}(1 - \hat{\boldsymbol{m}} \cdot \hat{\boldsymbol{n}}) \end{cases}} \tag{II.2.6}$$

and

$$\boxed{|-\hat{\boldsymbol{m}}\rangle \text{ into SG}\,\hat{\boldsymbol{n}} \quad \longrightarrow \quad \begin{cases} S_n = +\dfrac{\hbar}{2} & \text{with probability } \dfrac{1}{2}(1 - \hat{\boldsymbol{m}} \cdot \hat{\boldsymbol{n}}) \\ S_n = -\dfrac{\hbar}{2} & \text{with probability } \dfrac{1}{2}(1 + \hat{\boldsymbol{m}} \cdot \hat{\boldsymbol{n}}). \end{cases}} \tag{II.2.7}$$

Example: Suppose that a particle is subjected to an SG $\hat{\boldsymbol{x}}$ measurement and emerges with $S_x = +\hbar/2$. It is then subjected to an SG $\hat{\boldsymbol{y}}$ measurement. List the possible outcomes and the probabilities with which they occur.

Answer: If it emerges from the SG $\hat{\boldsymbol{x}}$ apparatus with $S_x = +\hbar/2$ then its state after the first measurement is $|+\hat{\boldsymbol{x}}\rangle$. Using Eq. (II.2.6) with $\hat{\boldsymbol{m}} = \hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{n}} = \hat{\boldsymbol{y}}$ one obtains outcome $S_y = +\hbar/2$ with probability 1/2 and $S_y = -\hbar/2$ with probability 1/2.

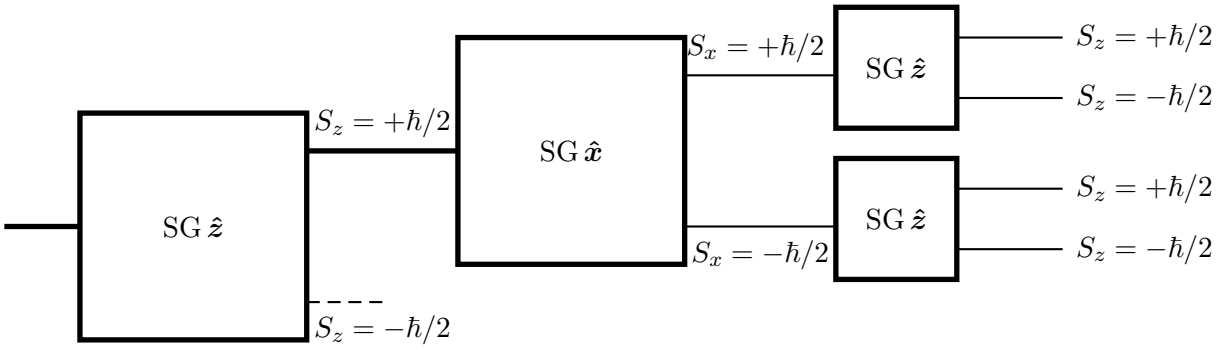


Figure II.2.11: Three successive Stern-Gerlach measurements for the circumstance in which the particle emerges with $S_z = +\hbar/2$ after the first SG $\hat{\boldsymbol{z}}$ measurement.

An unusual feature of quantum mechanics is evident when one considers two Stern-Gerlach measurements of the same type interspersed with a single Stern-Gerlach measurement along an orthogonal direction. An example is illustrated in Fig. II.2.11, in which it is assumed that the particle emerged with $S_z = +\hbar/2$ after the first SG $\hat{\boldsymbol{z}}$ measurement.

Applying the rules of (II.2.6) one can deduce that the probability with which the particle emerges with $S_z = +\hbar/2$ after either of the latter SG $\hat{\boldsymbol{z}}$ measurements is $1/2$. Similarly the probability with which the particle emerges with $S_z = -\hbar/2$ after either of the latter SG $\hat{\boldsymbol{z}}$ measurements is $1/2$.

Exercise: Consider the arrangement of Fig. II.2.11. Suppose that the particle emerges from the SG $\hat{\boldsymbol{x}}$ measurement with $S_x = +\hbar/2$. Show that the probability with which the particle emerges from the subsequent SG $\hat{\boldsymbol{z}}$ measurement with $S_z = +\hbar/2$ is $1/2$. Repeat this for a particle that emerges with $S_z = -\hbar/2$ is $1/2$. Repeat this entire analysis for the case where the particle emerges from the SG $\hat{\boldsymbol{x}}$ measurement with $S_x = -\hbar/2$.

Exercise: Repeat the argument of the previous exercise for the case where the particle emerges with $S_z = +\hbar/2$ after the first SG $\hat{\boldsymbol{z}}$ measurement.

The implication is that it makes sense to speak of a particle as having a definite value for S_z *in the context where the only subsequent operations are SG* $\hat{\boldsymbol{z}}$ *measurements.* However whenever subsequent operations include SG $\hat{\boldsymbol{n}}$ measurements where $\hat{\boldsymbol{n}}$ is distinct from $\hat{\boldsymbol{z}}$ (or $-\hat{\boldsymbol{z}}$), it is impossible to speak of a particle as having a definite value of S_z. Thus is quantum mechanics it is meaningful to speak of a particle as having a definite value for S_n only in certain very restricted contexts. One cannot universally ascribe a definite value of S_n to a particle.

2.4 Maximal description of physical states

It is possible to consider measuring more than one component of spin via successive Stern-Gerlach apparati as illustrated in Fig. II.2.12.

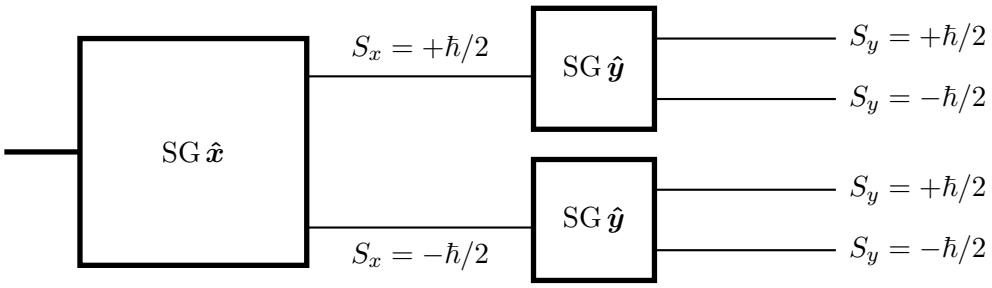


Figure II.2.12: Successive Stern-Gerlach measurements in an attempt to measure orthogonal spin components.

Any spin-1/2 particle to which these are applied will yield values for S_x and S_y. But this joint outcome can only be considered a measurement of a property of the particle if, when repeated again to a system that has undergone no extra measurements or interactions, it yields exactly the same value as it did previously.

Exercise: Consider the double SG experiment as illustrated in Fig. II.2.12 and a spin-1/2 particle that emerges from the uppermost output beam. Suppose that this is then reapplied to the same apparatus. Show that it does not emerge from the uppermost beam with certainty.

The preceding exercise indicates that two SG apparati cannot be combined to jointly measure two orthogonal components of spin in the sense that the measurement outcome is not repeatable. Thus, in the context of SG measurements it is not sensible to label the physical state of a particle via $|+\hat{\boldsymbol{m}}, +\hat{\boldsymbol{n}}\rangle$ where $\hat{\boldsymbol{m}}$ and $\hat{\boldsymbol{m}}$ are not parallel; such a state would indicate that a combined SG $\hat{\boldsymbol{m}}$ and SG $\hat{\boldsymbol{n}}$ measurement would yield $S_m = +\hbar/2$ and $S_n = +\hbar/2$ with certainty. However, a general argument similar to that of the preceding exercise indicates that this is not possible.

Thus, in terms of repeatable measurement outcomes, the only physically meaningful states are $|+\hat{\boldsymbol{n}}\rangle$ and $|-\hat{\boldsymbol{n}}\rangle$ where $\hat{\boldsymbol{n}}$ is any unit vector. The collection can be simplified by noting that, in terms of SG $\hat{\boldsymbol{n}}$ measurement outcomes, $|-\hat{\boldsymbol{z}}\rangle$ is equivalent to $|+(-\hat{\boldsymbol{n}})\rangle$.

Exercise: Show that $|+(-\hat{\boldsymbol{n}})\rangle$ yields $S_n = -\hbar/2$ with certainty. To do so define $\hat{\boldsymbol{m}} = -\hat{\boldsymbol{n}}$ and determine the probabilities with which $|+\hat{\boldsymbol{m}}\rangle$ yields $S_n = \pm\hbar/2$.

Thus the collection of all possible physically distinct states of spin-1/2 particles is $\{|+\hat{\boldsymbol{n}}\rangle\}$ where $\hat{\boldsymbol{n}}$ ranges through all unit vectors.

Mathematical Tools: Probability

Quantum mechanics provides a means for determining the probabilities with which the various outcomes of a measurement can occur. The intuitive motion of probability is established by considering a large number of identical measurements on copies of identically prepared systems. Suppose that N spin-1/2 particles are all prepared in the same state and subjected to the same SG $\hat{\boldsymbol{n}}$ measurement. If the the probability of attaining $S_n = +\hbar/2$ is p then in a typical run of N such measurements the outcome $S_n = +\hbar/2$ will occur approximately pN times, although this is not guaranteed. Deep theorems in probability such as the central limit theorem quantify such statements, and these become more precise as $N \to \infty$.

The general tools of probability theory can be illustrated by considering a generic measurement with n distinct outcomes, labeled $m_1, m_2, \ldots, m_n$. Denote the probability with which the outcome m_i occurs by p_i. Following the basic notion of a probability representing fractional occurrences of a given outcome after many measurements, it is required that the probabilities satisfy:

$$0 \leqslant p_i \leqslant 1 \tag{II.2.8}$$

and that

$$\sum_{i=1}^{n} p_i = 1. \tag{II.2.9}$$

The collection of all p_i is called a *probability distribution.*

Example: An unbiased die can give one of the following outcomes after being rolled: $1, 2, 3, 4, 5, 6$. Regarding a roll of the die as a measurement, the outcomes are $m_1 = 1, m_2 = 2, \ldots, m_6 = 6$. For an unbiased die the probabilities with which each outcome occurs are identical; thus $p_1 = 1/6, p_2 = 1/6$. It is easy to see that these satisfy Eqs. (II.2.9) and (II.2.9).

The mean of a distribution gives a notion of the average numerical outcome and is defined via:

$$\langle m \rangle := \sum_{i=1}^{n} m_i p_i. \tag{II.2.10}$$

In physics and particularly in quantum mechanics, this is known as the expectation value of the quantity which is being measured.

Example: For an unbiased die, the mean is:

$$\begin{aligned} \langle m \rangle &= \sum_{i=1}^{6} m_i p_i = \sum_{i=1}^{6} m_i \frac{1}{6} \\ &= \frac{1}{6} \sum_{i=1}^{6} m_i = \frac{1}{6} \left(1 + 2 + 3 + 4 + 5 + 6\right) \\ &= \frac{7}{2}. \end{aligned}$$

It is important to note that the mean is an idealized quantity and that a given run of measurements will not necessarily yield the mean when averaged. Suppose that the outcomes of an experiment involving N measurements are $s_1, s_2, s_3, \ldots s_N$. This constitutes a sample of the probability distribution and the *sample average* is defined as:

$$\overline{m} := \frac{1}{N} \sum_{i=1}^{N} s_i. \tag{II.2.11}$$

Typically for large N one expects that $\overline{m} \approx \langle m \rangle$.

Example: For an unbiased die, in one trial consisting of ten rolls, the following outcomes occurred: $2, 4, 4, 3, 3, 2, 1, 4, 4, 6$. The sample average of these is 3.3 which differs from the mean, 3.5.

Various theorems in probability theory quantify the extent to which a sample average approximates the mean. Generally as N increases the probability with which the sample average will be within a given range of the mean increases; the probability with which it is beyond the a certain range of the mean diminishes as $1/N$.

An important tool in quantifying such discrepancies and fluctuations away from the mean is the *standard deviation* or variance of a probability distribution. This is defined as:

$$\Delta m := \sqrt{\sum_{i=1}^{n} p_i \left(m_i - \langle m \rangle\right)^2} \tag{II.2.12}$$

and this quantifies the extent to which measurement outcomes typically deviate from the mean. A general result that simplifies this calculation is:

$$\Delta m := \sqrt{\langle m^2 \rangle - \langle m \rangle^2} \tag{II.2.13}$$

where

$$\langle m^2 \rangle := \sum_{i=1}^{n} m_i^2 p_i. \qquad \blacksquare$$

These notions can be applied to physical systems subject to the laws of quantum mechanics.

Example: An ensemble of spin-1/2 particles is subjected to an SG $\hat{\boldsymbol{m}}$ measurement where $\hat{\boldsymbol{m}} = \frac{1}{\sqrt{2}}\,\hat{\boldsymbol{x}} + \frac{1}{\sqrt{2}}\,\hat{\boldsymbol{y}}$. Only the particles that emerge from this measurement having given the outcome $S_m = +\hbar/2$ are retained. Each of these is subjected to an SG $\hat{\boldsymbol{x}}$ measurement. Determine $\langle S_x \rangle$ and ΔS_x.

Answer: The particles that emerge from the initial SG $\hat{\boldsymbol{m}}$ measurement with $S_m = +\hbar/2$ are in the state $|+\hat{\boldsymbol{m}}\rangle$. We can list the measurement outcomes and probabilities for the SG $\hat{\boldsymbol{x}}$ measurement performed on particles in this state:

Outcome (S_z)	Probability
$m_1 = +\dfrac{\hbar}{2}$	$p_1 = \dfrac{\sqrt{2}+1}{2\sqrt{2}}$
$m_2 = -\dfrac{\hbar}{2}$	$p_2 = \dfrac{\sqrt{2}-1}{2\sqrt{2}}$

where we have used

$$\Pr\left(S_x = +\hbar/2\right) = \frac{1}{2}\left(1 + \hat{\boldsymbol{m}}\cdot\hat{\boldsymbol{n}}\right)$$
$$\Pr\left(S_x = -\hbar/2\right) = \frac{1}{2}\left(1 - \hat{\boldsymbol{m}}\cdot\hat{\boldsymbol{n}}\right)$$

etc,.... Thus

$$\begin{aligned}\langle S_z\rangle &= \sum_{i=1}^{n} m_i p_i \\ &= +\frac{\hbar}{2}\,\frac{\sqrt{2}+1}{2\sqrt{2}} + \left(-\frac{\hbar}{2}\right)\,\frac{\sqrt{2}-1}{2\sqrt{2}} \\ &= \frac{\hbar}{2\sqrt{2}}.\end{aligned}$$

Then

$$\Delta S_z = \sqrt{\langle S_z^2\rangle - \langle S_z\rangle^2}.$$

Here

$$\begin{aligned}\langle S_z^2\rangle &= \sum_{i=1}^{n} m_i^2 p_i \\ &= \left(+\frac{\hbar}{2}\right)^2\,\frac{\sqrt{2}+1}{2\sqrt{2}} + \left(-\frac{\hbar}{2}\right)^2\,\frac{\sqrt{2}-1}{2\sqrt{2}} \\ &= \frac{\hbar^2}{4}.\end{aligned}$$

Thus

$$\Delta S_z = \sqrt{\frac{\hbar^2}{4} - \frac{\hbar^2}{8}} = \frac{\hbar}{2\sqrt{2}}.$$

3 Mathematical Description of Spin-1/2 particles

The physical states of spin-1/2 particles are represented in terms of kets which are interpreted in terms of the outcomes of spin measurements. Thus the collection of possible states for a spin-1/2 particle is

$$\{|+\hat{\boldsymbol{n}}\rangle \mid \quad \hat{\boldsymbol{n}} \text{ unit vectors } \} \tag{II.3.1}$$

where the set is constructed from all unit vectors, $\hat{\boldsymbol{n}}$. We shall now supply a mathematical structure to this set in such a way that the permissible mathematical manipulations result in predictions consistent with the experiments that were described earlier.

Taking a cue from the phenomena and tools of wave mechanics, we would like to be able to form superpositions of kets such as

$$\frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle \quad \text{or} \quad \frac{3}{5}\,|+\hat{\boldsymbol{x}}\rangle + \frac{4i}{5}\,|-\hat{\boldsymbol{x}}\rangle \quad \text{or} \quad \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{x}}\rangle + \frac{1}{2}\,|-\hat{\boldsymbol{z}}\rangle \quad \text{etc,...}$$

In general we hope to give mathematical and physical meanings to arbitrary superpositions such as

$$c_+\,|+\hat{\boldsymbol{z}}\rangle + c_-\,|-\hat{\boldsymbol{z}}\rangle$$

where c_+ and c_- are any complex constants. Mathematically this requires that, at least, the kets must be elements of a **complex vector space.**

Mathematical Tools: Vector spaces

An example of a vector space is the set of all directed arrows in the two dimensional plane. Two elements of this set are illustrated in Fig. II.3.1. The entire vector space consists of all such directed arrows whose tails are located at the origin.

An alternative and useful representation of any vector is in terms of the Cartesian coordinates of the tip of the corresponding arrow. These are the conventional components of the vectors and this representation is denoted

$$\mathbf{u} \leftrightarrow \begin{pmatrix} u_x \\ u_y \end{pmatrix}$$

Example: The components of the vector $\mathbf{u}$ in Fig. II.3.1 are $u_x = 2$ and $u_y = 3$.

Thus the set of vectors in the plane is equivalent to:

$$\left\{ \begin{pmatrix} u_x \\ u_y \end{pmatrix} \mid u_x \in \mathbb{R}, u_y \in \mathbb{R} \right\}$$

and is denoted $\mathbb{R}^2$.

This set of vectors is more than a mere collection; it is possible to add and subtract vectors and multiply these by real numbers. These operations allow various vectors to be related to each other; in this way, the set has a greater structure than the collection of elements from which it is composed. Addition and subtraction of vectors can be envisaged

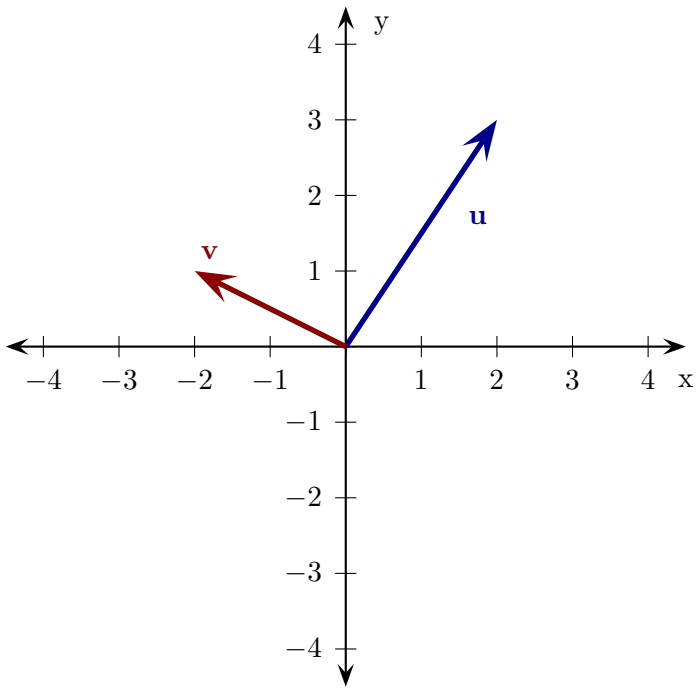


Figure II.3.1: Vectors in $\mathbb{R}^2$.

graphically using the elementary notions of parallel transport and the resulting "head-to-tail" construction. However, such operations are best realized mathematically by using the components of the vectors. The results are probably familiar. If

$$\mathbf{u} \leftrightarrow \begin{pmatrix} u_x \\ u_y \end{pmatrix} \quad \text{and} \quad \mathbf{v} \leftrightarrow \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

then

$$\mathbf{u} + \mathbf{v} \leftrightarrow \begin{pmatrix} u_x + v_x \\ u_y + v_y \end{pmatrix}$$
$$\mathbf{u} - \mathbf{v} \leftrightarrow \begin{pmatrix} u_x - v_x \\ u_y - v_y \end{pmatrix}$$
$$\alpha\mathbf{u} \leftrightarrow \begin{pmatrix} \alpha u_x \\ \alpha u_y \end{pmatrix}$$

where α is any real number. Note that among the vectors there is a special element (the zero vector)

$$\mathbf{0} \leftrightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which satisfies $\mathbf{0} + \mathbf{u} = \mathbf{u}$ for any vector $\mathbf{u}$. Also, for every vector $\mathbf{u}$, there is a vector $-\mathbf{u}$ such that $\mathbf{0} + (-\mathbf{u}) = \mathbf{0}$.

Vectors in three dimensional space obey the same rules; the only difference is that they require three components to mimic the correspondence used above; in fact, the operations on vectors in the plane generalize to many other sets of mathematical objects. The next step is to abstract these essential, common, general properties from the examples above. This leads to the definition of a *vector space* as a set, denoted V, which is equipped with addition and multiplication by elements of a set of scalars, F, which are frequently either the real numbers or the complex numbers. Addition gives

$$\boxed{\mathbf{u}, \mathbf{v} \in V \Rightarrow \mathbf{u} + \mathbf{v} \in V} \tag{II.3.2}$$

while multiplication gives

$$\boxed{\mathbf{u} \in V,\ \alpha \in \Rightarrow \alpha\mathbf{u} \in V} \tag{II.3.3}$$

for any $\mathbf{u}, \mathbf{v} \in V$ and any $\alpha \in F$. A number of technical requirements, which will be obvious for vectors in the plane, are essential for manipulating vectors and proving key results. There must be a zero vector $\mathbf{0}$. Addition must satisfy

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \tag{II.3.4a}$$
$$(\mathbf{u} + \mathbf{v}) + \boldsymbol{z} = \mathbf{u} + (\mathbf{v} + \boldsymbol{z}) \tag{II.3.4b}$$
$$\mathbf{u} + \mathbf{0} = \mathbf{u} \tag{II.3.4c}$$

for every $\mathbf{u}, \mathbf{v} \in V$. Additionally for every $\mathbf{u}$ there must be a negative vector $-\mathbf{u}$ such that

$$\mathbf{u} + (-\mathbf{u}) = \mathbf{0}. \tag{II.3.5}$$

Scalar multiplication is required to satisfy:

$$1\mathbf{u} = \mathbf{u} \tag{II.3.6a}$$
$$\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u} \tag{II.3.6b}$$
$$(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u} \tag{II.3.6c}$$
$$\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v} \tag{II.3.6d}$$

for any $\alpha, \beta \in F$ and $\mathbf{u}, \mathbf{v} \in V$.

These requirements should appear self-evident in the context of vectors in the plane. Here, once addition and multiplication have been defined via manipulations of components, it is straightforward to verify that the requirements for these operations are satisfied. However, note that they do not necessarily follow from the basic definitions of addition and multiplication.

The definitions and requirements of Eqs. (II.3.2) to (II.3.6) establish the mathematical structure of a vector space. In addition to two and three dimensional real space, examples are: continuous functions on an interval, matrices with a fixed number of rows and columns, etc, ... In all cases, the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$ would be candidates for F (these are not the only possibilities). Choosing one versus the other will give distinct vector spaces whose elements cannot be compared or combined in general.

Although there is a broad range of possibilities for vector spaces, most of the key features can be envisioned in terms of vectors in the real plane, i.e. in terms of the vector space $\mathbb{R}^2$. ■

The first requirement for describing a spin-1/2 particle mathematically is that the set of all kets be a vector space for which the scalars are taken from the complex numbers (i.e. F is $\mathbb{C}$). This means that kets such as $|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle, |+\hat{\boldsymbol{x}}\rangle, |-\hat{\boldsymbol{x}}\rangle, \ldots$ will all be regarded as vectors. However, any combination of vectors is again a vector and thus

$$\frac{1}{\sqrt{2}}|+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}}|-\hat{\boldsymbol{z}}\rangle \quad \text{or} \quad \frac{3}{5}|+\hat{\boldsymbol{z}}\rangle + \frac{4i}{5}|-\hat{\boldsymbol{z}}\rangle$$

etc, ... are also all vectors. Combinations such as those above are called *linear superpositions.* Every ket, including all linear superpositions, will correspond to a physical state of a spin-1/2 particle (we will describe how to ascribe meanings later). It will emerge that the *most general state of a spin-*1/2 *particle* can be expressed as

$$|\Psi\rangle = c_+|+\hat{\boldsymbol{z}}\rangle + c_-|-\hat{\boldsymbol{z}}\rangle \tag{II.3.7}$$

where c_+, c_- are complex numbers and Ψ serves as a label for a generic ket (much like $\mathbf{v}$ denotes a generic vector in the two dimensional plane). The rules for adding and multiplying such combinations are taken directly from the rules of vector addition. In particular,

$$\big(a_+|+\hat{\boldsymbol{z}}\rangle + a_-|-\hat{\boldsymbol{z}}\rangle\big) + \big(b_+|+\hat{\boldsymbol{z}}\rangle + b_-|-\hat{\boldsymbol{z}}\rangle\big) = (a_+ + b_+)|+\hat{\boldsymbol{z}}\rangle + (a_- + b_-)|-\hat{\boldsymbol{z}}\rangle \tag{II.3.8}$$

$$\alpha\big(a_+|+\hat{\boldsymbol{z}}\rangle + a_-|-\hat{\boldsymbol{z}}\rangle\big) = \alpha a_+|+\hat{\boldsymbol{z}}\rangle + \alpha a_-|-\hat{\boldsymbol{z}}\rangle\,. \tag{II.3.9}$$

Note that, for spin-1/2 quantum systems, the symbol inside the ket brackets (e.g. $+\hat{\boldsymbol{z}}$ in $|+\hat{\boldsymbol{z}}\rangle$) is *only a label.* Here the labels happen to be vectors from the real three dimensional space, which is entirely different from the vector space which the kets inhabit. One may be tempted to replace linear superpositions of kets by a single ket with the similar superposition of labels but this will be inconsistent with measurement outcomes and hence incorrect. In general

$$\alpha|+\hat{\boldsymbol{x}}\rangle + \beta|+\hat{\boldsymbol{y}}\rangle \neq |+\alpha\boldsymbol{x} + \beta\boldsymbol{y}\rangle\,.$$

At this stage we have only ascribed physical meaning, in terms of outcomes of measurements, to states of the form $|+\hat{\boldsymbol{n}}\rangle, |-\hat{\boldsymbol{n}}\rangle$ and it is unclear what superpositions such as that of Eq. (II.3.7) imply. For example, one could subject a spin-1/2 particle known to be in the state $|\Psi\rangle = 3/5\,|+\hat{\boldsymbol{z}}\rangle + 4i/5\,|-\hat{\boldsymbol{z}}\rangle$ to an SG $\hat{\boldsymbol{z}}$ measuring apparatus and ask what the probabilities of attaining $S_z = +\hbar/2$ or $S_z = -\hbar/2$ are. At the moment the answer should be unclear. Furthermore it should not be apparent that the state in Eq. (II.3.7) is the most general possibility for a state of a spin-1/2 particle.

Additional rules are necessary to resolve these issues and completely connect the mathematical structure to the physical picture. It will emerge that these are best answered in terms of the linear algebraic concepts of basis states and *inner product,* a generalization of the ordinary dot product.

Mathematical Tools: Bases, inner products and orthonormality

Every vector in $\mathbb{R}^2$ can be expressed uniquely as a linear superposition of the the two vectors $\{\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2\}$ illustrated in Fig. II.3.2. Thus

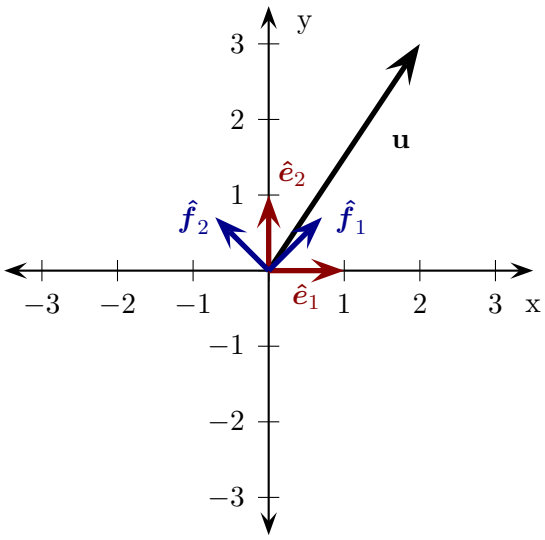


Figure II.3.2: Bases for vectors in $\mathbb{R}^2$.

$$\mathbf{u} = u_x\hat{\boldsymbol{e}}_1 + u_y\hat{\boldsymbol{e}}_2$$

where, for any given vector $\mathbf{u}$, u_x and u_y are real numbers, called the *components of* $\mathbf{u}$ *in the basis* $\{\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2\}$. Arranging the components as a column vector

$$\begin{pmatrix} u_x \\ u_y \end{pmatrix}$$

is called the *representation of* $\mathbf{u}$ *in the basis* $\{\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2\}$.

Example: For $\mathbf{u}$ in Fig. II.3.2, $\mathbf{u} = 2\hat{\boldsymbol{e}}_1 + 3\hat{\boldsymbol{e}}_2$.

Given a vector space, V, any set of vectors $\{\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \ldots, \hat{\boldsymbol{e}}_n\}$ which has the property that any $\mathbf{v} \in V$ can be expressed as $\mathbf{v} = \sum_{i=1}^{n} v^i\hat{\boldsymbol{e}}_i$, for a unique set of numbers v^i is called a *basis* for the vector space. The numbers $v^1, \ldots, v^n$ are called *components of* $\mathbf{u}$ *with respect to the basis* $\{\hat{\boldsymbol{e}}_1, \ldots, \hat{\boldsymbol{e}}_n\}$. Provided that one finite set of basis vectors can be found, it is readily shown that all other possible basis sets have the same number of vectors; this is called the *dimension of* V. The fact that there may be more than one possible basis set is evident from the example of Fig. II.3.2, in which

$$\hat{\boldsymbol{f}}_1 = \frac{1}{\sqrt{2}}(\hat{\boldsymbol{e}}_1 + \hat{\boldsymbol{e}}_2)$$
$$\hat{\boldsymbol{f}}_2 = \frac{1}{\sqrt{2}}(-\hat{\boldsymbol{e}}_1 + \hat{\boldsymbol{e}}_2)$$

can easily be seen to form a basis for the set of vectors in the plane.

An additional useful concept associated with vectors in the two dimensional plane is the dot product, which is a map from pairs of vectors onto the real numbers:

$$\mathbf{u}, \mathbf{v} \mapsto \mathbf{u} \cdot \mathbf{v} \in \mathbb{R}.$$

The important properties of the dot product are

$$\mathbf{u} \cdot (\alpha \mathbf{v} + \beta \boldsymbol{z}) = \alpha \mathbf{u} \cdot \mathbf{v} + \beta \mathbf{u} \cdot \boldsymbol{z}$$
$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$$

for any $\mathbf{u}, \mathbf{v}, \boldsymbol{z}$ and numbers $\alpha, \beta \in \mathbb{R}$. The dot product is useful in calculations and leads to the notion of orthogonality. Two vectors in the plane $\mathbf{u}, \mathbf{v}$ are *orthogonal* if $\mathbf{u} \cdot \mathbf{v} = 0$. Additionally, $\mathbf{u}$ is *normalized* if $\mathbf{u} \cdot \mathbf{u} = 1$. A collection of normalized basis vectors which are mutually orthogonal is called an *orthonormal basis.*

The inner product is a generalization of this to vector spaces for which the scalars are complex and is signified by a pair of brackets rather than the "dot." This an inner product is a map

$$\mathbf{u}, \mathbf{v} \in V \mapsto (\mathbf{u}, \mathbf{v}) \in \mathbb{C} \tag{II.3.10}$$

which satisfies

$$(\mathbf{u}, \alpha \mathbf{v} + \beta \boldsymbol{z}) = \alpha (\mathbf{u}, \mathbf{v}) + \beta (\mathbf{u}, \boldsymbol{z}) \tag{II.3.11a}$$
$$(\mathbf{v}, \mathbf{u}) = (\mathbf{u}, \mathbf{v})^* \tag{II.3.11b}$$
$$(\mathbf{u}, \mathbf{u}) \geqslant 0 \quad \text{with equality} \Leftrightarrow u = 0. \tag{II.3.11c}$$

Exercise: Let V be the set of all two-dimensional complex column vectors

$$V := \left\{ \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \mid u^1, u^2 \in \mathbb{C} \right\}.$$

For

$$\mathbf{u} = \begin{pmatrix} u^1 \\ u^2 \end{pmatrix} \quad \text{and} \quad \mathbf{v} = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix},$$

define the following candidate for an inner product

$$(\mathbf{u}, \mathbf{v}) := (u^1)^* v^1 + (u^2)^* v^2.$$

Show that this satisfies the requirements of Eqs. (II.3.11).

Example: Using the definition in the previous example, compute the inner product $(\mathbf{u}, \mathbf{v})$ where

$$\mathbf{u} = \begin{pmatrix} 1 \\ e^{i\pi/4} \end{pmatrix} \quad \text{and} \quad \mathbf{v} = \begin{pmatrix} 3i \\ 5 \end{pmatrix}.$$

Compute $(\mathbf{v}, \mathbf{u})$ and show that Eq. (II.3.11b) is satisfied.

Answer: We can use the definition of the inner product in the previous exercise to give

$$(\mathbf{u},\mathbf{v}) = (1)^* 3i + \left(e^{i\pi/4}\right)^* 5 = 3i + 5e^{-i\pi/4}.$$

Then using

$$e^{i\theta} = \cos\theta + i\sin\theta$$

we get

$$(\mathbf{v},\mathbf{u}) = 3i + 5\left(\frac{1}{\sqrt{2}} - i\frac{1}{\sqrt{2}}\right) = \frac{5}{\sqrt{2}} + \left(3 - \frac{5}{\sqrt{2}}\right) i.$$

Similarly

$$(\mathbf{v},\mathbf{u}) = (3i)^* 1 + 5^* e^{i\pi/4} = -3i + 5e^{i\pi/4} = \frac{5}{\sqrt{2}} - \left(3 - \frac{5}{\sqrt{2}}\right) i.$$

It is clear that $(\mathbf{v},\mathbf{u}) = (\mathbf{u},\mathbf{v})^*$.

Given that V has an inner product, two vectors $\mathbf{u},\mathbf{v}$ are called *orthogonal* if and only if

$$(\mathbf{u},\mathbf{v}) = 0 \tag{II.3.12}$$

and a vector is *normalized* if and only if

$$(\mathbf{u},\mathbf{u}) = 1. \tag{II.3.13}$$

A basis consisting of mutually orthogonal, normalized vectors is called an *orthonormal basis* and it can be shown that, for a finite dimensional vector space with an inner product, orthonormal bases always exist.

Exercise: Let V be the set of all two-dimensional complex column vectors. Consider the following basis vectors

$$\hat{\boldsymbol{f}}_1 := \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix} \quad \text{and} \quad \hat{\boldsymbol{f}}_2 := \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix}.$$

Show that this is an orthonormal basis. ■

The *inner product of the ket* $|\Phi\rangle$ *with the ket* $|\Psi\rangle$ will be denoted

$$\boxed{\langle\Phi|\Psi\rangle}$$

and this replaces the standard mathematical notation $(|\Phi\rangle, |\Psi\rangle)$. Note that this satisfies:

$$\boxed{\text{If } |\Psi\rangle = \alpha\,|\Psi_1\rangle + \beta\,|\Psi_2\rangle \text{ then } \langle\Phi|\Psi\rangle = \alpha\,\langle\Phi|\Psi_1\rangle + \beta\,\langle\Phi|\Psi_2\rangle} \tag{II.3.14}$$

where $|\Psi_1\rangle, |\Psi_2\rangle$ are kets and α,β are complex numbers. Additionally,

$$\boxed{\langle\Phi|\Psi\rangle = \big(\langle\Psi|\Phi\rangle\big)^*} \tag{II.3.15}$$

and the right hand side is usually written $\langle\Psi|\Phi\rangle^*$. Eqs. (II.3.14) and (II.3.15) imply

$$\boxed{\text{If } |\Phi\rangle = \alpha\,|\Phi_1\rangle + \beta\,|\Phi_2\rangle \text{ then } \langle\Phi|\Psi\rangle = \alpha^*\,\langle\Phi_1|\Psi\rangle + \beta^*\,\langle\Phi_2|\Psi\rangle} \tag{II.3.16}$$

where $|\Phi_1\rangle, |\Phi_2\rangle$ are kets.

Example: We will shortly see that the two kets $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ are orthonormal. Then consider

$$|\Phi\rangle = \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle + \frac{i}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle \quad \text{and}$$
$$|\Psi\rangle = \frac{3i}{5}\,|+\hat{\boldsymbol{z}}\rangle + \frac{4}{5}\,|-\hat{\boldsymbol{z}}\rangle\,.$$

Determine $\langle\Phi|\Psi\rangle$.

Answer: Using Eq. (II.3.16),

$$\langle\Phi|\Psi\rangle = \left(\frac{1}{\sqrt{2}}\right)^* \langle+\hat{\boldsymbol{z}}|\Psi\rangle + \left(\frac{i}{\sqrt{2}}\right)^* \langle-\hat{\boldsymbol{z}}|\Psi\rangle = \frac{1}{\sqrt{2}}\,\langle+\hat{\boldsymbol{z}}|\Psi\rangle - \frac{i}{\sqrt{2}}\,\langle-\hat{\boldsymbol{z}}|\Psi\rangle\,.$$

Now Eq. (II.3.14) gives

$$\langle+\hat{\boldsymbol{z}}|\Psi\rangle = \frac{3i}{5}\,\langle+\hat{\boldsymbol{z}}|+\hat{\boldsymbol{z}}\rangle + \frac{4}{5}\,\langle+\hat{\boldsymbol{z}}|-\hat{\boldsymbol{z}}\rangle \quad \text{and}$$
$$\langle-\hat{\boldsymbol{z}}|\Psi\rangle = \frac{3i}{5}\,\langle-\hat{\boldsymbol{z}}|+\hat{\boldsymbol{z}}\rangle + \frac{4}{5}\,\langle-\hat{\boldsymbol{z}}|-\hat{\boldsymbol{z}}\rangle\,.$$

But if the kets of $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ are orthonormal then

$$\langle+\hat{\boldsymbol{z}}|+\hat{\boldsymbol{z}}\rangle = 1$$
$$\langle+\hat{\boldsymbol{z}}|-\hat{\boldsymbol{z}}\rangle = 0$$
$$\langle-\hat{\boldsymbol{z}}|+\hat{\boldsymbol{z}}\rangle = 0$$
$$\langle-\hat{\boldsymbol{z}}|-\hat{\boldsymbol{z}}\rangle = 1.$$

Thus

$$\langle+\hat{\boldsymbol{z}}|\Psi\rangle = \frac{3i}{5} \quad \text{and}$$
$$\langle-\hat{\boldsymbol{z}}|\Psi\rangle = \frac{4}{5}.$$

This gives

$$\langle\Phi|\Psi\rangle = \frac{1}{\sqrt{2}}\,\frac{3i}{5} - \frac{i}{\sqrt{2}}\,\frac{4}{5} = -\frac{i}{5\sqrt{2}}.$$

The first step in relating the vector space structure of the set of kets to the physical states of the spin-1/2 particle connects possible measurement outcomes to orthonormal bases for the vector space. The fundamental requirement is based on considerations of measurements. The first part concerns any conceivable measurement.

Kets are associated with mutually incompatible outcomes of any *one* measurement if and only if they are orthogonal.

Here the ket associated with a particular measurement outcome is that which gives the outcome with certainty. Mutually incompatible outcomes means that one will never confuse one measurement outcome with another.

Example: For a spin-1/2 particle, consider measuring S_z via an ideal SG $\hat{\boldsymbol{z}}$ apparatus. The ket $|+\hat{\boldsymbol{z}}\rangle$ gives $S_z = +\hbar/2$ with certainty and $|-\hat{\boldsymbol{z}}\rangle$ gives $S_z = -\hbar/2$ with certainty. These outcomes are mutually incompatible since the trajectories in which the particle could emerge from the apparatus will not be confused. Thus $|+\hat{\boldsymbol{z}}\rangle$ and $|-\hat{\boldsymbol{z}}\rangle$ are orthogonal.

Exercise: Argue that for a spin-1/2 particle, the kets $|+\hat{\boldsymbol{n}}\rangle$ and $|-\hat{\boldsymbol{n}}\rangle$ where $\hat{\boldsymbol{n}}$ is any unit vector are orthogonal.

Exercise: For a spin-1/2 particle, consider the kets $|+\hat{\boldsymbol{x}}\rangle$ and $|+\hat{\boldsymbol{z}}\rangle$. Are these orthogonal? To answer the question, you will need to find a direction $\hat{\boldsymbol{n}}$ corresponding to measuring S_n using an SG $\hat{\boldsymbol{n}}$ apparatus. One of the states must be associated with the outcomes $S_n = +\hbar/2$ and the other with $S_n = -\hbar/2$ if they are to be orthogonal. Can you find such an $\hat{\boldsymbol{n}}$?

The second step in relating the vector space structure of the set of kets to the physical states of the spin-1/2 particle is somewhat subtle and deals with the wide variety of possible measurements that one can do. For example, a spin-1/2 particle can be subject to an ideal SG $\hat{\boldsymbol{n}}$ apparatus for any direction $\hat{\boldsymbol{n}}$ which then records the value of S_n. However, one can imagine that the apparatus is crudely constructed so that the trajectories that, in the ideal case, corresponded to $S_n = +\hbar/2$ and $S_n = -\hbar/2$ now overlap. For such a crude apparatus there is just one measurement outcome. Although this would be pointless considering crude measurements like this for a spin-1/2 particle, there are other quantum mechanical systems for which such measurements are worth contemplating. In fact, measuring the energy of the electron in a hydrogen atom is like this, since there are many distinct angular momentum values corresponding to one energy state. At present we shall consider only measurements which are "best," in the sense that they discriminate as finely as possible between possible states of the quantum mechanical system; these are called *maximal*. For a spin-1/2 particle, measuring any single component of spin, S_n, for any direction $\hat{\boldsymbol{n}}$ is maximal. Then for a *maximal measurement:*

The kets associated with mutually incompatible outcomes of any *one maximal* measurement form an orthonormal basis. Conversely the kets that make any orthonormal basis are associated with the mutually incompatible measurement outcomes of one maximal measurement.

Example: For a spin-1/2 particle, consider measuring S_z via an ideal SG $\hat{\boldsymbol{z}}$ apparatus. The ket $|+\hat{\boldsymbol{z}}\rangle$ gives $S_z = +\hbar/2$ with certainty and $|-\hat{\boldsymbol{z}}\rangle$ gives $S_z = -\hbar/2$ with certainty. Thus the kets $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ form an orthonormal basis.

It follows that the dimension of the vector space of kets for a single spin-1/2 particle is two. The orthogonality and normalization requirements imply that

$$\langle +\hat{\boldsymbol{z}}| + \hat{\boldsymbol{z}}\rangle = 1 \tag{II.3.17a}$$

$$\langle +\hat{\boldsymbol{z}}| - \hat{\boldsymbol{z}}\rangle = \langle -\hat{\boldsymbol{z}}| + \hat{\boldsymbol{z}}\rangle = 0 \tag{II.3.17b}$$

$$\langle -\hat{\boldsymbol{z}}| - \hat{\boldsymbol{z}}\rangle = 1. \tag{II.3.17c}$$

Clearly any ket can be expressed as the linear superposition

$$|\Psi\rangle = c_+ |+\hat{\boldsymbol{z}}\rangle + c_- |-\hat{\boldsymbol{z}}\rangle$$

where c_+ and c_- are complex numbers. The representation of this state in this basis establishes a correspondence between kets and column vectors with two complex numbers:

$$\boxed{|\Psi\rangle = c_+ |+\hat{\boldsymbol{z}}\rangle + c_- |-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} c_+ \\ c_- \end{pmatrix}.} \tag{II.3.18}$$

In particular,

$$\boxed{|+\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \, |-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.} \tag{II.3.19}$$

Exercise: For a spin-1/2 particle, argue that the kets $\{|+\hat{\boldsymbol{n}}\rangle, |-\hat{\boldsymbol{n}}\rangle\}$ form an orthonormal basis. Show that $\langle +\hat{\boldsymbol{n}}| + \hat{\boldsymbol{n}}\rangle = \langle -\hat{\boldsymbol{n}}| - \hat{\boldsymbol{n}}\rangle = 1$ and that $\langle +\hat{\boldsymbol{n}}| - \hat{\boldsymbol{n}}\rangle = 0$.

Exercise: For a spin-1 particle the maximally discriminating measurement is that of a single component of the particle's spin. For example, experiments show that a spin-1 particle subjected to an SG $\hat{\boldsymbol{z}}$ apparatus gives one of the following possible outcomes: $S_z = +\hbar$, or $S_z = 0$ or $S_z = -\hbar$. What must the dimension of the vector space of kets be for a spin-1 particle?

It is frequently convenient to work with the representation of a ket in a particular basis rather than using more abstract notions such as those of Eq. (II.3.14). For example, calculating inner products is often done in terms of components in the basis $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$.

Theorem: If $|\Psi\rangle = a_+ |+\hat{\boldsymbol{z}}\rangle + a_- |-\hat{\boldsymbol{z}}\rangle$ and $|\Phi\rangle = b_+ |+\hat{\boldsymbol{z}}\rangle + b_- |-\hat{\boldsymbol{z}}\rangle$ then

$$\boxed{\langle\Phi|\Psi\rangle = b_+^* a_+ + b_-^* a_-.} \tag{II.3.20}$$

Proof: First Eq. (II.3.14) gives

$$\langle\Phi|\Psi\rangle = a_+ \langle\Phi| + \hat{\boldsymbol{z}}\rangle + a_- \langle\Phi| - \hat{\boldsymbol{z}}\rangle .$$

Then Eqs. (II.3.16) and (II.3.17) imply

$$\langle\Phi| + \hat{\boldsymbol{z}}\rangle = b_+^* \langle +\hat{\boldsymbol{z}}| + \hat{\boldsymbol{z}}\rangle + b_-^* \langle -\hat{\boldsymbol{z}}| + \hat{\boldsymbol{z}}\rangle = b_+^*$$
$$\langle\Phi| - \hat{\boldsymbol{z}}\rangle = b_+^* \langle +\hat{\boldsymbol{z}}| - \hat{\boldsymbol{z}}\rangle + b_-^* \langle -\hat{\boldsymbol{z}}| - \hat{\boldsymbol{z}}\rangle = b_-^*.$$

Combining these gives

$$\langle\Phi|\Psi\rangle = b_+^* a_+ + b_-^* a_-$$

which proves the result. •

Given an arbitrary state such as $|\Psi\rangle = c_+ |+\hat{\boldsymbol{z}}\rangle + c_- |-\hat{\boldsymbol{z}}\rangle$, one physical question that one could attempt to answer is what the probabilities of outcomes of measurements of a given spin component are. The method of calculating this arises from the following stipulation:

> Suppose that a spin-1/2 particle is in state corresponding to the ket $|\Psi\rangle$ and is subjected to an SG $\hat{\boldsymbol{n}}$ apparatus for some direction $\hat{\boldsymbol{n}}$. The probabilities of attaining either of the measurement outcomes are:
>
> $$\Pr(S_n = +\hbar/2) = |\langle +\hat{\boldsymbol{n}}|\Psi\rangle|^2$$
> $$\Pr(S_n = -\hbar/2) = |\langle -\hat{\boldsymbol{n}}|\Psi\rangle|^2.$$

(II.3.21)

It is important to note that the above is an imposed requirement; it cannot be derived from anything which is not equivalent to assuming that Eq. (II.3.21) is true. This is really a description of how we would like to use the mathematics to compute physical quantities; it will be invoked subsequently to relate a general state such as $|+\hat{\boldsymbol{n}}\rangle$ to the states $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$. However, we do need to check that it is consistent in the following sense.

Exercise: Suppose that a spin-1/2 particle is in the state $|+\hat{\boldsymbol{n}}\rangle$ and is subject to an SG $\hat{\boldsymbol{n}}$ apparatus. Without using Eq. (II.3.21) determine the probabilities with which each of the outcomes $S_n = +\hbar/2$ and $S_n = -\hbar/2$ occurs. Now calculate these using Eq. (II.3.21). Verify that the results are consistent. Repeat this for a particle in the state $|-\hat{\boldsymbol{n}}\rangle$.

Exercise: Suppose that a spin-1/2 particle is in the state $|\Psi\rangle = c_+ |+\hat{\boldsymbol{z}}\rangle + c_- |-\hat{\boldsymbol{z}}\rangle$ and is subjected to measurement via an SG $\hat{\boldsymbol{z}}$ apparatus. Show that the probabilities of the two outcomes are $\Pr(S_z = +\hbar/2) = |c_+|^2$ and $\Pr(S_z = -\hbar/2) = |c_-|^2$.

One consequence of the result of the preceding exercise is that the components c_+, c_- in the expression for a general state $|\Psi\rangle = c_+ |+\hat{\boldsymbol{z}}\rangle + c_- |-\hat{\boldsymbol{z}}\rangle$ must be *normalized,* i.e.

$$\boxed{|c_+|^2 + |c_-|^2 = 1} \tag{II.3.22}$$

since it is certain that at least one of $S_z = +\hbar/2$ or $S_z = -\hbar/2$ will be attained. If measurement outcomes are to be predicted correctly using a ket vector, it is essential that the ket be normalized. An alternative way of stating this is that the kets representing the physical state of a system must satisfy

$$\boxed{\langle\Psi|\Psi\rangle = 1} \tag{II.3.23}$$

Frequently it will be necessary to adjust the normalization of a ket to ensure this.

Example: Suppose that a spin-1/2 particle is in the state described by

$$|\Psi\rangle = 3\,|+\hat{\boldsymbol{z}}\rangle + 4i\,|-\hat{\boldsymbol{z}}\rangle\,.$$

Is this normalized? If not what constant A will ensure that

$$A\big(3\,|+\hat{\boldsymbol{z}}\rangle + 4i\,|-\hat{\boldsymbol{z}}\rangle\big)$$

is normalized?

Answer: Applying the left hand side of Eq. (II.3.22) gives

$$|c_+|^2 + |c_-|^2 = |3|^2 + |4i|^2 = 25.$$

Thus the ket is not normalized. Now consider

$$A\big(3\,|+\hat{\boldsymbol{z}}\rangle + 4i\,|-\hat{\boldsymbol{z}}\rangle\big),$$

in which case $c_+ = 3A$ and $c_- = 4iA$. Then $|c_+|^2 + |c_-|^2 = 25A^2$ and any complex number of the form $A = e^{i\varphi}/5$ where φ is real will work. The simplest choice is $\varphi = 0$ in which case $A = 1/5$.

It still remains to express states of the form $\{|+\hat{\boldsymbol{n}}\rangle\,, |-\hat{\boldsymbol{n}}\rangle\}$ for any unit vector $\hat{\boldsymbol{n}}$ in terms of the basis $\{|+\hat{\boldsymbol{z}}\rangle\,, |-\hat{\boldsymbol{z}}\rangle\}$. This will be convenient since calculating inner products by representing states in terms of the basis $\{|+\hat{\boldsymbol{z}}\rangle\,, |-\hat{\boldsymbol{z}}\rangle\}$ can be done using Eq. (II.3.20). To this end we establish the following key result.

Theorem: Let $\hat{\boldsymbol{n}}$ be any unit vector in three dimensional space and suppose that the spherical polar coordinates for it are θ and ϕ, i.e.

$$\hat{\boldsymbol{n}} = \sin\theta\,\cos\phi\,\hat{\boldsymbol{x}} + \sin\theta\,\sin\phi\,\hat{\boldsymbol{y}} + \cos\theta\,\hat{\boldsymbol{z}}. \tag{II.3.24}$$

Then, for a spin-1/2 particle,

$$\boxed{\begin{aligned} |+\hat{\boldsymbol{n}}\rangle &= \cos\left(\frac{\theta}{2}\right)\,|+\hat{\boldsymbol{z}}\rangle + e^{i\phi}\,\sin\left(\frac{\theta}{2}\right)\,|-\hat{\boldsymbol{z}}\rangle \\ |-\hat{\boldsymbol{n}}\rangle &= \sin\left(\frac{\theta}{2}\right)\,|+\hat{\boldsymbol{z}}\rangle - e^{i\phi}\,\cos\left(\frac{\theta}{2}\right)\,|-\hat{\boldsymbol{z}}\rangle\,. \end{aligned}} \tag{II.3.25}$$

Proof: A given direction, $\hat{\boldsymbol{n}}$, can be represented in terms of spherical coordinates θ and ϕ. Using the fact that $\{|+\hat{\boldsymbol{z}}\rangle\,, |-\hat{\boldsymbol{z}}\rangle\}$ form a basis, there exist $a_+(\theta,\phi), a_-(\theta,\phi), b_+(\theta,\phi), b_-(\theta,\phi)$ such that

$$\begin{aligned} |+\hat{\boldsymbol{n}}\rangle &= a_+(\theta,\phi)\,|+\hat{\boldsymbol{z}}\rangle + a_-(\theta,\phi)\,|+\hat{\boldsymbol{z}}\rangle \\ |-\hat{\boldsymbol{n}}\rangle &= b_+(\theta,\phi)\,|+\hat{\boldsymbol{z}}\rangle + b_-(\theta,\phi)\,|+\hat{\boldsymbol{z}}\rangle \end{aligned}$$

In an exercise that follows consider a spin-1/2 particle is in the state $|+\hat{\boldsymbol{n}}\rangle$ subjected to an SG $\hat{\boldsymbol{z}}$ measurement. Applying Eqs. (II.2.6) and (II.3.21) yields $|a_+| = |\cos(\theta/2)|$ and $|a_-| = |\sin(\theta/2)|$. Similarly considering a particle in the state $|-\hat{\boldsymbol{n}}\rangle$ subjected to an SG $\hat{\boldsymbol{z}}$ measurement gives $|b_+| = |\sin(\theta/2)|$ and $|b_-| = |\cos(\theta/2)|$. By eliminating a global phase (described a few pages later), we can always arrange for

$$\begin{aligned} a_+(\theta,\phi) &= \cos(\theta/2) \\ b_+(\theta,\phi) &= \sin(\theta/2) \\ a_-(\theta,\phi) &= e^{i\alpha}\,\sin(\theta/2) \\ b_-(\theta,\phi) &= e^{i\beta}\,\cos(\theta/2) \end{aligned}$$

where α and β are functions of θ and ϕ. Now consider a particle in the state $|+\hat{\boldsymbol{n}}\rangle$ subjected to an SG $\hat{\boldsymbol{x}}$ measurement. Then Eqs. (II.2.6) implies that $\Pr(S_x = +\hbar/2) = (1+\cos\phi\sin\theta)/2$. On the other hand the above inference implies that

$$|+\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}}\,\left\{|+\hat{\boldsymbol{z}}\rangle + e^{i\alpha}\,|-\hat{\boldsymbol{z}}\rangle\right\}$$

and Eqs. (II.3.21) yields

$$\cos\left(\alpha(\pi/2,0) - \alpha(\theta,\phi)\right) = \cos\phi.$$

This implies that

$$\alpha(\theta,\phi) - \alpha(\pi/2,0) = \phi + 2\pi n$$

where n is an irrelevant integer. Thus

$$\alpha(\theta,\phi) = \phi + \gamma$$

where γ is independent of $\hat{\boldsymbol{n}}$. A similar argument involving the state $|-\hat{\boldsymbol{n}}\rangle$ gives

$$\beta(\theta,\phi) = \phi + \delta$$

where δ is independent of $\hat{\boldsymbol{n}}$. Requiring that $|-\hat{\boldsymbol{n}}\rangle$ be orthogonal to $|+\hat{\boldsymbol{n}}\rangle$ results in

$$\delta = \gamma + \pi.$$

At this stage the states have been restricted to:

$$|+\hat{\boldsymbol{n}}\rangle = \cos\left(\frac{\theta}{2}\right)|+\hat{\boldsymbol{z}}\rangle + e^{i\phi}\, e^{i\gamma}\, \sin\left(\frac{\theta}{2}\right)|-\hat{\boldsymbol{z}}\rangle$$
$$|-\hat{\boldsymbol{n}}\rangle = \sin\left(\frac{\theta}{2}\right)|+\hat{\boldsymbol{z}}\rangle - e^{i\phi}\, e^{i\gamma}\, \cos\left(\frac{\theta}{2}\right)|-\hat{\boldsymbol{z}}\rangle\,.$$

It is clear that for calculations using Eqs. (II.3.21) γ is irrelevant, provided that it is the same for all $\hat{\boldsymbol{n}}$. As a matter of convenience we choose $\gamma = 0$, thus yielding the desired expressions. •

Exercise: Suppose that a spin-1/2 particle is in the state $|+\hat{\boldsymbol{n}}\rangle = a_+\,|+\hat{\boldsymbol{z}}\rangle + a_-\,|-\hat{\boldsymbol{z}}\rangle$ where $\hat{\boldsymbol{n}}$ is any unit vector. Suppose that is subjected to an SG $\hat{\boldsymbol{z}}$ measurement. Using Eqs. (II.2.6) and (II.3.21) show that $|a_+| = \cos(\theta/2)$ and $|a_-| = \sin(\theta/2)$.

Exercise: Suppose that a spin-1/2 particle is in the state $|-\hat{\boldsymbol{n}}\rangle = b_+\,|+\hat{\boldsymbol{z}}\rangle + b_-\,|-\hat{\boldsymbol{z}}\rangle$ where $\hat{\boldsymbol{n}}$ is any unit vector. Suppose that is subjected to an SG $\hat{\boldsymbol{z}}$ measurement. Using Eqs. (II.2.6) and (II.3.21) show that $|b_+| = \sin(\theta/2)$ and $|b_-| = \cos(\theta/2)$.

One can perform some consistency checks on the expressions of Eqs. (II.3.21). For example, it is readily seen that when $\hat{\boldsymbol{n}} = \hat{\boldsymbol{z}}$, the expression for $|+\hat{\boldsymbol{n}}\rangle$ simplifies as expected.

Example: Suppose that a spin-1/2 particle is in the state $|+\hat{\boldsymbol{y}}\rangle$ and is subjected to an SG $\hat{\boldsymbol{z}}$ measuring device. Determine the probabilities with which the two measurement outcomes occur.

Answer: First Eq. (II.3.21) with $|\Psi\rangle = |+\hat{\boldsymbol{y}}\rangle$ implies

$$\Pr\left(S_z = +\hbar/2\right) = |\langle +\hat{\boldsymbol{z}}| + \hat{\boldsymbol{y}}\rangle|^2.$$

Now use Eq. (II.3.25) to express $|+\hat{\boldsymbol{y}}\rangle$ in terms of $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$. The spherical polar coordinates for $\hat{\boldsymbol{y}}$ are $\theta = \pi/2$ and $\phi = \pi/2$, giving

$$|+\hat{\boldsymbol{y}}\rangle = \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle + \frac{i}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle\,.$$

Applying Eq. (II.3.20) gives

$$\langle +\hat{\boldsymbol{z}}| + \hat{\boldsymbol{y}}\rangle = \frac{1}{\sqrt{2}}$$

and thus

$$\Pr\left(S_z = +\hbar/2\right) = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}.$$

Similarly

$$\Pr\left(S_z = -\hbar/2\right) = |\langle -\hat{\boldsymbol{z}}| + \hat{\boldsymbol{y}}\rangle|^2.$$

Now

$$\langle -\hat{\boldsymbol{z}}| + \hat{\boldsymbol{y}}\rangle = \frac{i}{\sqrt{2}}$$

and thus

$$\Pr\left(S_z = -\hbar/2\right) = \left|\frac{i}{\sqrt{2}}\right|^2 = \frac{1}{2}.$$

3.1 Equivalence between mathematical and physical states

The kets $|+\hat{\boldsymbol{n}}\rangle$ and $|-\hat{\boldsymbol{n}}\rangle$ have well defined *physical* meanings in terms of SG $\hat{\boldsymbol{n}}$ measurements. To recap, the physical meaning of $|+\hat{\boldsymbol{z}}\rangle$ (or $|-\hat{\boldsymbol{z}}\rangle$) is that when a particle in this state is subjected to an SG $\hat{\boldsymbol{n}}$ measurement, the outcome $S_z = +\hbar/2$ (or $S_z = -\hbar/2$) will occur with certainty. An alternative physical interpretation is that if an SG $\hat{\boldsymbol{n}}$ measurement yields outcome $S_z = +\hbar/2$ (or $S_z = -\hbar/2$), then that state of the particle immediately after measurement is $|+\hat{\boldsymbol{n}}\rangle$ (or $|-\hat{\boldsymbol{z}}\rangle$). These are illustrated in Figs. II.2.8 and II.2.9. Thus in terms of the basis, $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$, the general states for which such physical meanings based on measurement outcomes can be inferred are those given in Eqs. (II.3.25). However, the vector space will admit all superpositions of the form

$$|\Psi\rangle = c_+ |+\hat{\boldsymbol{z}}\rangle + c_- |-\hat{\boldsymbol{z}}\rangle$$

where c_+ and c_- are any complex numbers. The question now is whether physical meanings of the type given above can be associated with such an arbitrary state in the vector space. This amounts to asking whether every state of this form can be expressed in the form of those in Eqs. (II.3.25). If c_+ and c_- are unrestricted, then this is impossible. However, normalized states, i.e.

$$|c_+|^2 + |c_-|^2 = 1,$$

are the only possibilities that are consistent with the standard interpretation in terms of probabilities of measurement outcomes. For these the following result is important.

Theorem: If $|\Psi\rangle = c_+ |+\hat{\boldsymbol{z}}\rangle + c_- |-\hat{\boldsymbol{z}}\rangle$ satisfies

$$|c_+|^2 + |c_-|^2 = 1,$$

then there exist real numbers θ, ϕ, φ such that

$$c_+ = e^{i\varphi} \cos\left(\frac{\theta}{2}\right)$$
$$c_- = e^{i\varphi}\, e^{i\phi} \sin\left(\frac{\theta}{2}\right).$$

Proof: Any complex number can be expressed as $z = r\,e^{i\theta}$ where r and θ are real. It follows that

$$c_+ = a\,e^{i\alpha}$$
$$c_- = b\,e^{i\beta}$$

where $a, b, \alpha, \beta \in \mathbb{R}$. Normalization, Eq. (II.3.22) gives

$$|c_+|^2 + |c_-|^2 = 1 \quad \Rightarrow \quad a^2 + b^2 = 1$$

and it is always possible to find θ such that

$$a = \cos\left(\frac{\theta}{2}\right)$$
$$b = \sin\left(\frac{\theta}{2}\right)$$

Thus

$$c_+ = e^{i\alpha}\,\cos\left(\frac{\theta}{2}\right)$$
$$c_- = e^{i\alpha}\,e^{i(\beta-\alpha)}\,\sin\left(\frac{\theta}{2}\right)$$

and setting $\varphi := \alpha$ and $\phi := \beta - \alpha$ proves the result. •

It follows that the most general normalized state can be expressed as

$$|\Psi\rangle = e^{i\varphi}\left\{\cos\left(\frac{\theta}{2}\right)\,|+\hat{\boldsymbol{z}}\rangle + e^{i\phi}\,\sin\left(\frac{\theta}{2}\right)\,|-\hat{\boldsymbol{z}}\rangle\right\}. \tag{II.3.26}$$

The overall constant factor of the form $e^{i\varphi}$ is called a *global phase.*

Exercise: Consider the two states

$$|\Psi_1\rangle = e^{i\varphi}\left\{\cos\left(\frac{\theta}{2}\right)\,|+\hat{\boldsymbol{z}}\rangle + e^{i\phi}\,\sin\left(\frac{\theta}{2}\right)\,|-\hat{\boldsymbol{z}}\rangle\right\} \quad \text{and}$$
$$|\Psi_2\rangle = \cos\left(\frac{\theta}{2}\right)\,|+\hat{\boldsymbol{z}}\rangle + e^{i\phi}\,\sin\left(\frac{\theta}{2}\right)\,|-\hat{\boldsymbol{z}}\rangle\,.$$

Show that, if a particle in each state is subjected to a SG $\hat{\boldsymbol{n}}$ measurement for any direction $\hat{\boldsymbol{n}}$, the probabilities of the measurement outcome $S_n = +\hbar/2$ are identical for the two particles.

Example: Consider two particles in the states

$$|\Psi_1\rangle = \frac{1}{\sqrt{2}}\left\{|+\hat{\boldsymbol{z}}\rangle + |-\hat{\boldsymbol{z}}\rangle\right\} \quad \text{and}$$

$$|\Psi_2\rangle = e^{i\varphi}\frac{1}{\sqrt{2}}\left\{|+\hat{\boldsymbol{z}}\rangle + |-\hat{\boldsymbol{z}}\rangle\right\}.$$

Each particle is subjected to a SG $\hat{\boldsymbol{z}}$ measurement. Determine the probability with which $S_z = +\hbar/2$.

Answer: For a particle in state $|\Psi\rangle$,

$$\Pr\left(S_z = +\hbar/2\right) = |\langle+\hat{\boldsymbol{z}}|\Psi\rangle|^2.$$

For the particle in state $|\Psi_1\rangle$,

$$\Pr\left(S_z = +\hbar/2\right) = |\langle+\hat{\boldsymbol{z}}|\Psi_1\rangle|^2$$

and

$$\langle+\hat{\boldsymbol{z}}|\Psi_1\rangle = \frac{1}{\sqrt{2}}$$

giving

$$\Pr\left(S_z = +\hbar/2\right) = \frac{1}{2}.$$

For the particle in state $|\Psi_2\rangle$,

$$\Pr\left(S_z = +\hbar/2\right) = |\langle+\hat{\boldsymbol{z}}|\Psi_2\rangle|^2$$

and

$$\langle+\hat{\boldsymbol{z}}|\Psi_2\rangle = e^{i\varphi}\frac{1}{\sqrt{2}}$$

giving

$$\Pr\left(S_z = +\hbar/2\right) = e^{-i\varphi}\frac{1}{\sqrt{2}}e^{i\varphi}\frac{1}{\sqrt{2}} = \frac{1}{2}.$$

These are identical.

Thus the value of φ never affects the probabilities of the outcomes of any measurements. In this sense it is physically irrelevant and can be omitted. Ignoring global phases, the most general normalized state has form:

$$|\Psi\rangle = \cos\left(\frac{\theta}{2}\right)|+\hat{\boldsymbol{z}}\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|-\hat{\boldsymbol{z}}\rangle \tag{II.3.27}$$

for some angles θ and ϕ corresponding to a direction $\hat{\boldsymbol{n}}$; this is exactly the state $|+\hat{\boldsymbol{n}}\rangle$ of Eq. (II.3.25). It follows that any normalized ket, $|\Psi\rangle$ has a physical interpretation in terms of the outcomes of an SG $\hat{\boldsymbol{n}}$ measurement for some specific unit vector $\hat{\boldsymbol{n}}$, which depends on

$|\Psi\rangle$, in the sense that the measurement will yield the outcome $S_n = +\hbar/2$ with certainty. In general for spin-1/2 particle:

> Any normalized state, $|\Psi\rangle$, i.e. satisfying $\langle\Psi|\Psi\rangle = 1$, has the physical interpretation that there is some measurement such that, when it is performed on a particle in this state exactly one of the possible outcomes is attained with certainty.

In the case of spin-1/2 particles the possible measurements are Stern-Gerlach measurements of spin components; for a given state $|\Psi\rangle$, it is necessary to find θ and ϕ so that the state can be expressed as $|+\hat{\boldsymbol{n}}\rangle$ according to Eq. (II.3.25). The relevant measurement is then that of the spin component S_n, which is equivalent to submitting the particle to an SG $\hat{\boldsymbol{n}}$ apparatus. The definite measurement outcome which will result is $S_n = +\hbar/2$.

3.2 Bras: Calculating inner products

Determining probabilities of measurement outcomes in quantum mechanics hinges on the ability to calculate inner products of ket vectors. While the definition of Eq. (II.3.20) is adequate for this, there is a more convenient technique which only uses standard vector and matrix operations. Recall that for

$$|\Psi\rangle = a_+ |+\hat{\boldsymbol{z}}\rangle + a_- |-\hat{\boldsymbol{z}}\rangle$$
$$|\Phi\rangle = b_+ |+\hat{\boldsymbol{z}}\rangle + b_- |-\hat{\boldsymbol{z}}\rangle$$

the inner product is

$$\langle\Phi|\Psi\rangle = b_+^* a_+ + b_-^* a_-.$$

It would be convenient to be able to compute this directly from the column vectors representing the kets in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis:

$$|\Psi\rangle \leftrightarrow \begin{pmatrix} a_+ \\ a_- \end{pmatrix}$$
$$|\Phi\rangle \leftrightarrow \begin{pmatrix} b_+ \\ b_- \end{pmatrix}.$$

There is no previously defined operation on such column vectors which will result in a scalar. The basic idea will be to convert one of the vectors to a row vector and apply matrix multiplication to the two.

Mathematical Tools: Row vectors and dual vector spaces

It is conventional to represent any real vector in two dimensions, $\mathbf{u} = u_x \hat{\boldsymbol{e}}_1 + u_y \hat{\boldsymbol{e}}_2$, via a column vector

$$\mathbf{u} \leftrightarrow \begin{pmatrix} u_x \\ u_y \end{pmatrix}.$$

An alternative possibility is to use a row vector representation

$$\mathbf{u} \leftrightarrow (u_x \;\; u_y)\,.$$

There is clearly a one-to-one correspondence between such row and column vectors and for this reason these two type of vectors are often not regarded as distinct. However, their appearance is different and they are distinct mathematical entities. This can be illustrated by the process of adding two row vectors, which is defined via

$$(u_x \;\; u_y) + (v_x \;\; v_y) := (u_x + v_x \;\; u_y + v_y)\,. \tag{II.3.28}$$

This raises the question of the possibility of the following addition

$$(u_x \;\; u_y) + \begin{pmatrix} v_x \\ v_y \end{pmatrix} =?$$

Neither the definition of addition of row vectors or that of column vectors is any use here. To complete this sum, it would be necessary to *convert one of the vectors into the same type as the other*. Thus it is necessary to regard row and column vectors as distinct mathematical quantities. Note that it is possible to define scalar multiplication of a row vector,

$$\alpha\,(u_x \;\; u_y) := (\alpha u_x \;\; \alpha u_y)\,. \tag{II.3.29}$$

Then the set of all row vectors equipped with addition as defined by Eq. (II.3.28) and multiplication by Eq. (II.3.29) can be shown to satisfy the requirements of a vector space.

One purpose which the row vectors serve, and this is their essential relationship to the column vectors, is as *operators on the column vectors*. The action of a row vector $(u_x \;\; u_y)$ on a column vector is typically defined, so as to produce an inner product, via:

$$(u_x \;\; u_y) \begin{pmatrix} v_x \\ v_y \end{pmatrix} = u_x v_x + u_y v_y. \tag{II.3.30}$$

By this definition, a row vector operates on column vectors so as to produce a scalar; alternatively the row vector acts as a function which takes an entire column vector as its input and produces a scalar. The operation can be denoted by replacing the column vector by the actual two-dimensional vector which it represents, i.e.

$$(u_x \;\; u_y)\,\mathbf{v} = u_x v_x + u_y v_y.$$

Schematically for a given vector,

$$(u_x \;\; u_y) : \text{column vector } \mathbf{v} \rightarrow \text{scalar}$$

and for the entire set of vectors, V

$$\boxed{(u_x \;\; u_y) : V \rightarrow \mathbb{R}.} \tag{II.3.31}$$

Example: Consider the row vector $(-2\ \ 3)$. Then for $\mathbf{v} = 5\hat{\boldsymbol{e}}_1 + 1\hat{\boldsymbol{e}}_2$,

$$\begin{aligned}(-2\ \ 3)\,\mathbf{v} &= (-2\ \ 3)\begin{pmatrix}5\\1\end{pmatrix}\\ &= -7.\end{aligned}$$

Similarly for $\boldsymbol{w} = 1\hat{\boldsymbol{e}}_1 + -2\hat{\boldsymbol{e}}_2$,

$$\begin{aligned}(-2\ \ 3)\,\boldsymbol{w} &= (-2\ \ 3)\begin{pmatrix}1\\-2\end{pmatrix}\\ &= -8.\end{aligned}$$

Thus this particular row vector maps

$$\begin{aligned}\mathbf{v} &\to -7 \quad \text{and}\\ \boldsymbol{w} &\to -8.\end{aligned}$$

The most important property of the operation of a row vector is that it is linear, i.e. for any $\alpha, \beta \in \mathbb{R}$ and any two vectors $\mathbf{v}, \boldsymbol{w}$

$$(u_x\ \ u_y)\,(\alpha\mathbf{v} + \beta\boldsymbol{w}) = \alpha\,(u_x\ \ u_y)\,\mathbf{v} + \beta\,(u_x\ \ u_y)\,\boldsymbol{w}. \tag{II.3.32}$$

Example: Consider the row vector $(-2\ \ 3)$ and the vectors $\mathbf{v} = 5\hat{\boldsymbol{e}}_1 + 1\hat{\boldsymbol{e}}_2, \boldsymbol{w} = 1\hat{\boldsymbol{e}}_1 + -2\hat{\boldsymbol{e}}_2$. Then

$$\begin{aligned}(-2\ \ 3)\,(-4\mathbf{v} + 2\boldsymbol{w}) &= (-2\ \ 3)\begin{pmatrix}-4v_x + 2w_x\\-4v_y + 2w_y\end{pmatrix}\\ &= (-2\ \ 3)\begin{pmatrix}-18\\-8\end{pmatrix}\\ &= 12.\end{aligned}$$

For individual vectors

$$\begin{aligned}(-2\ \ 3)\,\mathbf{v} &= 7\\ (-2\ \ 3)\,\boldsymbol{w} &= -8\end{aligned}$$

and thus

$$-4\,(-2\ \ 3)\,\mathbf{v} + 2\,(-2\ \ 3)\,\boldsymbol{w} = 12.$$

The set of all operators that map the vector space V (of column vectors) to the scalars linearly is called the dual vector space and is denoted V^*.

For two dimensional real vectors there is clearly a one-to-one correspondence between column vectors and row vectors:

$$(u_x\ \ u_y) \leftrightarrow \begin{pmatrix}u_x\\u_y\end{pmatrix}.$$

This correspondence happens to depend on the basis used to describe the vector components; under a change of basis there is no guarantee that the action of a given row vector on a column vector will always produce the same scalar. However, provided that orthonormal basis vectors are used, it can be shown that the correspondence above guarantees that the operation of a given row vector on a column vector always returns the same result. The true method for associating a row vector with a column vector emerges after a consideration of the matrix transposition and adjoint operations.

Before considering these operations, we review matrix multiplication. Two matrices A and B can be multiplied in the order AB whenever the number of columns in A equals the number of rows in B. For example consider the two matrices

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix} \quad \text{and}$$
$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{pmatrix}.$$

Then the product is defined as:

$$AB := \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \end{pmatrix}. \tag{II.3.33}$$

Here the entry in the first row and first column of AB is attained by operating with the row vector formed from the first row of A with the column vector formed by the first column of B. Similarly, the entry in the first row and second column of AB is attained by using the first row of A and the second column of B. This type of matrix multiplication can be generalized to matrices of arbitrary size, provided that the number of columns in A equals the number of rows in B.

Example: Let

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Determine AB and BA and check whether these are equal.

Answer: Using Eq. (II.3.33),

$$AB = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$= \begin{pmatrix} 1\times 1 + 1\times 0 & 0\times 1 + 1\times -1 \\ 3\times 1 + 3\times 0 & 3\times 1 + 3\times -1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 3 & 0 \end{pmatrix}.$$

Similarly, using Eq. (II.3.33),

$$BA = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}$$
$$= \begin{pmatrix} 1 \times 1 + 0 \times 3 & 1 \times 1 + 0 \times 3 \\ 0 \times 1 + -1 \times 3 & 0 \times 1 + -1 \times 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -3 & -3 \end{pmatrix}$$

and clearly $AB \neq BA$.

Note that the action of a row vector on a column vector is a special example of matrix multiplication,

$$(a_1 \;\; a_2) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 b_1 + a_2 b_2.$$

This association would be complete if there was a method prescribed for converting a column vector into a row vector. In fact, a more general tool of this type is available. The *transpose* of a matrix A, denoted A^T is defined as follows. If A is an m (number of rows) by n (number of columns) matrix then the transpose A^T is an n by m matrix with entries $(A^T)_{ij} = A_{ji}$. For example:

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \\ A_{13} & A_{23} \end{pmatrix}$$

Example: Let

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}.$$

Determine A^T.

Answer: A is a 2 by 2 matrix and thus A^T will also be a 2 by 2 matrix. Using $(A^T)_{ij} = A_{ji}$ gives

$$A^T = \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix}.$$

One special example is the transpose of a vector $\mathbf{u} = u_x \hat{\boldsymbol{e}}_1 + u_y \hat{\boldsymbol{e}}_2$,

$$\mathbf{u}^T = \begin{pmatrix} u_x \\ u_y \end{pmatrix}^T = (u_x \;\; u_y).$$

Thus given two vectors which are elements of a two dimensional real vector space

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$
$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

it is clear that

$$\mathbf{u}^T\mathbf{v} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}^T \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \mathbf{u} \cdot \mathbf{v}.$$

The important point is that the dot product can be computed by just using the operations of matrix multiplication and transposition:

$$\boxed{\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v}.} \tag{II.3.34}$$

The transpose has the following useful properties:

Theorem: For matrices A and B and any $\lambda \in \mathbb{R}$

$$(A+B)^T = A^T + B^T \tag{II.3.35a}$$
$$(\lambda A)^T = \lambda A^T \tag{II.3.35b}$$
$$(AB)^T = B^T A^T \tag{II.3.35c}$$
$$\left(A^T\right)^T = A. \tag{II.3.35d}$$

Proof: Exercise.•

Exercise: Let

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Verify each of Eqs. (II.3.35), using $\lambda = 4$.

These tools can be extended to complex vector spaces and their inner products by performing both complex conjugation and transposition. To this end define the *adjoint* operation (or alternatively complex conjugate transpose) on a matrix A, denoted $A^\dagger$, as follows. The matrix entries of the adjoint are $(A^\dagger)_{ij} = A^*_{ji}$. For example:

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}^\dagger = \begin{pmatrix} A^*_{11} & A^*_{21} \\ A^*_{12} & A^*_{22} \\ A^*_{13} & A^*_{23} \end{pmatrix}.$$

In terms of matrix entries $A^\dagger = A^{*T}$.

Example: Let

$$A = \begin{pmatrix} 1 & i \\ 3 & e^{i\phi} \end{pmatrix}$$

and determine $A^\dagger$.

Answer: Using $A^\dagger = A^{*T}$ gives

$$\begin{aligned} A^\dagger &= \begin{pmatrix} 1 & i \\ 3 & e^{i\phi} \end{pmatrix}^{*T} \\ &= \begin{pmatrix} 1 & -i \\ 3 & e^{-i\phi} \end{pmatrix}^{T} \\ &= \begin{pmatrix} 1 & 3 \\ -i & e^{-i\phi} \end{pmatrix}. \end{aligned}$$

The relevance for inner products of two dimensional complex vectors is that

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}^\dagger \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1^* & a_2^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1^* b_1 + a_2^* b_2$$

which is exactly the inner product as previously defined.

The adjoint has the following properties:

Theorem: For matrices A and B and any $\lambda \in \mathbb{C}$

$$(A+B)^\dagger = A^\dagger + B^\dagger \tag{II.3.36}$$

$$(\lambda A)^\dagger = \lambda^* A^\dagger \tag{II.3.37}$$

$$(AB)^\dagger = B^\dagger A^\dagger \tag{II.3.38}$$

$$\left(A^\dagger\right)^\dagger = A. \tag{II.3.39}$$

Proof: Exercise. •

Note that the order of multiplication is reversed after performing the adjoint operation. ■

So far the state of a quantum system has been represented by a ket, a two dimensional complex vector. These are conventionally represented by column vectors

$$|\Psi\rangle = b_+ |+\hat{\boldsymbol{z}}\rangle + b_- |-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} b_+ \\ b_- \end{pmatrix}.$$

A vector in the dual vector space, which is represented by a row vector, is called a "bra" and is generically denoted by the symbol $\langle\Phi|$. Just as for the case of two-dimensional vectors, the role of any bra vector is to operate on a ket vector and produce a complex scalar. Thus $\langle\Phi|\,|\Psi\rangle$ is understood to mean that the bra $\langle\Phi|$ acts on the ket $|\Psi\rangle$ to produce a complex number. Schematically

$$\boxed{\langle\Phi| : \text{all kets} \to \mathbb{C}.} \tag{II.3.40}$$

We can view a bra as a device that takes a ket as an input and produces a complex number; this is illustrated schematically in Fig. II.3.3.

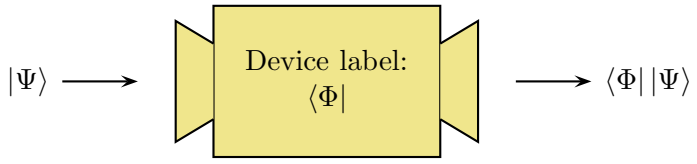


Figure II.3.3: A bra, $\langle\Phi|$ can be regarded as a device which takes a ket as an input, processes it and produces a complex number as an output.

By definition, and as illustrated in Fig. II.3.4, bra vectors act linearly. Thus

$$\boxed{\langle\Phi|\left(\alpha\,|\Psi_1\rangle+\beta\,|\Psi_2\rangle\right)=\alpha\,\langle\Phi|\,|\Psi_1\rangle+\beta\,\langle\Phi|\,|\Psi_2\rangle\,.} \tag{II.3.41}$$

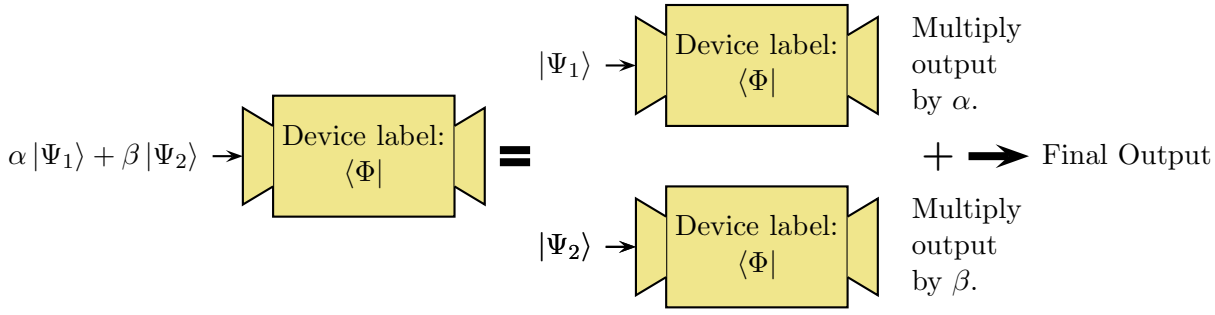


Figure II.3.4: A schematic illustrating the linearity of the operation of the bra $\langle\Phi|$.

Example: Suppose that a particular bra, $\langle\Phi|$, is known to act on the basis kets of $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ as follows

$$\langle\Phi|\,|+\hat{\boldsymbol{z}}\rangle=\frac{1}{\sqrt{2}}$$
$$\langle\Phi|\,|-\hat{\boldsymbol{z}}\rangle=\frac{1}{\sqrt{2}}$$

Determine $\langle\Phi|\,|+\hat{\boldsymbol{x}}\rangle$ and $\langle\Phi|\,|-\hat{\boldsymbol{x}}\rangle$.

Answer: First for $\hat{\boldsymbol{x}}$, the spherical coordinates are $\theta = \pi/2$ and $\phi = 0$. Thus

$$|+\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle$$
$$|-\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle - \frac{1}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle\,.$$

Then the linearity of the action of the bra implies that

$$\begin{aligned}\langle\Phi|\,|+\hat{\boldsymbol{x}}\rangle &= \frac{1}{\sqrt{2}}\,\langle\Phi|\,|+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}}\,\langle\Phi|\,|-\hat{\boldsymbol{z}}\rangle \\ &= \frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}} \\ &= 1.\end{aligned}$$

Similarly

$$\begin{aligned}\langle\Phi|\,|-\hat{\boldsymbol{x}}\rangle &= \frac{1}{\sqrt{2}}\,\langle\Phi|\,|+\hat{\boldsymbol{z}}\rangle - \frac{1}{\sqrt{2}}\,\langle\Phi|\,|-\hat{\boldsymbol{z}}\rangle \\ &= \frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\,\frac{1}{\sqrt{2}} \\ &= 0.\end{aligned}$$

Bra vectors may be combined linearly. Thus given $\langle\Phi_1|\,, \langle\Phi_2|$ and complex numbers α, β, the meaning of $\alpha\,\langle\Phi_1| + \beta\,\langle\Phi_2|$ is attained by defining its action on a ket as

$$\boxed{\Big(\alpha\,\langle\Phi_1| + \beta\,\langle\Phi_2|\Big)\,|\Psi\rangle := \alpha\,\langle\Phi_1|\,|\Psi\rangle + \beta\,\langle\Phi_2|\,|\Psi\rangle\,.} \tag{II.3.42}$$

Example: Suppose that particular bras, $\langle\Phi_1|\,, \langle\Phi_2|$ are known to act on $|+\hat{\boldsymbol{z}}\rangle$ as follows

$$\langle\Phi_1|\,|+\hat{\boldsymbol{z}}\rangle = \frac{3}{\sqrt{5}}$$
$$\langle\Phi_2|\,|-\hat{\boldsymbol{z}}\rangle = \frac{4i}{\sqrt{5}}$$

Determine $\Big(\frac{1}{\sqrt{2}}\,\langle\Phi_1| + \frac{1}{\sqrt{2}}\,\langle\Phi_2|\Big)\,|+\hat{\boldsymbol{z}}\rangle\,.$

Answer: Then the addition of bras implies that

$$\begin{aligned}\Big(\frac{1}{\sqrt{2}}\,\langle\Phi_1| + \frac{1}{\sqrt{2}}\,\langle\Phi_2|\Big)\,|+\hat{\boldsymbol{x}}\rangle &= \frac{1}{\sqrt{2}}\,\langle\Phi_1|\,|+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}}\,\langle\Phi_2|\,|-\hat{\boldsymbol{z}}\rangle \\ &= \frac{1}{\sqrt{2}}\,\frac{3}{5} + \frac{1}{\sqrt{2}}\,\frac{4i}{5} \\ &= \frac{3+4i}{5\sqrt{2}}.\end{aligned}$$

It remains to associate a bra and ket vectors; that is, given any ket $|\Phi\rangle$ what is the unique associated bra, $\langle\Phi|$? We would like the association to mirror that for the column and row vectors. Thus if

$$|\Phi\rangle = b_+ |+\hat{\boldsymbol{z}}\rangle + b_- |-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} b_+ \\ b_- \end{pmatrix}$$

then the associated bra will be defined to be

$$\langle\Phi| \leftrightarrow \begin{pmatrix} b_+^* & b_-^* \end{pmatrix}.$$

In fact, this association is really defined via the inner product:

> For any ket $|\Phi\rangle$, the associated bra $\langle\Phi|$ is such that
>
> $$\langle\Phi|\,|\Psi\rangle = \langle\Phi|\Psi\rangle \tag{II.3.43}$$
>
> where $|\Psi\rangle$ is any ket.

Using this definition the action of a bra on a ket is written using the notation for the inner product (i.e. there is only one and not two vertical line between the bra and ket label). This correspondence utilizes the adjoint and this is expressed via:

$$\langle\Phi| = |\Phi\rangle^\dagger \quad \text{and} \tag{II.3.44}$$

$$|\Phi\rangle = \langle\Phi|^\dagger. \tag{II.3.45}$$

This is frequently convenient for formal manipulations. Elementary, but important, consequences of Eq. (II.3.43) and the fact that $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ are orthonormal are

$$\langle+\hat{\boldsymbol{z}}|+\hat{\boldsymbol{z}}\rangle = 1 \tag{II.3.46a}$$

$$\langle+\hat{\boldsymbol{z}}|-\hat{\boldsymbol{z}}\rangle = 0 \tag{II.3.46b}$$

$$\langle-\hat{\boldsymbol{z}}|+\hat{\boldsymbol{z}}\rangle = 0 \tag{II.3.46c}$$

$$\langle-\hat{\boldsymbol{z}}|-\hat{\boldsymbol{z}}\rangle = 1. \tag{II.3.46d}$$

These enable explicit calculations of the action of bras on kets.

Example: The bra $\langle+\hat{\boldsymbol{z}}|$ acts on the ket $|\Psi\rangle = b_+ |+\hat{\boldsymbol{z}}\rangle + b_- |-\hat{\boldsymbol{z}}\rangle$ as follows

$$\begin{aligned} \langle+\hat{\boldsymbol{z}}|\,|\Psi\rangle &= \langle+\hat{\boldsymbol{z}}|\left(b_+ |+\hat{\boldsymbol{z}}\rangle + b_- |-\hat{\boldsymbol{z}}\rangle\right) \\ &= b_+ \langle+\hat{\boldsymbol{z}}|\,|+\hat{\boldsymbol{z}}\rangle + b_- \langle+\hat{\boldsymbol{z}}|\,|-\hat{\boldsymbol{z}}\rangle \\ &= b_+ \langle+\hat{\boldsymbol{z}}|+\hat{\boldsymbol{z}}\rangle + b_- \langle+\hat{\boldsymbol{z}}|-\hat{\boldsymbol{z}}\rangle \\ &= b_+. \end{aligned}$$

Any bra can be expressed as a linear combination of $\langle+\hat{\boldsymbol{z}}|$ and $\langle-\hat{\boldsymbol{z}}|$,

$$\langle\Phi| = b_+ \langle+\hat{\boldsymbol{z}}| + b_- \langle-\hat{\boldsymbol{z}}|$$

where $b_\pm$ are complex numbers. The representation of this bra in terms of a row vector is:

$$\boxed{\langle\Phi| = b_+ \langle+\hat{\boldsymbol{z}}| + b_- \langle-\hat{\boldsymbol{z}}| \leftrightarrow (b_+ \;\; b_-)\,.} \tag{II.3.47}$$

The following special cases are important.

$$\boxed{\langle+\hat{\boldsymbol{z}}| \leftrightarrow (1 \;\; 0) \quad \text{and} \quad \langle-\hat{\boldsymbol{z}}| \leftrightarrow (0 \;\; 1)} \tag{II.3.48}$$

In general the bra associated with a ket can be determined by applying the following theorem.

Theorem: If $|\Phi\rangle = \alpha\,|\Phi_1\rangle + \beta\,|\Phi_2\rangle$ then

$$\langle\Phi| = \alpha^*\,\langle\Phi_1| + \beta^*\,\langle\Phi_2|\,. \tag{II.3.49}$$

Proof: Let $|\Psi\rangle$ be any ket. Then

$$\begin{aligned}\langle\Phi|\,|\Psi\rangle &= \langle\Phi|\Psi\rangle \\ &= \alpha^*\,\langle\Phi_1|\Psi\rangle + \beta^*\,\langle\Phi_2|\Psi\rangle\,,\end{aligned}$$

which follows from the anti-linear property of the inner product, Eq (II.3.16). Thus

$$\begin{aligned}\langle\Phi|\,|\Psi\rangle &= \langle\Phi|\Psi\rangle \\ &= \alpha^*\,\langle\Phi_1|\Psi\rangle + \beta^*\,\langle\Phi_2|\Psi\rangle\,, \\ &= \alpha^*\,\langle\Phi_1|\,|\Psi\rangle + \beta^*\,\langle\Phi_2|\,|\Psi\rangle \\ &= [\alpha^*\,\langle\Phi_1| + \beta^*\,\langle\Phi_2|]\,|\Psi\rangle\,.\end{aligned}$$

This is true for any ket, $|\Psi\rangle$ and thus

$$\langle\Phi| = \alpha^*\,\langle\Phi_1| + \beta^*\,\langle\Phi_2|\,. \qquad \bullet$$

An informal method of proving this theorem uses properties of the adjoint. Thus

$$\begin{aligned}\langle\Phi| &= |\Phi\rangle^\dagger \\ &= [\alpha\,|\Phi_1\rangle + \beta\,|\Phi_2\rangle]^\dagger \\ &= [\alpha\,|\Phi_1\rangle]^\dagger + [\beta\,|\Phi_2\rangle]^\dagger \\ &= \alpha^*\,|\Phi_1\rangle^\dagger + \beta^*\,|\Phi_2\rangle^\dagger \\ &= \alpha^*\,\langle\Phi_1| + \beta^*\,\langle\Phi_2|\,.\end{aligned}$$

This is conveniently expressed in terms of row and column representations. Thus if

$$|\Phi\rangle = \alpha\,|\Phi_1\rangle + \beta\,|\Phi_2\rangle \leftrightarrow \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

then

$$\langle\Phi| = \alpha^*\,\langle\Phi_1| + \beta^*\,\langle\Phi_2| \leftrightarrow (\alpha^* \;\; \beta^*) = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}^\dagger$$

and this could be reached using $\langle\Phi| = |\Phi\rangle^\dagger$ directly on the column representation.

Example: Let $|\Phi\rangle = a_+ |+\hat{\boldsymbol{z}}\rangle + a_- |-\hat{\boldsymbol{z}}\rangle$. Then

$$\langle\Phi| = a_+^* \langle+\hat{\boldsymbol{z}}| + a_-^* \langle-\hat{\boldsymbol{z}}| \leftrightarrow \begin{pmatrix} a_+^* & a_-^* \end{pmatrix} = \begin{pmatrix} a_+ \\ a_- \end{pmatrix}^\dagger .$$

Bra vectors are an important tool in calculating inner products and probabilities in quantum mechanics.

Example: Suppose that a particle in state $|+\hat{\boldsymbol{x}}\rangle$ is subjected to an SG $\hat{\boldsymbol{y}}$ measurement. Determine the probability with which the measurement outcome is $S_y = +\hbar/2$.

Answer: First

$$\Pr\left(S_y = +\hbar/2\right) = |\langle+\hat{\boldsymbol{y}}| + \hat{\boldsymbol{x}}\rangle|^2.$$

Then

$$|+\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}} |+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}} |-\hat{\boldsymbol{z}}\rangle$$
$$|+\hat{\boldsymbol{y}}\rangle = \frac{1}{\sqrt{2}} |+\hat{\boldsymbol{z}}\rangle + \frac{i}{\sqrt{2}} |-\hat{\boldsymbol{z}}\rangle$$

Now

$$\langle+\hat{\boldsymbol{y}}| + \hat{\boldsymbol{x}}\rangle = \langle+\hat{\boldsymbol{y}}| \, |+\hat{\boldsymbol{x}}\rangle$$

and

$$\begin{aligned} \langle+\hat{\boldsymbol{y}}| &= |+\hat{\boldsymbol{y}}\rangle^\dagger \\ &= \left(\frac{1}{\sqrt{2}} |+\hat{\boldsymbol{z}}\rangle + \frac{i}{\sqrt{2}} |-\hat{\boldsymbol{z}}\rangle\right)^\dagger \\ &= \left(\frac{1}{\sqrt{2}}\right)^* |+\hat{\boldsymbol{z}}\rangle^\dagger + \left(\frac{i}{\sqrt{2}}\right)^* |-\hat{\boldsymbol{z}}\rangle^\dagger \\ &= \frac{1}{\sqrt{2}} \langle+\hat{\boldsymbol{z}}| - \frac{i}{\sqrt{2}} \langle-\hat{\boldsymbol{z}}| . \end{aligned}$$

Thus

$$\begin{aligned} \langle+\hat{\boldsymbol{y}}| \, |+\hat{\boldsymbol{x}}\rangle &= \left(\frac{1}{\sqrt{2}} \langle+\hat{\boldsymbol{z}}| - \frac{i}{\sqrt{2}} \langle-\hat{\boldsymbol{z}}|\right)\left(\frac{1}{\sqrt{2}} |+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}} |-\hat{\boldsymbol{z}}\rangle\right) \\ &= \frac{1}{2}\Big(\langle+\hat{\boldsymbol{z}}| \, |+\hat{\boldsymbol{z}}\rangle + \langle+\hat{\boldsymbol{z}}| \, |-\hat{\boldsymbol{z}}\rangle - i \langle-\hat{\boldsymbol{z}}| \, |+\hat{\boldsymbol{z}}\rangle - i \langle-\hat{\boldsymbol{z}}| \, |-\hat{\boldsymbol{z}}\rangle\Big) \\ &= \frac{1}{2}\left(1 + 0 - 0 - i\right) \end{aligned}$$

since $\langle+\hat{\boldsymbol{z}}| \, |+\hat{\boldsymbol{z}}\rangle = \langle+\hat{\boldsymbol{z}}| + \hat{\boldsymbol{z}}\rangle = 1$, etc,.... Thus

$$|\langle+\hat{\boldsymbol{y}}| + \hat{\boldsymbol{x}}\rangle|^2 = \left(\frac{1-i}{2}\right)\left(\frac{1-i}{2}\right)^* = \frac{1}{2}.$$

Another useful application of the relationship between bra and ket vectors is

$$\begin{aligned}\left(\langle\Phi|\Psi\rangle\right)^* = \left(\langle\Phi|\Psi\rangle\right)^\dagger &= \left(\langle\Phi|\,|\Psi\rangle\right)^\dagger \\ &= |\Psi\rangle^\dagger\,\langle\Phi|^\dagger \\ &= \langle\Psi|\,|\Phi\rangle \\ &= \langle\Psi|\Phi\rangle\,.\end{aligned}$$

Thus

$$\boxed{\langle\Psi|\Phi\rangle = \left(\langle\Phi|\Psi\rangle\right)^*} \tag{II.3.50}$$

A more concrete method of using bra and ket vectors involves using their representations in terms of row and column vectors

Example: A spin-1/2 particle in the state $|+\hat{\boldsymbol{x}}\rangle$ is subjected to an SG $\hat{\boldsymbol{y}}$ measurement. Determine the probability with which the measurement outcome is $S_y = -\hbar/2$.

Answer: Again

$$\Pr\left(S_y = -\hbar/2\right) = |\langle -\hat{\boldsymbol{y}}| + \hat{\boldsymbol{x}}\rangle|^2.$$

Then

$$|+\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$$

$$|-\hat{\boldsymbol{y}}\rangle = \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle - \frac{i}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix}\frac{1}{\sqrt{2}}\\ -\frac{i}{\sqrt{2}}\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$$

Then we need to determine the result of the operation of the bra $\langle -\hat{\boldsymbol{y}}|$ on the ket $|+\hat{\boldsymbol{x}}\rangle$. The bra is represented via:

$$\begin{aligned}\langle -\hat{\boldsymbol{y}}| &= |-\hat{\boldsymbol{y}}\rangle^\dagger \\ &\leftrightarrow \left(\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}\right)^\dagger \\ &= \left(\frac{1}{\sqrt{2}}\right)^* \begin{pmatrix}1\\i\end{pmatrix}^T \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix}1 & i\end{pmatrix}.\end{aligned}$$

Thus

$$\langle -\hat{\boldsymbol{y}}| + \hat{\boldsymbol{x}}\rangle = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{2}(1+i).$$

This leaves

$$\Pr\left(S_y = -\hbar/2\right) = |\langle -\hat{\boldsymbol{y}}| + \hat{\boldsymbol{x}}\rangle|^2 = \frac{1}{2}.$$

3.3 Measurement operators

Given that a spin-1/2 particle in the state $|\Psi\rangle$ is subjected to an SG $\hat{\boldsymbol{n}}$ measurement, the probability of the outcome $S_n = +\hbar/2$ is

$$\begin{aligned} \Pr\left(S_n = +\hbar/2\right) &= |\langle +\hat{\boldsymbol{n}}|\Psi\rangle|^2 \\ &= \left(\langle +\hat{\boldsymbol{n}}|\Psi\rangle\right)^* \langle +\hat{\boldsymbol{n}}|\Psi\rangle . \end{aligned}$$

Now, according to Eq. (II.3.50), this gives

$$\Pr\left(S_n = +\hbar/2\right) = \langle \Psi| + \hat{\boldsymbol{n}}\rangle \langle +\hat{\boldsymbol{n}}|\Psi\rangle .$$

The next step is to formally separate the parts that pertain to the state from those pertaining to the measurement outcome. Thus

$$\Pr\left(S_n = +\hbar/2\right) = \langle \Psi| \left[|+\hat{\boldsymbol{n}}\rangle \langle +\hat{\boldsymbol{n}}|\right] |\Psi\rangle \tag{II.3.51}$$

Evidently the entity $|+\hat{\boldsymbol{n}}\rangle \langle +\hat{\boldsymbol{n}}|$ pertains just to the measurement type and at least the state related to the outcome $S_n = +\hbar/2$, while the remaining terms in Eq. (II.3.51) refer to the state of the particle entering the measuring device. The issue now is to assign a mathematical meaning to this entity. The straightforward way to do so is to consider it as an operator that maps kets onto kets. Diagrammatically this is shown in Fig. II.3.5.

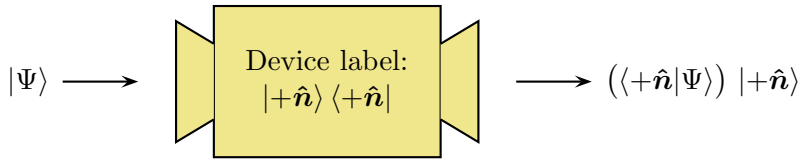


Figure II.3.5: An operator of the form $|+\hat{\boldsymbol{n}}\rangle \langle +\hat{\boldsymbol{n}}|$ can be regarded as a device which takes a ket as an input, processes it and produces a ket as an output.

The rule according to which $|+\hat{\boldsymbol{n}}\rangle \langle +\hat{\boldsymbol{n}}|$ operates is

$$\boxed{\left[|+\hat{\boldsymbol{n}}\rangle \langle +\hat{\boldsymbol{n}}|\right] |\Psi\rangle := \left(\langle +\hat{\boldsymbol{n}}|\Psi\rangle\right) |+\hat{\boldsymbol{n}}\rangle .} \tag{II.3.52}$$

The term $\langle +\hat{\boldsymbol{n}}|\Psi\rangle$ on the right hand side is a complex number and hence the entire right hand side is a ket. This can be remembered via the heuristic argument:

$$\begin{aligned}\left[|+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}|\right]|\Psi\rangle &= |+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}|\Psi\rangle \\ &= \left(\langle+\hat{\boldsymbol{n}}|\Psi\rangle\right)|+\hat{\boldsymbol{n}}\rangle\,.\end{aligned}$$

The operator $|+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}|$ is denoted

$$\boxed{\hat{P}_{+n} := |+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}|} \tag{II.3.53}$$

where the "hat" indicates an operator. Similarly there is an operator associated with $S_n = -\hbar/2$ and is denoted

$$\boxed{\hat{P}_{-n} := |-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}|\,.} \tag{II.3.54}$$

These are called the projection operators associated with the measurement of S_n.

In order to actually compute the action of such operators on kets, it is helpful to use the fact that they are linear. Thus for any kets $|\Psi_1\rangle, |\Psi_2\rangle$ and complex numbers α, β,

$$\boxed{\hat{P}_{\pm n}\Big(\alpha\,|\Psi_1\rangle + \beta\,|\Psi_2\rangle\Big) = \alpha\hat{P}_{\pm n}\,|\Psi_1\rangle + \beta\hat{P}_{\pm n}\,|\Psi_2\rangle} \tag{II.3.55}$$

Exercise: Prove the result of Eq. (II.3.55)

The linearity of the operators $\hat{P}_{\pm n}$ means that the action of the operator can be reduced ultimately to operations on basic vectors as the example that follows indicates. Additionally one can picture this action diagrammatically via Fig. II.3.6.

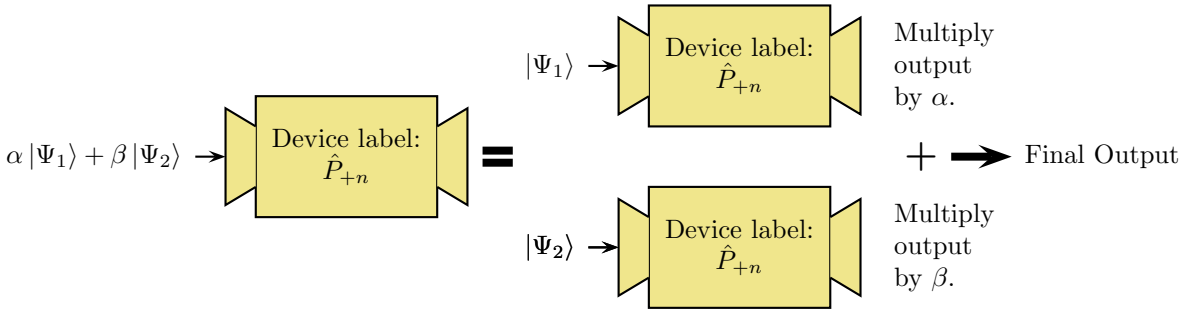


Figure II.3.6: A schematic illustrating the linearity of the operator $\hat{P}_{+n}$.

Example: Determine the action of $\hat{P}_{+z}$ acting on $|\Psi\rangle = |+\hat{\boldsymbol{y}}\rangle$.

Answer: First

$$|+\hat{\boldsymbol{y}}\rangle = \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle + \frac{i}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle\,.$$

Thus

$$\begin{aligned}\hat{P}_{+z}\left|+\hat{\boldsymbol{y}}\right\rangle &= \hat{P}_{+z}\left(\frac{1}{\sqrt{2}}\left|+\hat{\boldsymbol{z}}\right\rangle + \frac{i}{\sqrt{2}}\left|-\hat{\boldsymbol{z}}\right\rangle\right)\\ &= \frac{1}{\sqrt{2}}\hat{P}_{+z}\left|+\hat{\boldsymbol{z}}\right\rangle + \frac{i}{\sqrt{2}}\hat{P}_{+z}\left|-\hat{\boldsymbol{z}}\right\rangle\\ &= \frac{1}{\sqrt{2}}\left|+\hat{\boldsymbol{z}}\right\rangle\left\langle+\hat{\boldsymbol{z}}\right|+\hat{\boldsymbol{z}}\right\rangle + \frac{i}{\sqrt{2}}\left|+\hat{\boldsymbol{z}}\right\rangle\left\langle+\hat{\boldsymbol{z}}\right|-\hat{\boldsymbol{z}}\right\rangle\\ &= \frac{1}{\sqrt{2}}\left\langle+\hat{\boldsymbol{z}}\right|+\hat{\boldsymbol{z}}\right\rangle\left|+\hat{\boldsymbol{z}}\right\rangle + \frac{i}{\sqrt{2}}\left\langle+\hat{\boldsymbol{z}}\right|-\hat{\boldsymbol{z}}\right\rangle\left|+\hat{\boldsymbol{z}}\right\rangle\\ &= \frac{1}{\sqrt{2}}\left|+\hat{\boldsymbol{z}}\right\rangle.\end{aligned}$$

Additionally one can show that if $\left|+\hat{\boldsymbol{n}}\right\rangle = c_+\left|+\hat{\boldsymbol{z}}\right\rangle + c_-\left|-\hat{\boldsymbol{z}}\right\rangle$ then

$$\begin{aligned}\left|+\hat{\boldsymbol{n}}\right\rangle\left\langle+\hat{\boldsymbol{n}}\right| &= \left(c_+\left|+\hat{\boldsymbol{z}}\right\rangle + c_-\left|-\hat{\boldsymbol{z}}\right\rangle\right)\left(c_+\left|+\hat{\boldsymbol{z}}\right\rangle + c_-\left|-\hat{\boldsymbol{z}}\right\rangle\right)^\dagger\\ &= \left(c_+\left|+\hat{\boldsymbol{z}}\right\rangle + c_-\left|-\hat{\boldsymbol{z}}\right\rangle\right)\left(c_+^*\left\langle+\hat{\boldsymbol{z}}\right| + c_-\left\langle-\hat{\boldsymbol{z}}\right|\right)\\ &= |c_+|^2\left|+\hat{\boldsymbol{z}}\right\rangle\left\langle+\hat{\boldsymbol{z}}\right| + c_+c_-^*\left|+\hat{\boldsymbol{z}}\right\rangle\left\langle-\hat{\boldsymbol{z}}\right| + c_+^*c_-\left|-\hat{\boldsymbol{z}}\right\rangle\left\langle+\hat{\boldsymbol{z}}\right| + |c_-|^2\left|-\hat{\boldsymbol{z}}\right\rangle\left\langle-\hat{\boldsymbol{z}}\right|.\end{aligned} \tag{II.3.56}$$

Results of this type are conveniently illustrated by representing kets and bras in terms of row and column vectors.

Example: Represent $\hat{P}_{+x}$ as a matrix in the $\{\left|+\hat{\boldsymbol{z}}\right\rangle, \left|-\hat{\boldsymbol{z}}\right\rangle\}$ basis and determine its effect on $\left|+\hat{\boldsymbol{z}}\right\rangle$.

Answer: Here

$$\hat{P}_{+x} := \left|+\hat{\boldsymbol{x}}\right\rangle\left\langle+\hat{\boldsymbol{x}}\right|.$$

Then in the $\{\left|+\hat{\boldsymbol{z}}\right\rangle, \left|-\hat{\boldsymbol{z}}\right\rangle\}$ basis

$$\left|+\hat{\boldsymbol{x}}\right\rangle = \frac{1}{\sqrt{2}}\left|+\hat{\boldsymbol{z}}\right\rangle + \frac{1}{\sqrt{2}}\left|-\hat{\boldsymbol{z}}\right\rangle \leftrightarrow \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$$

and similarly

$$\begin{aligned}\left\langle+\hat{\boldsymbol{x}}\right| &\leftrightarrow \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}^\dagger\\ &= \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}.\end{aligned}$$

This gives

$$\begin{aligned}\hat{P}_{+x} &= |+\hat{\boldsymbol{x}}\rangle\langle+\hat{\boldsymbol{x}}| \\ &\leftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\end{aligned}$$

where the last step was attained using the usual matrix multiplication rules. Thus the projection operators are represented by matrices. Now the action of this on $|+\hat{\boldsymbol{z}}\rangle$ is determine by noting that

$$|+\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix}1\\0\end{pmatrix},$$

giving

$$\begin{aligned}\hat{P}_{+x}|+\hat{\boldsymbol{z}}\rangle &\leftrightarrow \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}\end{aligned}$$

which is the same as $\frac{1}{2}|+\hat{\boldsymbol{x}}\rangle$. Thus, by working with representations in terms of matrices and vectors, we arrive at

$$\hat{P}_{+x}|+\hat{\boldsymbol{z}}\rangle = \frac{1}{2}|+\hat{\boldsymbol{x}}\rangle.$$

Exercise: Show that $\hat{P}_{+x}|-\hat{\boldsymbol{z}}\rangle = \frac{1}{2}|-\hat{\boldsymbol{x}}\rangle$.

In general, for spin-1/2 particles the projection operators can be represented by 2×2 matrices. Explicitly these are constructed by multiplying a column vector with a row vector as the example above illustrates. These techniques can also be used to determine expectation values. The general rule is

For a particle in the state described by $|\Psi\rangle$ and subjected to SG $\hat{\boldsymbol{n}}$, the probabilities of the two measurement outcomes are

$$\begin{aligned}\Pr(S_n = +\hbar/2) &= \langle\Psi|\hat{P}_{+n}|\Psi\rangle \\ \Pr(S_n = -\hbar/2) &= \langle\Psi|\hat{P}_{-n}|\Psi\rangle.\end{aligned}$$

(II.3.57)

Example: An ensemble of spin-1/2 particles are all in the state $|+\hat{\boldsymbol{z}}\rangle$ and each is subjected to an SG $\hat{\boldsymbol{x}}$ measurement. Determine $\Pr(S_x = +\hbar/2)$.

Answer: Applying the general rule,

$$\begin{aligned}\Pr\left(S_x = +\hbar/2\right) &= \langle\Psi|\,\hat{P}_{+x}\,|\Psi\rangle \\ &= \begin{pmatrix}1 & 0\end{pmatrix}\frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix} \\ &= \begin{pmatrix}1 & 0\end{pmatrix}\frac{1}{2}\begin{pmatrix}1\\ 1\end{pmatrix} \\ &= \frac{1}{2}.\end{aligned}$$

The projection operators associated with a measurement are useful for computing expectation values. Applying the general rule for computing the mean of a distribution, Eq. (II.2.10), gives for an ensemble of particles each in state $|\Psi\rangle$

$$\begin{aligned}\langle S_n\rangle &= \frac{\hbar}{2}\Pr\left(S_n = +\hbar/2\right) + \left(\frac{\hbar}{2}\right)\Pr\left(S_n = -\hbar/2\right) \\ &= \frac{\hbar}{2}\left[\langle\Psi|\,\hat{P}_{+n}\,|\Psi\rangle - \langle\Psi|\,\hat{P}_{-n}\,|\Psi\rangle\right] \\ &= \langle\Psi|\left[\frac{\hbar}{2}\,\hat{P}_{+n} - \frac{\hbar}{2}\,\hat{P}_{-n}\right]|\Psi\rangle \qquad \text{(II.3.58)}\end{aligned}$$

which leads to the definition of *the operator associated with the measurement of the* $\hat{\boldsymbol{n}}$ *component of the spin,*

$$\boxed{\hat{S}_n := \frac{\hbar}{2}\,\hat{P}_{+n} - \frac{\hbar}{2}\,\hat{P}_{-n}.} \qquad \text{(II.3.59)}$$

Alternatively

$$\boxed{\hat{S}_n := \frac{\hbar}{2}\,|+\hat{\boldsymbol{n}}\rangle\,\langle+\hat{\boldsymbol{n}}| - \frac{\hbar}{2}\,|-\hat{\boldsymbol{n}}\rangle\,\langle-\hat{\boldsymbol{n}}|\,.} \qquad \text{(II.3.60)}$$

This spin operator patently includes all the information about the measurement. The two outcomes, $S_n = \pm\hbar/2$ and the two associated states, $|+\hat{\boldsymbol{n}}\rangle$ and $|-\hat{\boldsymbol{n}}\rangle$ are clearly present in the expression of Eq. (II.3.60). On important use of such operators is described by:

> For an ensemble of particles, each in the state described by $|\Psi\rangle$ and subjected to SG $\hat{\boldsymbol{n}}$, the expectation value of the outcomes is
>
> $$\langle S_n\rangle = \langle\Psi|\,\hat{S}_n\,|\Psi\rangle\,.$$ (II.3.61)

It is usually convenient to work with spin operators by representing them as matrices.

Example: Find the matrix representation of $\hat{S}_z$ in the $\{|+\hat{\boldsymbol{z}}\rangle\,, |-\hat{\boldsymbol{z}}\rangle\}$ basis.

Answer: According to Eq. (II.3.60):

$$\hat{S}_z := \frac{\hbar}{2} \left|+\hat{\boldsymbol{z}}\right\rangle \left\langle+\hat{\boldsymbol{z}}\right| - \frac{\hbar}{2} \left|-\hat{\boldsymbol{z}}\right\rangle \left\langle-\hat{\boldsymbol{z}}\right|$$

The representation is arrived at conveniently by representing the bra and ket vectors. To this end,

$$\left|+\hat{\boldsymbol{z}}\right\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and}$$
$$\left|-\hat{\boldsymbol{z}}\right\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Similarly

$$\left\langle+\hat{\boldsymbol{z}}\right| \leftrightarrow (1 \;\; 0) \quad \text{and}$$
$$\left\langle-\hat{\boldsymbol{z}}\right| \leftrightarrow (0 \;\; 1).$$

Thus

$$\left|+\hat{\boldsymbol{z}}\right\rangle \left\langle+\hat{\boldsymbol{z}}\right| \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \;\; 0)$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$\left|-\hat{\boldsymbol{z}}\right\rangle \left\langle-\hat{\boldsymbol{z}}\right| \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \;\; 1)$$
$$= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

These give

$$\hat{S}_z \leftrightarrow \frac{\hbar}{2} \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right]$$
$$= \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In a similar fashion it can be shown that

The operators corresponding to measurements of the components of spin along $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$ are represented in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis by

$$\hat{S}_x \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\hat{S}_y \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{II.3.62}$$

$$\hat{S}_z \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Exercise: Prove that

$$\hat{S}_x \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \hat{S}_y \leftrightarrow \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

A remarkable fact is that the component of the operator associated with the $\hat{\boldsymbol{n}}$ component of the spin can be expressed as a simple combination of the operators associated with the $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{z}}$ components of spin. This is the content of

The operator corresponding to measurement of the component of spin along $\hat{\boldsymbol{n}}$ where using standard spherical coordinates

$$\hat{\boldsymbol{n}} = \sin\theta\,\cos\phi\,\hat{\boldsymbol{x}} + \sin\theta\,\sin\phi\,\hat{\boldsymbol{y}} + \cos\theta\,\hat{\boldsymbol{z}} \tag{II.3.63}$$

is

$$\hat{S}_n = \sin\theta\,\cos\phi\,\hat{S}_x + \sin\theta\,\sin\phi\,\hat{S}_y + \cos\theta\,\hat{S}_z.$$

Exercise: Prove Eq. (II.3.63).

The spin operators are frequently used to determine expectation values of measurements of spin components and this can be done conveniently using representations of the entities in terms of a basis.

Example: An ensemble of spin-1/2 particles, with each prepared in the state $|+\hat{\boldsymbol{n}}\rangle$ is subjected to an SG $\hat{\boldsymbol{z}}$ measurement. Determine the expectation value of the measurement outcome.

Answer: According to Eq. (II.3.61),

$$\langle S_z \rangle = \langle -\hat{\boldsymbol{n}}| \hat{S}_z |+\hat{\boldsymbol{n}}\rangle .$$

The representations in terms of the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ can be reached using

$$\begin{aligned} |+\hat{\boldsymbol{n}}\rangle &= \cos\left(\frac{\theta}{2}\right) |+\hat{\boldsymbol{z}}\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |-\hat{\boldsymbol{z}}\rangle \\ &\leftrightarrow \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi} \sin(\theta/2) \end{pmatrix}. \end{aligned}$$

and thus

$$\begin{aligned} \langle +\hat{\boldsymbol{n}}| &\leftrightarrow \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi} \sin(\theta/2) \end{pmatrix}^\dagger \\ &= \begin{pmatrix} \cos(\theta/2) & e^{-i\phi} \sin(\theta/2) \end{pmatrix}. \end{aligned}$$

This gives

$$\begin{aligned} \langle S_z \rangle &= \langle -\hat{\boldsymbol{n}}| \hat{S}_z |+\hat{\boldsymbol{n}}\rangle \\ &= \begin{pmatrix} \cos(\theta/2) & e^{-i\phi} \sin(\theta/2) \end{pmatrix} \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi} \sin(\theta/2) \end{pmatrix} \\ &= \frac{\hbar}{2} \begin{pmatrix} \cos(\theta/2) & e^{-i\phi} \sin(\theta/2) \end{pmatrix} \begin{pmatrix} \cos(\theta/2) \\ -e^{i\phi} \sin(\theta/2) \end{pmatrix} \\ &= \frac{\hbar}{2} \left[\cos^2(\theta/2) - \sin^2(\theta/2)\right] \\ &= \frac{\hbar}{2} \cos\theta. \end{aligned}$$

The spin operators have an important property, namely that they are *Hermitian,* whose meaning is given by the following.

$$\boxed{\text{An operator } \hat{A} \text{ is Hermitian} \Leftrightarrow \hat{A}^\dagger = \hat{A}.} \tag{II.3.64}$$

Thus

$$\boxed{\hat{S}_n^\dagger = \hat{S}_n} \tag{II.3.65}$$

for any direction $\hat{\boldsymbol{n}}$. This stems from a similar result for the operators $\hat{P}_{\pm n}$, which is a requirement for these operators to produce probability values that are real.

Exercise: Prove Eq. (II.3.65).

Example: Show that $\hat{S}_y$ is Hermitian.

Answer: For $\hat{S}_y$,

$$\begin{aligned}\hat{S}_y^\dagger &\leftrightarrow \left[\frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\right]^\dagger \\ &= \frac{\hbar^*}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}^\dagger \\ &= \frac{\hbar}{2}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} \\ &\leftrightarrow \hat{S}_y.\end{aligned}$$

This shows that $\hat{S}_y^\dagger = \hat{S}_y$.

Hermitian operators play an important role in the description of measurements in quantum mechanics and are called *observables*. We have shown that the operator associated with measurements of any component of spin is Hermitian, and hence is an observable. It remains to demonstrate that any observable corresponds to some measurement.

Outcome	Associated State
λ_1	$\|\phi_1\rangle$
λ_2	$\|\phi_2\rangle$

Table II.3.1: Table of measurement outcomes and states for measurement of A.

Consider a physically measurable quantity, denoted A. Suppose that the measurement outcomes and associated states are given by Table 3.3. The meaning of the associated state $|\phi_i\rangle$ is that, for a particle in this state, measurement of A yields λ_i with certainty (for each of $i = 1, 2$. The operator corresponding to this measurement would be constructed in the same way as that for any component of spin. Thus

$$\hat{A} = \lambda_1 |\phi_1\rangle \langle\phi_1| + \lambda_2 |\phi_2\rangle \langle\phi_2| \tag{II.3.66}$$

is the observable associated with A.

Exercise: Show that $\hat{A}$ as defined in Eq. (II.3.66) is Hermitian and show that the expectation value of A for an ensemble of particles, each in the state $|\Psi\rangle$ is $\langle A\rangle = \langle\Psi|\hat{A}|\Psi\rangle$.

Now consider an arbitrary observable. Is it possible to determine the measurement outcomes and associated states and thereby reconstruct Table 3.3? Consider first measurements of S_z.

Example: Consider measurements of S_z. Show that

$$\hat{S}_z \left|+\hat{\boldsymbol{z}}\right\rangle = \lambda_1 \left|+\hat{\boldsymbol{z}}\right\rangle$$
$$\hat{S}_z \left|-\hat{\boldsymbol{z}}\right\rangle = \lambda_2 \left|-\hat{\boldsymbol{z}}\right\rangle$$

where λ_i are two constants. Find their values.

Answer: For measurements of S_z,

$$\hat{S}_z \left|+\hat{\boldsymbol{z}}\right\rangle \leftrightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$= \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Thus

$$\hat{S}_z \left|+\hat{\boldsymbol{z}}\right\rangle = \frac{\hbar}{2} \left|+\hat{\boldsymbol{z}}\right\rangle.$$

Similarly

$$\hat{S}_z \left|-\hat{\boldsymbol{z}}\right\rangle = -\frac{\hbar}{2} \left|-\hat{\boldsymbol{z}}\right\rangle.$$

It follows that $\lambda_1 = +\hbar/2$ and $\lambda_2 = -\hbar/2$.

In the previous example each of the states that are associated with measurement outcomes satisfy an *eigenvalue equation*

$$\hat{S}_z \left|\phi\right\rangle = \lambda \left|\phi\right\rangle$$

where λ is a number. The important feature of this equation is that the same state appears on both sides. In general this is not true of an arbitrary state. For example,

$$\hat{S}_z \left|+\hat{\boldsymbol{x}}\right\rangle \leftrightarrow \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$
$$= \frac{\hbar}{2} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and thus

$$\hat{S}_z \left|+\hat{\boldsymbol{x}}\right\rangle = \frac{\hbar}{2} \left|-\hat{\boldsymbol{x}}\right\rangle.$$

Distinct states appear on the two sides of this and $\left|+\hat{\boldsymbol{x}}\right\rangle$ does not satisfy an eigenvalue equation for $\hat{S}_z$.

Exercise: Show that

$$\hat{S}_x \left|+\hat{\boldsymbol{x}}\right\rangle = \frac{\hbar}{2} \left|+\hat{\boldsymbol{x}}\right\rangle$$

and

$$\hat{S}_x \left|-\hat{\boldsymbol{x}}\right\rangle = -\frac{\hbar}{2} \left|-\hat{\boldsymbol{x}}\right\rangle.$$

The previous examples indicate a possible general relationship between observables and the associated measurement outcomes and states. In both cases the state associated with the measurement outcome λ satisfies an eigenvalue equation

$$\boxed{\hat{A}\,|\phi\rangle = \lambda\,|\phi\rangle} \tag{II.3.67}$$

where λ is called an *eigenvalue* and $|\phi\rangle$ and *eigenstate*. There may be more than one eigenstate associated with a given observable $\hat{A}$, and for each eigenvalue there is an associated eigenstate. In the cases that we have considered, the method for reconstructing Table 3.3 is to find the eigenvalues and associated eigenstates of the observable; the eigenvalues are the possible measurement outcomes and the eigenstates are the associated states.

Example: Find the eigenvalues and eigenstates of

$$\hat{A} \leftrightarrow \frac{\hbar}{2}\frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}.$$

Answer: It is easy to show that $\hat{A}^\dagger = \hat{A}$. The eigenvalue equation is

$$\hat{A}\,|\phi\rangle = \lambda\,|\phi\rangle$$

and a general result from linear algebra asserts that the eigenvalues satisfy

$$\det\left(\hat{A} - \lambda\hat{I}\right) = 0$$

where "det" signifies the determinant and

$$\hat{I} \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is the *identity* operator. For 2×2 matrices,

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc.$$

Here

$$\begin{aligned}
\det\left(\hat{A} - \lambda\hat{I}\right) = 0 &\Longleftrightarrow \det\left[\frac{\hbar}{2}\frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = 0 \\
&\Longleftrightarrow \det\begin{pmatrix} \frac{4\hbar}{10} - \lambda & \frac{3\hbar}{10} \\ \frac{3\hbar}{10} & -\frac{4\hbar}{10} - \lambda \end{pmatrix} = 0 \\
&\Longleftrightarrow \left(\frac{4\hbar}{10} - \lambda\right)\left(-\frac{4\hbar}{10} - \lambda\right) - \left(\frac{3\hbar}{10}\right)^2 = 0.
\end{aligned}$$

Answer: Thus the eigenvalues satisfy

$$\lambda^2 - \left(\frac{4\hbar}{10}\right)^2 - \left(\frac{3\hbar}{10}\right)^2 = 0,$$

from which it follows that

$$\lambda^2 - \frac{\hbar^2}{4} = 0.$$

Thus the two eigenvalues are

$$\lambda_1 = +\frac{\hbar}{2} \quad \text{and}$$
$$\lambda_2 = -\frac{\hbar}{2}.$$

Now consider the process of finding the eigenstates. First for the eigenvalue $\lambda_1 = +\hbar/2$, the eigenstate $|\phi_1\rangle$ satisfies

$$\hat{A}\,|\phi_1\rangle = \frac{\hbar}{2}\,|\phi_1\rangle\,.$$

The standard approach is to represent $|\phi_1\rangle$ in terms of the $\{|+\hat{\boldsymbol{z}}\rangle\,, |-\hat{\boldsymbol{z}}\rangle\}$ basis,

$$|\phi_1\rangle = c_+\,|+\hat{\boldsymbol{z}}\rangle + c_-\,|-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} c_+ \\ c_- \end{pmatrix}$$

where $c_\pm$ are complex numbers that need to still be determined. Thus the eigenvalue equation becomes

$$\frac{\hbar}{2}\frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} c_+ \\ c_- \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} c_+ \\ c_- \end{pmatrix}$$

which simplifies to

$$\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} c_+ \\ c_- \end{pmatrix} = 5\begin{pmatrix} c_+ \\ c_- \end{pmatrix}.$$

This leaves the pair of equations

$$4c_+ + 3c_- = 5c_+$$
$$3c_+ - 4c_- = 5c_-$$

which simplify to two equivalent equations:

$$-c_+ + 3c_- = 0$$
$$3c_+ - 9c_- = 0$$

Taking the first gives

$$c_+ = 3c_-$$

and thus the eigenstate is

$$|\phi_1\rangle = c_-\big(3\,|+\hat{\boldsymbol{z}}\rangle + |-\hat{\boldsymbol{z}}\rangle\big).$$

The constant c_- can be fixed by requiring that $\langle\phi|\phi\rangle = 1$. This leaves

$$|c_-| = \frac{1}{\sqrt{10}}$$

and the most convenient choice is

$$c_- = \frac{1}{\sqrt{10}}.$$

Thus the eigenstate corresponding to $\lambda_1 = +\hbar/2$ is

$$|\phi_1\rangle = \frac{3}{\sqrt{10}}\,|+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{10}}\,|-\hat{\boldsymbol{z}}\rangle\,.$$

A similar derivation leads to the fact that the eigenstate corresponding to $\lambda_2 = -\hbar/2$ is

$$|\phi_2\rangle = \frac{1}{\sqrt{10}}\,|+\hat{\boldsymbol{z}}\rangle - \frac{3}{\sqrt{10}}\,|-\hat{\boldsymbol{z}}\rangle\,.$$

The connection between the eigenvalues and measurement outcomes is enabled via the following theorems.

Theorem: The eigenvalues of an Hermitian operator are real.

Proof: The eigenvalue equation is

$$\hat{A}\,|\phi\rangle = \lambda\,|\phi\rangle$$

and thus

$$\langle\phi|\,\hat{A}\,|\phi\rangle = \lambda\,\langle\phi|\phi\rangle\,.$$

Then

$$\begin{aligned}\Big(\langle\phi|\,\hat{A}\,|\phi\rangle\Big)^\dagger &= \Big(\lambda\,\langle\phi|\phi\rangle\Big)^\dagger \\ &= \lambda^*\Big(\langle\phi|\phi\rangle\Big)^\dagger \\ &= \lambda^*\big(\langle\phi|\phi\rangle\big)^* \\ &= \lambda^*\,\langle\phi|\phi\rangle\end{aligned}$$

since $\langle\phi|\phi\rangle$ is real. Returning to the left hand side,

$$\begin{aligned}\left(\langle\phi|\,\hat{A}\,|\phi\rangle\right)^{\dagger} &= |\phi\rangle^{\dagger}\,\hat{A}^{\dagger}\,\langle\phi|^{\dagger} \\ &= \langle\phi|\,\hat{A}\,|\phi\rangle \\ &= \lambda\,\langle\phi|\phi\rangle\,.\end{aligned}$$

Equating the sides gives $\lambda^* = \lambda$ and thus the eigenvalue is real. •

Theorem: If λ_i are distinct eigenvalues of an Hermitian operator then the corresponding eigenstates are orthogonal.

Proof: Suppose that

$$\begin{aligned}\hat{A}\,|\phi_1\rangle &= \lambda_1\,|\phi_1\rangle \\ \hat{A}\,|\phi_2\rangle &= \lambda_2\,|\phi_2\rangle\,.\end{aligned}$$

Then

$$\begin{aligned}\langle\phi_1|\,\hat{A}\,|\phi_2\rangle &= \langle\phi_1|\,\lambda_1\,|\phi_2\rangle \\ &= \lambda_1\,\langle\phi_1|\phi_2\rangle\end{aligned}$$

and similarly

$$\begin{aligned}\langle\phi_2|\,\hat{A}\,|\phi_1\rangle &= \langle\phi_2|\,\lambda_2\,|\phi_1\rangle \\ &= \lambda_2\,\langle\phi_2|\phi_1\rangle\,.\end{aligned}$$

However,

$$\begin{aligned}\left(\langle\phi_1|\,\hat{A}\,|\phi_2\rangle\right)^{\dagger} &= \langle\phi_2|\,\hat{A}^{\dagger}\,|\phi_1\rangle \\ &= \langle\phi_2|\,\hat{A}\,|\phi_1\rangle\end{aligned}$$

since $\hat{A}^{\dagger} = \hat{A}$. This implies that

$$\left(\lambda_1\,\langle\phi_1|\phi_2\rangle\right)^{\dagger} = \lambda_2\,\langle\phi_2|\phi_1\rangle$$

which is equivalent to

$$\lambda_1^*\,\langle\phi_2|\phi_1\rangle = \lambda_2\,\langle\phi_2|\phi_1\rangle$$

and since eigenvalues of Hermitian operators are real,

$$\lambda_1\,\langle\phi_2|\phi_1\rangle = \lambda_2\,\langle\phi_2|\phi_1\rangle\,.$$

The only way that this can be satisfied when $\lambda_1 \neq \lambda_2$ is for $\langle\phi_2|\phi_1\rangle = 0$. Thus the states corresponding to distinct eigenvalues are orthogonal. •

In the most straightforward cases, the eigenvalues of $\hat{A}$ are all distinct. For a spin-1/2 system there can at most be two of these (since the vectors are orthogonal and there can at most be two orthogonal vectors). It is possible to express $\hat{A}$ in terms of the eigenstates in exactly the form of Eq. (II.3.66). Thus the eigenvalues are the measurement outcomes and the eigenstates are the associated measurement states.

If $\hat{A}$ is an observable with eigenvalues and eigenstates given by

$$\hat{A}\,|\phi_i\rangle = \lambda_i\,|\phi_i\rangle$$

then the outcomes and states associated with the measurement of A are

Outcome	Associated State
λ_1	$\|\phi_1\rangle$
λ_2	$\|\phi_2\rangle$
$\vdots$	$\vdots$

For a spin-1/2 system a schematic version of the measurement process is provided by Fig. II.3.7. Measurement of the spin component, S_n, is clearly an example of this.

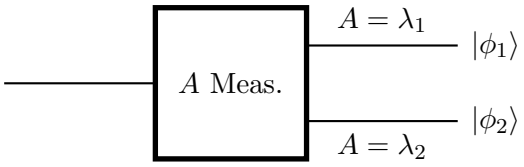


Figure II.3.7: Schematic diagram of a measurement corresponding to the observable $\hat{A}$. The eigenvalues of $\hat{A}$ are λ_1 and λ_2 with corresponding eigenstates $|\phi_1\rangle$ and $|\phi_2\rangle$.

Exercise: Let $|\phi_i\rangle$ be an eigenstate of an observable $\hat{A}$ corresponding to eigenvalue λ_i. Show that $\langle A\rangle = \lambda_i$ and that $\Delta A = 0$. Thus given a single particle in the state $|\phi_i\rangle$, subjected to a measurement of A, which outcomes can occur and what are the probabilities with which they occur?

Even for spin-1/2 systems, there are measurable physical quantities other than spin components. The most important is the energy of the system. The route to the corresponding observable can be ascertained by considering the relationship between the energy and spin of a classical charged dipole. When placed in a magnetic field $\mathbf{B}$ the dipole has potential energy

$$U = -\boldsymbol{\mu} \cdot \mathbf{B} \tag{II.3.68}$$

and as this is the only contribution to the total energy, the energy is

$$E = -\boldsymbol{\mu} \cdot \mathbf{B}. \tag{II.3.69}$$

In the absence of an external magnetic field it is meaningless to speak of energy. The route to an observable for energy in the quantum mechanical description will be to relate the magnetic dipole moment to the particle spin and to apply results for quantum mechanical measurements of spin to the problem. For a classical particle of mass m, charge q and g-factor g, the dipole moment is related to the spin via

$$\boldsymbol{\mu} = \frac{gq}{2m}\,\mathbf{S} \tag{II.3.70}$$

where $\mathbf{S}$ is the classical spin of the particle. Thus

$$E = -\frac{gq}{2m}\,\mathbf{B} \cdot \mathbf{S}. \tag{II.3.71}$$

Suppose that the magnetic field is oriented along the $\hat{\boldsymbol{z}}$ direction. Then $\mathbf{B} = B_z\hat{\boldsymbol{z}}$ and

$$E = -\frac{gq}{2m}\,B_z S_z. \tag{II.3.72}$$

It follows that given a value for the component of the spin, the energy of the particle can be computed via Eq. (II.3.72). Thus for a particle in the state $|+\hat{\boldsymbol{z}}\rangle$ a measurement of the z component of spin gives $S_z = +\hbar/2$ with certainty. It follows that the energy of a particle in this state must be

$$E = -\frac{gq}{2m}\,B_z\frac{\hbar}{2}$$

and an analogous inference applies to a particle in the state $|-\hat{\boldsymbol{z}}\rangle$. These are the only two measurement outcomes possible and the measurement details are listed in Table II.3.2. The corresponding observable, called the *Hamiltonian*, is constructed via the usual scheme and is

$$\begin{aligned}
\hat{H} &= -\frac{gq}{2m}\,B_z\frac{\hbar}{2}\,|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + \frac{gq}{2m}\,B_z\frac{\hbar}{2}\,|-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| \\
&= -\frac{gq}{2m}\,B_z\frac{\hbar}{2}\Big[|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| - |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\Big] \\
&= -\frac{gq}{2m}\,B_z\hat{S}_z.
\end{aligned}$$

Similarly if a spin-1/2 particle of mass m, charge q and g-factor g, placed in an arbitrary magnetic field, $\mathbf{B} = B\hat{\boldsymbol{n}} = B_x\hat{\boldsymbol{x}} + B_y\hat{\boldsymbol{y}} + B_z\hat{\boldsymbol{z}}$, then the corresponding Hamiltonian is

$$\hat{H} = -\frac{gq}{2m}\,B\hat{S}_n.$$

Outcome (E)	Associated State
$-\frac{gq}{2m} B_z \frac{\hbar}{2}$	$\|+\hat{\boldsymbol{z}}\rangle$
$+\frac{gq}{2m} B_z \frac{\hbar}{2}$	$\|+\hat{\boldsymbol{z}}\rangle$

Table II.3.2: Table of measurement outcomes and states for measurement of the energy of a spin-1/2 particle of mass m, charge q and g-factor g, placed in a magnetic field $\mathbf{B} = B_z\hat{\boldsymbol{z}}$.

Using the expression for $\hat{S}_n$ given by Eq. (II.3.63) this implies the following.

The Hamiltonian for a spin-1/2 particle of mass m, charge q and g-factor g, placed in the magnetic field $\mathbf{B}$ is

$$\hat{H} = -\frac{gq}{2m}\,\mathbf{B}\cdot\hat{\mathbf{S}} \qquad \text{(II.3.73)}$$

where

$$\mathbf{B}\cdot\hat{\mathbf{S}} := B_x\hat{S}_x + B_y\hat{S}_y + B_z\hat{S}_z. \qquad \text{(II.3.74)}$$

Note that the Hamiltonian can be constructed from the classical expression for energy by replacing quantum mechanical variables (spin components, in this case) with the corresponding observables. Thus

$$E = -\frac{gq}{2m}\,\mathbf{B}\cdot\mathbf{S} \quad \rightarrow \quad \hat{H} = -\frac{gq}{2m}\,\mathbf{B}\cdot\hat{\mathbf{S}}. \qquad \text{(II.3.75)}$$

3.4 Compatible and incompatible measurements

According to the above construction, if a spin-1/2 particle is placed in a magnetic field along $\hat{\boldsymbol{z}}$, then there are two states, namely $|+\hat{\boldsymbol{z}}\rangle$ and $|-\hat{\boldsymbol{z}}\rangle$, each of which gives definite outcomes for measurements of *both* S_z and E. In this respect S_z and E are *compatible measurements* (provided that the magnetic field is along $\hat{\boldsymbol{z}}$). However, there is no state which will give definite outcomes (i.e. outcomes with certainty) for measurements of *both* S_z and S_x. In this respect S_z and S_x are *incompatible measurements.* The question arises of whether it is possible to determine these facts given the associated observables. A hint at the answer to this emerges from the following theorem.

Theorem (Uncertainty Relation): For an ensemble of particles, each in the state $|\Psi\rangle$,

and any two observables $\hat{A}$ and $\hat{B}$,

$$\boxed{\Delta A\,\Delta B \geqslant \frac{1}{2}\left|\langle\Psi|\,[\hat{A},\hat{B}]\,|\Psi\rangle\right|} \tag{II.3.76}$$

where the commutator of $\hat{A}$ and $\hat{B}$ is

$$[\hat{A},\hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{II.3.77}$$

Proof: The following property of the inner product, called the *Schwarz inequality,* is crucial to this proof. Let $|\Omega\rangle$ and $|\Phi\rangle$ be any two states. Then, the Schwarz inequality states that

$$|\langle\Omega|\Phi\rangle|^2 \leqslant \langle\Omega|\Omega\rangle\,\langle\Phi|\Phi\rangle\,.$$

Now consider the states

$$|\Omega\rangle = \left(\hat{A} - \langle A\rangle\,\hat{I}\right)|\Psi\rangle$$
$$|\Phi\rangle = \left(\hat{B} - \langle B\rangle\,\hat{I}\right)|\Psi\rangle$$

Then

$$\begin{aligned}
\langle\Omega|\Omega\rangle &= \langle\Psi|\left(\hat{A} - \langle A\rangle\,\hat{I}\right)\left(\hat{A} - \langle A\rangle\,\hat{I}\right)|\Psi\rangle \\
&= \langle\Psi|\,\hat{A}^2\,|\Psi\rangle - \langle\Psi|\,\langle A\rangle\,\hat{I}\hat{A}\,|\Psi\rangle - \langle\Psi|\,\langle A\rangle\,\hat{A}\hat{I}\,|\Psi\rangle + \langle\Psi|\,\langle A\rangle^2\,\hat{I}\hat{I}\,|\Psi\rangle \\
&= \langle\Psi|\,\hat{A}^2\,|\Psi\rangle - \langle\Psi|\,\langle A\rangle\,\hat{A}\,|\Psi\rangle - \langle\Psi|\,\langle A\rangle\,\hat{A}\,|\Psi\rangle + \langle\Psi|\,\langle A\rangle^2\,\hat{I}\,|\Psi\rangle \\
&= \left\langle A^2\right\rangle - 2\,\langle A\rangle\,\langle\Psi|\,\hat{A}\,|\Psi\rangle + \langle A\rangle^2\,\langle\Psi|\,\hat{I}\,|\Psi\rangle \\
&= \left\langle A^2\right\rangle - \langle A\rangle^2 \\
&= (\Delta A)^2\,.
\end{aligned}$$

Similarly

$$\langle\Phi|\Phi\rangle = (\Delta B)^2\,.$$

Thus the Schwarz inequality gives

$$|\langle\Omega|\Phi\rangle|^2 \leqslant (\Delta A)^2\,(\Delta B)^2\,.$$

which leaves

$$|\langle\Omega|\Phi\rangle| \leqslant \Delta A\,\Delta B.$$

The left hand side of the previous equation is a complex number. For any complex number, z,

$$|z|^2 = [\mathrm{Re}(z)]^2 + [\mathrm{Im}(z)]^2 = \left[\frac{z+z^*}{2}\right]^2 + \left[\frac{z-z^*}{2i}\right]^2 \geqslant \left[\frac{z-z^*}{2i}\right]^2 = \left|\frac{z-z^*}{2i}\right|^2.$$

Thus

$$|z| \geqslant \left|\frac{z-z^*}{2i}\right| = \frac{1}{2}\,|z-z^*|$$

In this case,

$$|\langle\Omega|\Phi\rangle| \geqslant \frac{1}{2}\,|\langle\Omega|\Phi\rangle - \langle\Omega|\Phi\rangle^*| = \frac{1}{2}\,|\langle\Omega|\Phi\rangle - \langle\Phi|\Omega\rangle|\,.$$

Then

$$\begin{aligned}
\langle\Omega|\Phi\rangle &= \langle\Psi|\left(\hat{A} - \langle A\rangle\,\hat{I}\right)\left(\hat{B} - \langle B\rangle\,\hat{I}\right)|\Psi\rangle \\
&= \langle\Psi|\,\hat{A}\hat{B}\,|\Psi\rangle - \langle A\rangle\,\langle\Psi|\,\hat{I}\hat{B}\,|\Psi\rangle - \langle B\rangle\,\langle\Psi|\,\hat{A}\hat{I}\,|\Psi\rangle + \langle A\rangle\,\langle B\rangle\,\langle\Psi|\,\hat{I}\hat{I}\,|\Psi\rangle \\
&= \langle\Psi|\,\hat{A}\hat{B}\,|\Psi\rangle - \langle A\rangle\,\langle B\rangle - \langle B\rangle\,\langle A\rangle + \langle A\rangle\,\langle B\rangle \\
&= \langle\Psi|\,\hat{A}\hat{B}\,|\Psi\rangle - \langle A\rangle\,\langle B\rangle
\end{aligned}$$

and similarly

$$\langle\Phi|\Omega\rangle = \langle\Psi|\,\hat{B}\hat{A}\,|\Psi\rangle - \langle B\rangle\,\langle A\rangle\,.$$

Thus

$$\begin{aligned}
|\langle\Omega|\Phi\rangle| &\geqslant \frac{1}{2}\left|\langle\Psi|\,\hat{A}\hat{B}\,|\Psi\rangle - \langle A\rangle\,\langle B\rangle - \langle\Psi|\,\hat{B}\hat{A}\,|\Psi\rangle + \langle B\rangle\,\langle A\rangle\right| \\
&= \frac{1}{2}\left|\langle\Psi|\,\hat{A}\hat{B}\,|\Psi\rangle - \langle\Psi|\,\hat{B}\hat{A}\,|\Psi\rangle\right| \\
&= \frac{1}{2}\left|\langle\Psi|\,\left[\hat{A},\hat{B}\right]|\Psi\rangle\right|
\end{aligned}$$

which completes the proof. •

Commutators of observables occur repeatedly throughout quantum mechanics; in some instances it provides a good starting point for establishing a quantum mechanical description of a physical system. The following properties of the commutator are easily established.

$$[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}] \tag{II.3.78a}$$

$$[\alpha\hat{A}, \hat{B}] = \alpha[\hat{A}, \hat{B}] \tag{II.3.78b}$$

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}] \tag{II.3.78c}$$

where $\hat{A}, \hat{B}$ and $\hat{C}$ are any operators and α is any complex number.

Exercise: Prove the above.

Example: Suppose that a spin-1/2 particle is in a magnetic field $\mathbf{B} = B_z\hat{\boldsymbol{z}}$. Determine $\left[\hat{H}, \hat{S}_x\right]$.

Answer: First Eq. (II.3.73) implies that

$$\hat{H} = -\frac{gq}{2m}\,B\hat{S}_z.$$

Thus Eq. (II.3.78b) implies

$$\begin{aligned}\left[\hat{H}, \hat{S}_x\right] &= -\frac{gq}{2m} B\left[\hat{S}_z, \hat{S}_z\right] \\ &= -\frac{gq}{2m} B\left(\hat{S}_z\hat{S}_z - \hat{S}_z\hat{S}_z\right) \\ &= 0.\end{aligned}$$

An important example involves commutators of the observables corresponding to distinct spin components.

Example: Determine $\left[\hat{S}_z, \hat{S}_x\right]$.

Answer: The definition of the commutator gives

$$\left[\hat{S}_z, \hat{S}_x\right] = \hat{S}_z, \hat{S}_x - \hat{S}_x, \hat{S}_z.$$

One technique for computing this and which does not rely on matrix representations would be to use

$$\begin{aligned}\hat{S}_z &= \frac{\hbar}{2} \left|+\hat{\boldsymbol{z}}\right\rangle \left\langle+\hat{\boldsymbol{z}}\right| - \frac{\hbar}{2} \left|-\hat{\boldsymbol{z}}\right\rangle \left\langle-\hat{\boldsymbol{z}}\right| \\ \hat{S}_x &= \frac{\hbar}{2} \left|+\hat{\boldsymbol{x}}\right\rangle \left\langle+\hat{\boldsymbol{x}}\right| - \frac{\hbar}{2} \left|-\hat{\boldsymbol{x}}\right\rangle \left\langle-\hat{\boldsymbol{x}}\right|\end{aligned}$$

together with repeated applications of the inner product rule. A more concrete approach uses matrix representations in the $\left\{\left|+\hat{\boldsymbol{z}}\right\rangle, \left|-\hat{\boldsymbol{z}}\right\rangle\right\}$ basis. Thus

$$\begin{aligned}\left[\hat{S}_z, \hat{S}_x\right] &= \hat{S}_z, \hat{S}_x - \hat{S}_x, \hat{S}_z \\ &\leftrightarrow \frac{\hbar}{2}\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}0 & 1 \\ 1 & 0\end{pmatrix} - \frac{\hbar}{2}\begin{pmatrix}0 & 1 \\ 1 & 0\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix} \\ &= \frac{\hbar^2}{4}\begin{pmatrix}0 & 1 \\ -1 & 0\end{pmatrix} - \frac{\hbar^2}{4}\begin{pmatrix}0 & -1 \\ 1 & 0\end{pmatrix} \\ &= \frac{\hbar^2}{2}\begin{pmatrix}0 & 1 \\ -1 & 0\end{pmatrix} \\ &= i\hbar\begin{pmatrix}0 & -i \\ i & 0\end{pmatrix} \\ &\leftrightarrow i\hbar\hat{S}_y.\end{aligned}$$

Thus

$$\left[\hat{S}_z, \hat{S}_x\right] = i\hbar\hat{S}_y.$$

Similar derivations enable one to show that

$$\boxed{\begin{aligned}\left[\hat{S}_x, \hat{S}_y\right] &= i\hbar\hat{S}_z \\ \left[\hat{S}_y, \hat{S}_z\right] &= i\hbar\hat{S}_x \\ \left[\hat{S}_z, \hat{S}_x\right] &= i\hbar\hat{S}_y.\end{aligned}} \tag{II.3.79}$$

An application to the uncertainties in measurement outcomes is as follows.

Example: Suppose that an ensemble of spin-1/2 particles are each in the state $|+\hat{\boldsymbol{z}}\rangle$ and that either S_x or S_y is measured. Determine ΔS_x and ΔS_y and verify that the uncertainty principle is valid.

Answer: First

$$|+\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix}1\\0\end{pmatrix}.$$

Thus

$$\begin{aligned}\langle S_x\rangle &= \langle -\hat{\boldsymbol{z}}|\,\hat{S}_x\,|+\hat{\boldsymbol{z}}\rangle \\ &= \begin{pmatrix}1 & 0\end{pmatrix}\frac{\hbar}{2}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} \\ &= 0.\end{aligned}$$

Then

$$\begin{aligned}\left\langle S_x^2\right\rangle &= \langle -\hat{\boldsymbol{z}}|\,\hat{S}_x^2\,|+\hat{\boldsymbol{z}}\rangle \\ &= \begin{pmatrix}1 & 0\end{pmatrix}\frac{\hbar^2}{4}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} \\ &= \begin{pmatrix}1 & 0\end{pmatrix}\frac{\hbar^2}{4}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} \\ &= \frac{\hbar^2}{4}\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} \\ &= \frac{\hbar^2}{4}.\end{aligned}$$

Thus

$$\begin{aligned}\Delta S_x &= \sqrt{\left\langle S_x^2\right\rangle - \langle S_x\rangle^2} \\ &= \frac{\hbar}{2}.\end{aligned}$$

Similarly

$$\Delta S_y = \frac{\hbar}{2}.$$

Now the right hand side of Eq. (II.3.76) requires

$$\begin{aligned}\langle +\hat{\boldsymbol{z}}| \, [\hat{S}_x, \hat{S}_y] \, |+\hat{\boldsymbol{z}}\rangle &= \langle +\hat{\boldsymbol{z}}| \, i\hbar \hat{S}_z \, |+\hat{\boldsymbol{z}}\rangle \\ &= i\hbar \, \frac{\hbar}{2} \, \langle +\hat{\boldsymbol{z}}| + \hat{\boldsymbol{z}}\rangle \\ &= i\frac{\hbar^2}{2}.\end{aligned}$$

Thus the right hand side of Eq. (II.3.76) gives $\hbar^2/4$. In this case the uncertainty principle reduces to an equality.

In the previous example the right hand side of the uncertainty principle returned a non-zero value and this necessarily implied that for the state $|\Psi\rangle = |+\hat{\boldsymbol{z}}\rangle$, measurements of S_x or S_y could not return particular values with certainty. However, for arbitrary states the right hand side of Eq. (II.3.76) is not necessarily zero.

Exercise: Consider an ensemble in the state $|+\hat{\boldsymbol{x}}\rangle$. Show that the right hand side of Eq. (II.3.76) returns 0.

For the state of the previous exercise, one can show that $\Delta x = 0$. Thus the uncertainty principle cannot be used to extract any information about Δy. The following general theorem from linear algebra clarifies such situations.

Theorem: Two observables, $\hat{A}$ and $\hat{B}$ satisfy

$$[\hat{A}, \hat{B}] = 0$$

if and only if there exist a complete set of states $\{|\phi_{a,b}\rangle\}$ which are simultaneously eigenstates of both operators,

$$\begin{aligned}\hat{A} \, |\phi_{a,b}\rangle &= a \, |\phi_{a,b}\rangle \\ \hat{B} \, |\phi_{a,b}\rangle &= b \, |\phi_{a,b}\rangle\end{aligned}$$

Proof: Later, . . . •

It follows that if such a simultaneous set of eigenstates exist then they are the states that correspond to definite measurement outcomes of both quantities A and B. Applied to spin-1/2, this means that there are no states which will give definite outcomes for measurements of both S_x and S_y since $[\hat{S}_x, \hat{S}_y] \neq 0$. This is clearly a stronger statement than that offered by the uncertainty principle.

4 Dynamics of Spin-1/2 systems

The mathematical framework of the previous section is sufficient to answer questions regarding the kinematics of quantum systems. In particular it provides methods for determining the outcomes of measurements and the probabilities with which these occur. In the context of spin-1/2 systems one can use this to determine the probability with which a particle in a known initial state will emerge from a sequence of Stern-Gerlach measurements as illustrated in Fig. II.4.1.

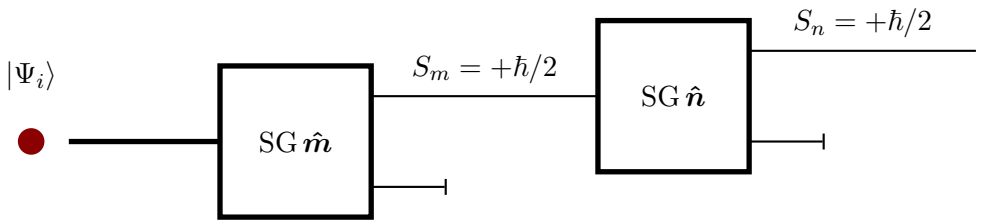


Figure II.4.1: Archetypal quantum kinematics situation. A particle in a known initial state, $|\Psi_i\rangle$, is subjected to a sequence of Stern-Gerlach devices in known directions, $\hat{\boldsymbol{m}}$ and $\hat{\boldsymbol{n}}$, and with certain output beams blocked. The task is to determine the probability with which it emerges from the last device and the state with which it emerges from the final Stern-Gerlach device.

Exercise: With reference to Fig. II.4.1, suppose that $\hat{\boldsymbol{m}} = \hat{\boldsymbol{z}}$ and $\hat{\boldsymbol{n}} = \hat{\boldsymbol{y}}$ and that $|\Psi_i\rangle = |+\hat{\boldsymbol{x}}\rangle$. Show that the probability with which the particle emerges from the last Stern-Gerlach apparatus is 1/4. If the particle does emerge from the last Stern-Gerlach apparatus, determine its state.

These questions can only be answered using the framework for dealing with measurements outcomes and probabilities *provided that between successive measurements the particle is subject to no outside influences that could alter its state.* The subject of dynamics deals with describing the evolution of the state of a spin-1/2 particle in situations where outside influences will change its state.

4.1 Evolution of classical magnetic dipoles.

A magnetic field $\mathbf{B}$ will exert a torque on classical magnetic dipole with dipole moment. The torque is given by

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B} \tag{II.4.1}$$

where $\boldsymbol{\mu}$ is the magnetic dipole moment of the particle. The torque determine the change in the spin angular momentum via

$$\frac{d\mathbf{S}}{dt} = \boldsymbol{\tau}. \tag{II.4.2}$$

Eqs. (II.1.5) and (II.4.1) then imply that for a particle of mass, M and charge Q,

$$\frac{d\boldsymbol{\mu}}{dt} = \frac{gQ}{2M}\,\boldsymbol{\mu} \times \mathbf{B} \tag{II.4.3}$$

The important conceptual consequence of this is that *a magnetic field generally alters the magnetic dipole moment vector of a particle placed in that field.*

Exercise: A magnetic dipole is placed in a constant magnetic field $\mathbf{B} = B_0\hat{\boldsymbol{z}}$. Show that the dipole moment components satisfy:

$$\mu_x(t) = \cos(\omega t)\mu_x(0) + \sin(\omega t)\mu_y(0) \tag{II.4.4a}$$
$$\mu_y(t) = -\sin(\omega t)\mu_x(0) + \cos(\omega t)\mu_y(0) \tag{II.4.4b}$$
$$\mu_z(t) = \mu_z(0) \tag{II.4.4c}$$

Find an expression for ω in terms of q, Q, M and B_0.

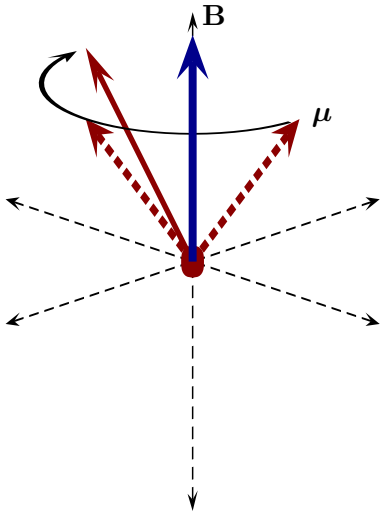


Figure II.4.2: Precession of a classical magnetic dipole moment $\boldsymbol{\mu}$ about an applied external magnetic field $\mathbf{B}$.

In the example of a magnetic dipole in a homogeneous external magnetic field oriented along $\hat{\boldsymbol{z}}$, Eqs. (II.4.4) indicate that the magnetic dipole moment maintains a constant magnitude and that the component along the direction of the magnetic field remains constant. A projection onto the xy plane would show that the component in this plane rotates in the plane with angular frequency ω. Thus the magnetic dipole moment rotates with an angular frequency ω about an axis directed along the magnetic field. This behavior of dipole moments is frequently encountered in Nuclear Magnetic Resonance (NMR) and sometimes in other branches of atomic and molecular physics and is called *Larmor precession* or *Zeeman precession.*

Eqs. (II.4.4) refer to an initial magnetic dipole moment. Rather than track the progress of a magnetic dipole moment, starting from a specified initial state, we shall consider the process by which any dipole moment will change when placed in this field.

Exercise: A magnetic dipole is placed in a homogeneous external magnetic field $\mathbf{B} = B_0\hat{\boldsymbol{z}}$. Show that

$$\begin{pmatrix}\mu_x(t)\\ \mu_y(t)\\ \mu_z(t)\end{pmatrix} = \begin{pmatrix}\cos(\omega t) & \sin(\omega t) & 0\\ -\sin(\omega t) & \cos(\omega t) & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\mu_x(0)\\ \mu_y(0)\\ \mu_z(0)\end{pmatrix} \tag{II.4.5}$$

where ω is a constant which depends on q, Q, M and B_0.

The preceding exercise shows it is possible to isolate the information about the initial state, $\boldsymbol{\mu}(0)$, in the description of the evolution of the magnetic dipole moment. More precisely, the evolution of a magnetic dipole moment in a homogeneous external magnetic field along $\hat{\boldsymbol{z}}$ can be described via

$$\boldsymbol{\mu}(t) = R\,\boldsymbol{\mu}(0) \tag{II.4.6}$$

where

$$R = \begin{pmatrix}\cos(\omega t) & \sin(\omega t) & 0\\ -\sin(\omega t) & \cos(\omega t) & 0\\ 0 & 0 & 1\end{pmatrix}. \tag{II.4.7}$$

Eq. (II.4.6) clearly separates the information about the initial state, $\boldsymbol{\mu}(0)$, and the mathematical mechanism by which this state is to be transformed, i.e. the matrix R. The matrix R only depends on the external magnetic field, the coupling between the field and the magnetic dipole moment and the duration of the evolution. It is *the same regardless of the initial state of the magnetic dipole moment* and, as such, describes time evolution of the dipole in a state-independent way.

Mathematically the matrix of Eq. (II.4.6) describes a rotation of a three dimensional vector through angle ωt in a clockwise sense about the $\hat{\boldsymbol{z}}$ axis. An initial attempt at understanding a rotation typically involves applying the rotation to one or more particular vectors and observing the resulting changes in these vectors. However, a rotation has an identity independent of any particular vector to which it may be applied and rotations can be combined and described without any reference to the vectors on which they operate.

In the same sense, the evolution matrix R has an identity independent of any initial magnetic dipole moment and one can describe the evolution of the dipole by only referring to this matrix. The central issue in the evolution of classical magnetic dipole moments is to determine the relevant rotation matrices, which can be complicated for time varying magnetic fields. This is a description of the *evolution process* rather than the evolution of a magnetic dipole moment known to start in a particular initial state. These notions can all be translated directly into statements about spin angular momenta. Thus

$$\mathbf{S}(t) = R\,\mathbf{S}(0) \tag{II.4.8}$$

where, again

$$R = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) & 0 \\ -\sin(\omega t) & \cos(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{II.4.9}$$

The process of classical evolution can be placed into the context of pre- and post-evolution measurements and their outcomes in a way which will be useful for comparison with quantum mechanics. A schematic view is provided by Fig. II.4.3.

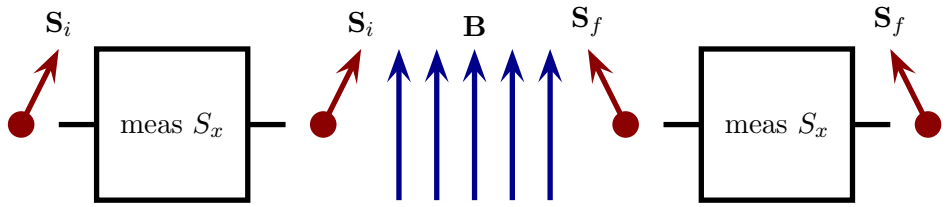


Figure II.4.3: Classical magnetic dipole, with spin angular momentum initially $\mathbf{S}_i$ passing through measuring devices and a magnetic field. The measuring devices do not change the spin of the particle. However, the spin precesses in the magnetic field and prior to the second measuring device is $\mathbf{S}_f$. A consequence of this is that measurements of S_x before and after the field regions will yield different outcomes. Mathematically the effects of the field are capture by a matrix R and $\mathbf{S}_f = R\,\mathbf{S}_i$. In the archetypal kinematics discussion the field is absent and $\mathbf{S}$ is unaltered.

One key feature of the evolution of a classical magnetic dipole moment that transfers to the quantum mechanical description is that the evolution is linear, i.e. if $\boldsymbol{\mu}(0) = a\boldsymbol{\mu}_1 + b\boldsymbol{\mu}_2$ then

$$\boldsymbol{\mu}(t) = R\,\boldsymbol{\mu}(0) = aR\,\boldsymbol{\mu}_1 + bR\,\boldsymbol{\mu}_2 \tag{II.4.10}$$

for any real numbers a and b.

4.2 Evolution of quantum mechanical systems

Given that an external magnetic field changes the state of a classical magnetic dipole and that measurement outcomes for spin components after the period of evolution will be typically differ from those prior to the period of evolution, it is reasonable to expect that the state of the quantum mechanical system must undergo a change. This is to be described in terms of a change from an initial state $|\Psi_i\rangle$ before the evolution to a final state $|\Psi_f\rangle$ after the evolution. A schematic diagram illustrates this in terms of measurement outcomes in Fig. II.4.4.

The general prescription for the evolution of spin-1/2 system is resembles that for a classical dipole and is given by

$$|\Psi_f\rangle = \hat{U}\,|\Psi_i\rangle \tag{II.4.11}$$

where $\hat{U}$ is a *linear operator* on the space of kets. This is called an *evolution operator* and it will be shown to have additional constraining properties. In the context of representations

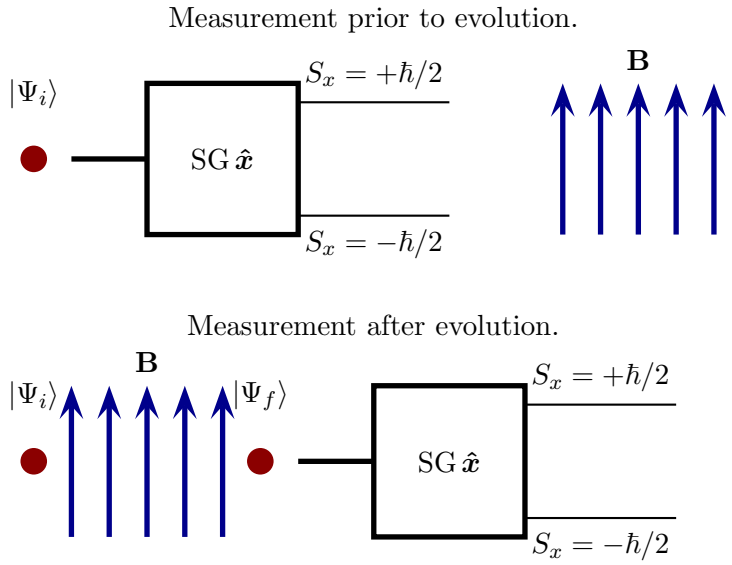


Figure II.4.4: Quantum mechanical measurements on a magnetic dipole, initially in the state $|\Psi_i\rangle$. In one case the measurement is made prior to evolution in the external magnetic field. In the other case after the dipole has passed through the external field and its state has been transformed from $|\Psi_i\rangle$ to $|\Psi_f\rangle$. Note that the process in which the particle is first subjected to an SG $\hat{\boldsymbol{x}}$ measurement, then allowed to pass through the field and then subjected to a second SG $\hat{\boldsymbol{x}}$ measurement is distinct from each of the two displayed.

in terms of vectors and matrices, the evolution operator is a 2×2 complex matrix. For a spin-1/2 system in an external magnetic field, the evolution operator will depend on the external magnetic field's strength and direction as well as the duration for which the magnetic field is applied. Varying these parameters will generally result in distinct evolution operators.

Example: Represented in terms of the basis $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$, one possibility for a evolution operator is

$$\hat{U} \leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Given that $|\Psi_i\rangle = |+\hat{\boldsymbol{z}}\rangle$, determine $|\Psi_f\rangle$.

Answer: Here

$$\begin{aligned} |\Psi_f\rangle &= \hat{U}\,|+\hat{\boldsymbol{z}}\rangle \\ &\leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &\leftrightarrow |-\hat{\boldsymbol{z}}\rangle\,. \end{aligned}$$

Thus $|\Psi_f\rangle = \hat{U}\,|+\hat{\boldsymbol{z}}\rangle = |-\hat{\boldsymbol{z}}\rangle\,.$

The effects of an evolution operator on a general state can be described by using matrix and vector representations of kets and operators in a basis. For example, in the $\{|+\hat{\boldsymbol{z}}\rangle\,, |-\hat{\boldsymbol{z}}\rangle\}$ basis, a general initial state is represented as

$$|\Psi_i\rangle = c_+\,|+\hat{\boldsymbol{z}}\rangle + c_-\,|-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} c_+ \\ c_- \end{pmatrix} \tag{II.4.12}$$

where $c_\pm$ are complex. Thus

$$|\Psi_f\rangle = \hat{U}\,|\Psi_i\rangle \leftrightarrow \begin{pmatrix} U_{++} & U_{+-} \\ U_{-+} & U_{--} \end{pmatrix} \begin{pmatrix} c_+ \\ c_- \end{pmatrix} \tag{II.4.13}$$

where $U_{++}, \ldots$ are complex numbers which will depend on the details of the external influences that drive the evolution. The matrix

$$\begin{pmatrix} U_{++} & U_{+-} \\ U_{-+} & U_{--} \end{pmatrix}$$

used here is the representation of $\hat{U}$ in the $\{|+\hat{\boldsymbol{z}}\rangle\,, |-\hat{\boldsymbol{z}}\rangle\}$ basis.

A convenient technique for representing an operator proceeds as follows. Consider

$$\begin{aligned} \hat{U}\,|+\hat{\boldsymbol{z}}\rangle &\leftrightarrow \begin{pmatrix} U_{++} & U_{+-} \\ U_{-+} & U_{--} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} U_{++} \\ U_{-+} \end{pmatrix}. \end{aligned}$$

Thus

$$\begin{aligned} \langle+\hat{\boldsymbol{z}}|\,\hat{U}\,|+\hat{\boldsymbol{z}}\rangle &\leftrightarrow {}^{(1\ \ 0)} \begin{pmatrix} U_{++} \\ U_{-+} \end{pmatrix} \\ &= U_{++}. \end{aligned}$$

Using all combinations of $\{|+\hat{\boldsymbol{z}}\rangle\,, |-\hat{\boldsymbol{z}}\rangle\}$ gives

$$U_{++} = \langle+\hat{\boldsymbol{z}}|\,\hat{U}\,|+\hat{\boldsymbol{z}}\rangle \tag{II.4.14a}$$

$$U_{+-} = \langle+\hat{\boldsymbol{z}}|\,\hat{U}\,|-\hat{\boldsymbol{z}}\rangle \tag{II.4.14b}$$

$$U_{-+} = \langle-\hat{\boldsymbol{z}}|\,\hat{U}\,|+\hat{\boldsymbol{z}}\rangle \tag{II.4.14c}$$

$$U_{--} = \langle-\hat{\boldsymbol{z}}|\,\hat{U}\,|-\hat{\boldsymbol{z}}\rangle\,. \tag{II.4.14d}$$

These offer an algebraic method of determining the matrix representation of any operator $\hat{U}$. Thus

$$\hat{U} \leftrightarrow \begin{pmatrix} \langle +\hat{\boldsymbol{z}}| \hat{U} |+\hat{\boldsymbol{z}}\rangle & \langle +\hat{\boldsymbol{z}}| \hat{U} |-\hat{\boldsymbol{z}}\rangle \\ \langle -\hat{\boldsymbol{z}}| \hat{U} |+\hat{\boldsymbol{z}}\rangle & \langle -\hat{\boldsymbol{z}}| \hat{U} |-\hat{\boldsymbol{z}}\rangle \end{pmatrix}. \tag{II.4.15}$$

Example: Determine the matrix representation of

$$\hat{U} = |+\hat{\boldsymbol{x}}\rangle \langle +\hat{\boldsymbol{x}}| - |-\hat{\boldsymbol{x}}\rangle \langle -\hat{\boldsymbol{x}}|$$

in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis.

Answer: Eq. (II.4.14) requires

$$\langle +\hat{\boldsymbol{z}}| \hat{U} |+\hat{\boldsymbol{z}}\rangle = \langle +\hat{\boldsymbol{z}}| + \hat{\boldsymbol{x}}\rangle \langle +\hat{\boldsymbol{x}}| + \hat{\boldsymbol{z}}\rangle - \langle +\hat{\boldsymbol{z}}| - \hat{\boldsymbol{x}}\rangle \langle -\hat{\boldsymbol{x}}| + \hat{\boldsymbol{z}}\rangle$$

The inner products are calculated using

$$|+\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}} |+\hat{\boldsymbol{x}}\rangle + \frac{1}{\sqrt{2}} |-\hat{\boldsymbol{x}}\rangle$$
$$|-\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}} |+\hat{\boldsymbol{x}}\rangle - \frac{1}{\sqrt{2}} |-\hat{\boldsymbol{x}}\rangle.$$

Thus

$$\langle +\hat{\boldsymbol{z}}| + \hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}}.$$

A similar series of calculations yields

$$\langle +\hat{\boldsymbol{z}}| \hat{U} |+\hat{\boldsymbol{z}}\rangle = \frac{1}{2} - \frac{1}{2} = 0$$
$$\langle +\hat{\boldsymbol{z}}| \hat{U} |-\hat{\boldsymbol{z}}\rangle = \frac{1}{2} + \frac{1}{2} = 1$$
$$\langle -\hat{\boldsymbol{z}}| \hat{U} |+\hat{\boldsymbol{z}}\rangle = \frac{1}{2} + \frac{1}{2} = 1$$
$$\langle -\hat{\boldsymbol{z}}| \hat{U} |-\hat{\boldsymbol{z}}\rangle = \frac{1}{2} - \frac{1}{2} = 0.$$

Thus

$$\hat{U} \leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

4.3 Unitary Operators

Given that an evolution operator is a linear operator, one question that arises is whether any linear operator represents a possible evolution of a quantum system. For example $|+\hat{\boldsymbol{z}}\rangle \langle +\hat{\boldsymbol{z}}|$ is a linear operator. One requirement for evolution operators is that for a normalized initial state, the resulting final the state must be normalized. Thus for any $|\Psi_i\rangle$, the state $|\Psi_f\rangle$ satisfies

$$\langle \Psi_f | \Psi_f \rangle = 1.$$

Now

$$\begin{aligned}\langle\Psi_f| &= |\Psi_f\rangle^\dagger \\ &= \left[\hat{U}\,|\Psi_i\rangle\right]^\dagger \\ &= |\Psi_i\rangle^\dagger\,\hat{U}^\dagger \\ &= \langle\Psi_i|\,\hat{U}^\dagger.\end{aligned}$$

Thus a requirement is that for any $|\Psi_i\rangle$,

$$\langle\Psi_i|\,\hat{U}^\dagger\hat{U}\,|\Psi_i\rangle = 1.$$

Clearly one way to satisfy this is

$$\hat{U}^\dagger\hat{U} = \hat{I}.$$

The following theorem guarantees that this is also a necessary requirement.

Theorem: Any evolution operator for an isolated quantum system is a *unitary operator,* i.e. it satisfies

$$\hat{U}^\dagger\hat{U} = \hat{I} \tag{II.4.16}$$

Proof: The evolution operator must satisfy

$$\langle\Psi_i|\,\hat{U}^\dagger\hat{U}\,|\Psi_i\rangle = 1.$$

for any state $|\Psi_i\rangle$. Let $\hat{N} = \hat{U}^\dagger\hat{U}$. This operator is normal, which means that its satisfies

$$\hat{N}^\dagger\hat{N} = \hat{N}\hat{N}^\dagger.$$

A general theorem from linear algebra states that a normal $\hat{N}$ operator can be diagonalized, meaning that there exists an orthonormal basis $\{|\phi_1\rangle, |\phi_2\rangle\}$ such that

$$\hat{N} = \lambda_1\,|\phi_1\rangle\,\langle\phi_1| + \lambda_2\,|\phi_2\rangle\,\langle\phi_2|$$

where λ_1, λ_2 are complex numbers. Now,

$$\langle\phi_i|\,\hat{N}\,|\phi_i\rangle = \lambda_i.$$

But this implies that

$$\langle\phi_i|\,\hat{U}^\dagger\hat{U}\,|\phi_i\rangle = \lambda_i,$$

giving $\lambda_i = 1$. Thus

$$\hat{U}^\dagger\hat{U} = |\phi_1\rangle\,\langle\phi_1| + |\phi_2\rangle\,\langle\phi_2|$$

and this implies that

$$\hat{U}^\dagger\hat{U} = \hat{I}$$

which completes the proof. •

The previous theorem restricts evolution operators to unitary operators. The converse question could be asked: "Given any unitary operator, does this represent a conceivable evolution of a quantum system?" For spin-1/2 particles we will demonstrate explicitly that this is true by characterizing all possible unitary transformations for spin-1/2 particles.

Example: Determine whether $\hat{A} := |+\hat{\boldsymbol{z}}\rangle \langle +\hat{\boldsymbol{z}}|$ is unitary.

Answer: The matrix representation of $|+\hat{\boldsymbol{z}}\rangle \langle +\hat{\boldsymbol{z}}|$ is

$$|+\hat{\boldsymbol{z}}\rangle |+\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} \hat{A}^\dagger \hat{A} \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^\dagger \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \\ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

This is not the identity operator and thus $\hat{A}$ is not unitary.

As the above example shows, the unitary requirement for an evolution operator is fairly restrictive. We now aim to characterize all unitary operators for spin-1/2 systems in terms of rotation operators.

4.4 Rotations

The notion of rotations applied to kets arises from the following considerations. Suppose that a spin-1/2 system is in the state $|+\hat{\boldsymbol{x}}\rangle$. Physically this means that if S_x is measured via a SG $\hat{\boldsymbol{x}}$ apparatus then the outcome $S_x = +\hbar/2$ will occur with certainty. If both the physical system and the measuring apparatus are rotated through φ about the $\hat{\boldsymbol{z}}$ then the same outcome will occur with certainty. However, the apparatus has been rotated and now the measurement is no longer of the component along $\hat{\boldsymbol{x}}$ but of the component of spin along the direction $\hat{\boldsymbol{n}} = \cos\varphi\hat{\boldsymbol{x}} + \sin\varphi\hat{\boldsymbol{y}}$. It follows that the state after rotation must be $|+\hat{\boldsymbol{n}}\rangle$.

For example, with $\varphi = \pi/2$, the measurement is now done by an SG $\hat{\boldsymbol{y}}$ apparatus and the outcome is $S_y = +\hbar/2$ with certainty. Thus the state is after this rotation must be $|+\hat{\boldsymbol{y}}\rangle$. The rule for rotating states is as follows:

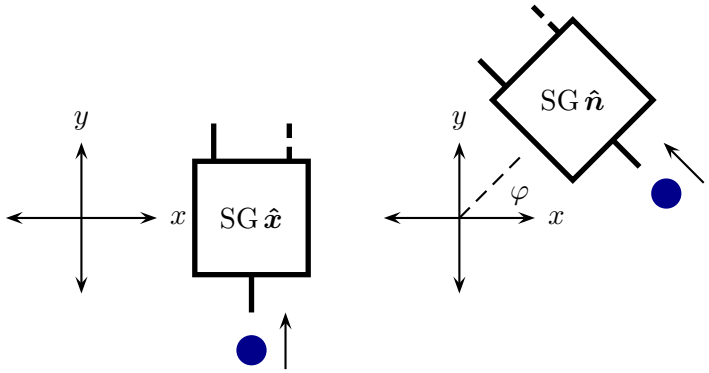


Figure II.4.5: Rotation of a Stern-Gerlach measuring apparatus and the particle through φ about $\hat{\boldsymbol{z}}$. If the initial state of the particle was $|+\hat{\boldsymbol{x}}\rangle$ then the particle emerges from the SG $\hat{\boldsymbol{x}}$ apparatus in the upper (left in this diagram) beam with certainty. It must again do so if both the particle and apparatus are rotated. However the Stern-Gerlach is now oriented along a new axis $\hat{\boldsymbol{n}} = \cos\varphi\hat{\boldsymbol{x}} + \sin\varphi\hat{\boldsymbol{y}}$. Thus the state after rotation is $|+\hat{\boldsymbol{n}}\rangle$.

If a particle is initially in the state $|+\hat{\boldsymbol{n}}_i\rangle$ and is subjected to a rotation through angle φ about axis $\hat{\boldsymbol{m}}$ then the state transforms as:

$$|+\hat{\boldsymbol{n}}_i\rangle \mapsto |+\hat{\boldsymbol{n}}_f\rangle \tag{II.4.17}$$

where $\hat{\boldsymbol{n}}_f$ is the vector obtained by rotating $\hat{\boldsymbol{n}}_i$ about $\hat{\boldsymbol{m}}$ through angle φ.

The operation of a rotation through angle φ about axis $\hat{\boldsymbol{n}}$ on kets is denoted $\hat{R}(\varphi\boldsymbol{n})$. The following theorem guarantees that these are linear transformations.

Theorem: The transformation of states as described by Eq. (II.4.17) is linear.

Proof: To follow. •

Example: Determine that state produced by the rotation $\hat{R}(\frac{\pi}{4}\boldsymbol{z})$ acting on $|+\hat{\boldsymbol{x}}\rangle$.

Answer: According to Eq (II.4.17)

$$\hat{R}(\tfrac{\pi}{4}\boldsymbol{z})\,|+\hat{\boldsymbol{x}}\rangle = |+\hat{\boldsymbol{n}}_f\rangle$$

where $\hat{\boldsymbol{n}}_f$ is the vector obtained by rotating $\hat{\boldsymbol{x}}$ through angle $\pi/4$. Standard geometrical considerations imply that $\hat{\boldsymbol{n}}_f = 1/\sqrt{2}\,\hat{\boldsymbol{x}}+1/\sqrt{2}\,\hat{\boldsymbol{y}}$. The spherical angle parameters for this are $\theta = \pi/2$ and $\phi = \pi/4$. Thus

$$\begin{aligned}|+\hat{\boldsymbol{n}}_f\rangle &= \cos\left(\frac{\pi}{4}\right)\,|+\hat{\boldsymbol{z}}\rangle + e^{i\pi/4}\,\sin\left(\frac{\pi}{4}\right)\,|-\hat{\boldsymbol{z}}\rangle \\ &= \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle + e^{i\pi/4}\,\frac{1}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle\,.\end{aligned}$$

Thus

$$\hat{R}(\tfrac{\pi}{4}\boldsymbol{z})\,|+\hat{\boldsymbol{x}}\rangle = \frac{1}{\sqrt{2}}\,|+\hat{\boldsymbol{z}}\rangle + e^{i\pi/4}\,\frac{1}{\sqrt{2}}\,|-\hat{\boldsymbol{z}}\rangle\,.$$

Given the geometrical definition of the rotation operator on kets it remains to show that the resulting rotation operator is unitary. The following theorem guarantees that this is true.

Theorem: The rotation operators as defined by Eq (II.4.17) are unitary.

Proof: For any input state $|+\hat{\boldsymbol{n}}_i\rangle$ the rotation is a linear transformation that produces $|+\hat{\boldsymbol{n}}_f\rangle = \hat{R}(\varphi\boldsymbol{n})\,|+\hat{\boldsymbol{n}}_i\rangle$. Then $\langle+\hat{\boldsymbol{n}}_f| + \hat{\boldsymbol{n}}_f\rangle = 1$ by the properties of the state $|+\hat{\boldsymbol{n}}_f\rangle$. Thus $\hat{R}(\varphi\boldsymbol{n})$ is unitary.

It will be useful to be able to construct rotation operators in purely algebraic terms without resorting to geometrical notions such as those above. To motivate the general result, consider the rotation about the z axis through angle φ as applied to the state $|+\hat{\boldsymbol{n}}_i\rangle$. Let θ and ϕ be the standard spherical angle parameters for $\hat{\boldsymbol{n}}_i$. Thus

$$|+\hat{\boldsymbol{n}}_i\rangle = \cos\left(\frac{\theta}{2}\right)\,|+\hat{\boldsymbol{z}}\rangle + e^{i\phi}\,\sin\left(\frac{\theta}{2}\right)\,|-\hat{\boldsymbol{z}}\rangle$$

By the geometrical definition of $\hat{R}(\varphi\boldsymbol{z})$, the standard spherical angles for $\hat{\boldsymbol{n}}_f$ are θ and $\phi+\varphi$ provided that $0 \leqslant \phi+\varphi \leqslant 2\pi$ (cases beyond this range can be accommodated by determining the remainder after dividing by 2π). Thus

$$\begin{aligned}\hat{R}(\varphi\boldsymbol{z})\,|+\hat{\boldsymbol{n}}_i\rangle &= |+\hat{\boldsymbol{n}}_f\rangle \\ &= \cos\left(\frac{\theta}{2}\right)\,|+\hat{\boldsymbol{z}}\rangle + e^{i(\phi+\varphi)}\,\sin\left(\frac{\theta}{2}\right)\,|-\hat{\boldsymbol{z}}\rangle\end{aligned}$$

The form $\hat{R}(\varphi\boldsymbol{z})$ can be ascertained by considering its representation in the $\{|+\hat{\boldsymbol{z}}\rangle\,,|-\hat{\boldsymbol{z}}\rangle\}$ basis. Suppose that

$$\hat{R}(\varphi\boldsymbol{z}) \leftrightarrow \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22}\end{pmatrix}.$$

Then Eq. (4.4) implies that

$$\begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ e^{i(\phi+\varphi)}\sin\left(\frac{\theta}{2}\right) \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) \end{pmatrix}$$

for any choices of θ and ϕ. Choosing $\theta = 0$ implies that $R_{21} = 0$ and choosing $\theta = \pi$ implies that $R_{12} = 0$. This implies that $R_{11} = 1$ and $R_{22} = e^{i\varphi}$. Thus

$$\hat{R}(\varphi \boldsymbol{z}) \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix}.$$

The corresponding operator is

$$\hat{R}(\varphi \boldsymbol{z}) = |+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + e^{i\varphi}|-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|.$$

Now, note that this can be expressed in a more symmetric form

$$\hat{R}(\varphi \boldsymbol{z}) \leftrightarrow e^{i\varphi/2}\begin{pmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{pmatrix}.$$

The overall factor $e^{i\varphi/2}$ provides a global phase to the states generated using $\hat{R}(\varphi \boldsymbol{z})$. This can be ignored and thus an alternative realization of this operator is

$$\hat{R}(\varphi \boldsymbol{z}) = e^{-i\varphi/2}|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + e^{i\varphi/2}|-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|.$$

The preceding example suggests the following algebraic form for rotation operators.

The operator corresponding to a rotation about $\hat{\boldsymbol{n}}$ through angle φ is

$$\hat{R}(\varphi \boldsymbol{n}) = e^{-i\varphi/2}|+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}| + e^{i\varphi/2}|-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}|. \tag{II.4.18}$$

Exercise: Show that $\hat{R}(\varphi \boldsymbol{n})$ as defined in Eq. (II.4.18) is unitary.

It is important to show that the operator of Eq. (II.4.18) is consistent with the requirement of Eq. (II.4.17). The following theorem guarantees this.

Theorem: Given an initial state $|+\hat{\boldsymbol{m}}_i\rangle$, the state produced by $\hat{R}(\varphi \boldsymbol{n})$,

$$|\Psi_f\rangle = \hat{R}(\varphi \boldsymbol{n})|+\hat{\boldsymbol{m}}_i\rangle$$

where $\hat{R}(\varphi \boldsymbol{n})$ is implemented using

$$\hat{R}(\varphi \boldsymbol{n}) = e^{-i\varphi/2}|+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}| + e^{i\varphi/2}|-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}|.$$

is such that

$$|\Psi_f\rangle = e^{i\alpha}\ |+\hat{\boldsymbol{m}}_f\rangle$$

$+\hat{\boldsymbol{m}}_f$ is the vector attained by rotating $+\hat{\boldsymbol{m}}_i$ about the axis $\hat{\boldsymbol{n}}$ through angle φ and α is real.

Proof: First, a geometrical construction demonstrates that rotating $\mathbf{v}_i$ about the axis $\hat{\boldsymbol{n}}$ through angle φ results in the vector

$$\mathbf{v}_f = \cos\varphi\,\mathbf{v}_i + \sin\varphi\,\hat{\boldsymbol{n}}\times\mathbf{v}_i + (1-\cos\varphi)\ (\hat{\boldsymbol{n}}\cdot\mathbf{v}_i)\,\hat{\boldsymbol{n}}.$$

The general strategy for the proof considers operators that are constructed from states. Specifically consider an operator that corresponds to the initial state, $|+\hat{\boldsymbol{m}}_i\rangle\langle+\hat{\boldsymbol{m}}_i|$. Denote the state produced from $|+\hat{\boldsymbol{m}}_i\rangle$ by application of $\hat{R}(\varphi\boldsymbol{n})$ by

$$|\Psi_f\rangle = \hat{R}(\varphi\boldsymbol{n})\,|+\hat{\boldsymbol{m}}_i\rangle\,.$$

We shall determine an expression for the operator corresponding to $|\Psi_i\rangle\langle\Psi_f|$ and the bulk of the proof will involve showing that

$$|\Psi_f\rangle\langle\Psi_f| = |+\hat{\boldsymbol{m}}_f\rangle\langle+\hat{\boldsymbol{m}}_f|$$

where $\hat{\boldsymbol{m}}_f$ is the vector obtained by rotating $+\hat{\boldsymbol{m}}_i$ about the axis $\hat{\boldsymbol{n}}$ through angle φ. Here this requires that,

$$\hat{\boldsymbol{m}}_f = \cos\varphi\,\hat{\boldsymbol{m}}_i + \sin\varphi\,\hat{\boldsymbol{n}}\times\hat{\boldsymbol{m}}_i + (1-\cos\varphi)\ (\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{m}}_i)\,\hat{\boldsymbol{n}}.$$

Given that we can show that $|\Psi_f\rangle\langle\Psi_f| = |+\hat{\boldsymbol{m}}_f\rangle\langle+\hat{\boldsymbol{m}}_f|$, can one conclude that $|\Psi_f\rangle = |+\hat{\boldsymbol{m}}_f\rangle$? To check this, we can express

$$|\Psi_f\rangle = \alpha\,|+\hat{\boldsymbol{m}}_f\rangle + \beta\,|\hat{\boldsymbol{m}}_{f\perp}\rangle$$

where $\langle+\hat{\boldsymbol{m}}_f|\hat{\boldsymbol{m}}_{f\perp}\rangle = 0$ and α,β are complex numbers that satisfy $|\alpha|^2+|\beta|^2=1$. Now

$$\langle+\hat{\boldsymbol{m}}_f|\Psi_f\rangle\langle\Psi_f|+\hat{\boldsymbol{m}}_f\rangle = \langle+\hat{\boldsymbol{m}}_f|+\hat{\boldsymbol{m}}_f\rangle\langle+\hat{\boldsymbol{m}}_f|+\hat{\boldsymbol{m}}_f\rangle = 1.$$

But by the previous expansion $\langle+\hat{\boldsymbol{m}}_f|\Psi_f\rangle = \alpha$. Thus $|\alpha|^2=1$ and $\beta=0$. Then $\alpha = e^{i\gamma}$ for some real γ. It follows that if we can show that $|\Psi_f\rangle\langle\Psi_f| = |+\hat{\boldsymbol{m}}_f\rangle\langle+\hat{\boldsymbol{m}}_f|$, then

$$|\Psi_f\rangle = e^{i\gamma}\ |+\hat{\boldsymbol{m}}_f\rangle$$

The crux of the proof is to determine how the operator $|+\hat{\boldsymbol{m}}_i\rangle\langle+\hat{\boldsymbol{m}}_i|$ transforms. To this end,

$$|\Psi_f\rangle\langle\Psi_f| = \hat{R}(\varphi\boldsymbol{n})\,|+\hat{\boldsymbol{m}}_i\rangle\langle+\hat{\boldsymbol{m}}_i|\,\hat{R}(\varphi\boldsymbol{n})^\dagger.$$

and the operator on the left can be computed by matrix representations of the three individual operators on the right. This is most readily done by representing operators and states in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis. It can be shown that

$$e^{-i\varphi/2}\ |+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}| + e^{i\varphi/2}\ |-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}| \leftrightarrow \begin{pmatrix} \cos\frac{\varphi}{2} - i\sin\frac{\varphi}{2}\cos\theta & \sin\frac{\varphi}{2}e^{-i\phi}\ \sin\theta \\ \sin\frac{\varphi}{2}e^{i\phi}\ \sin\theta & \cos\frac{\varphi}{2} + i\sin\frac{\varphi}{2} \end{pmatrix}$$

Now the axis of rotation is

$$\hat{\boldsymbol{n}} = \sin\theta\cos\phi\hat{\boldsymbol{x}} + \sin\theta\sin\phi\hat{\boldsymbol{x}} + \cos\theta,$$

where θ, ϕ are the spherical coordinates of $\hat{\boldsymbol{n}}$, and thus its components are

$$\begin{aligned} n_x &= \sin\theta\cos\phi \\ n_y &= \sin\theta\sin\phi \\ n_z &= \cos\theta. \end{aligned}$$

It follows that

$$\hat{R}(\varphi\boldsymbol{n}) \leftrightarrow \begin{pmatrix} \cos\frac{\varphi}{2} - i\sin\frac{\varphi}{2}\, n_z & \sin\frac{\varphi}{2}\,(n_x - in_y) \\ \sin\frac{\varphi}{2}\,(n_x + in_y) & \cos\frac{\varphi}{2} + i\sin\frac{\varphi}{2} n_z \end{pmatrix}$$

A convenient representation of this is

$$\hat{R}(\varphi\boldsymbol{n}) \leftrightarrow \cos\frac{\varphi}{2}\,\hat{I} - i\sin\frac{\varphi}{2}\,(n_x\hat{\sigma}_x + n_y\hat{\sigma}_y + n_z\hat{\sigma}_z)$$

where

$$\begin{aligned} \hat{\sigma}_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \hat{\sigma}_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \\ \hat{\sigma}_x &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Using $\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{\sigma}} := n_x\hat{\sigma}_x + n_y\hat{\sigma}_y + n_z\hat{\sigma}_z$ gives

$$\hat{R}(\varphi\boldsymbol{n}) \leftrightarrow \cos\frac{\varphi}{2}\,\hat{I} - i\sin\frac{\varphi}{2}\,\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{\sigma}}.$$

Similarly

$$\hat{R}(\varphi\boldsymbol{n})^\dagger \leftrightarrow \cos\frac{\varphi}{2}\,\hat{I} + i\sin\frac{\varphi}{2}\,\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{\sigma}}.$$

Additionally a straightforward calculation gives

$$\begin{aligned} |+\hat{\boldsymbol{m}}_i\rangle\langle+\hat{\boldsymbol{m}}_i| &\leftrightarrow \frac{1}{2}\begin{pmatrix} 2+\cos\theta_m & \frac{1}{2}\,e^{-i\phi_m}\sin\theta_m \\ \frac{1}{2}\,e^{i\phi_m}\sin\theta_m & 2 - cos\theta_m \end{pmatrix} \\ &= \frac{1}{2}\,\hat{I} + \frac{1}{2}\begin{pmatrix} m_z & m_x - im_y \\ m_x + im_y & -m_z \end{pmatrix} \\ &= \frac{1}{2}\,\hat{I} + \frac{1}{2}\,\hat{\boldsymbol{m}}_i\cdot\hat{\boldsymbol{\sigma}} \end{aligned}$$

where m_x, m_y, m_z are the components of $\hat{\boldsymbol{m}}_i$. In these terms

$$|\Psi_f\rangle\langle\Psi_f| = \left[\cos\frac{\varphi}{2}\,\hat{I} - i\sin\frac{\varphi}{2}\,\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{\sigma}}\right]\left[\frac{1}{2}\,\hat{I} + \frac{1}{2}\,\hat{\boldsymbol{m}}_i\cdot\hat{\boldsymbol{\sigma}}\right]\left[\cos\frac{\varphi}{2}\,\hat{I} + i\sin\frac{\varphi}{2}\,\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{\sigma}}\right].$$

The resulting matrix multiplication is greatly simplified by the readily proved fact that, for any vectors $\mathbf{u}, \mathbf{v}$,

$$[\mathbf{u} \cdot \hat{\boldsymbol{\sigma}}] [\mathbf{v} \cdot \hat{\boldsymbol{\sigma}}] = \mathbf{u} \cdot \mathbf{v} \hat{I} + i (\mathbf{u} \times \mathbf{v}) \cdot \hat{\boldsymbol{\sigma}}.$$

repeated application of this yields,

$$|\Psi_f\rangle \langle \Psi_f| = \frac{1}{2} \hat{I} + \frac{1}{2} \left[\cos\varphi \, \hat{\boldsymbol{m}}_i \sin\varphi \, \hat{\boldsymbol{n}} \times \hat{\boldsymbol{m}}_i + (1 - \cos\varphi) \, (\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{m}}_i) \, \hat{\boldsymbol{n}} \right] \cdot \hat{\boldsymbol{\sigma}}$$

Thus

$$|\Psi_f\rangle \langle \Psi_f| = \frac{1}{2} \hat{I} + \frac{1}{2} \hat{\boldsymbol{m}}_f \cdot \hat{\boldsymbol{\sigma}}$$

which implies that

$$|\Psi_f\rangle \langle \Psi_f| = |+\hat{\boldsymbol{m}}_f\rangle \langle +\hat{\boldsymbol{m}}_f| .$$

The rest of the proof follows by the argument earlier.•

Example: Determine the matrix representing $\hat{R}(\frac{\pi}{2}\boldsymbol{z})$ in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis and apply this to the states $|+\hat{\boldsymbol{z}}\rangle$ and $|+\hat{\boldsymbol{x}}\rangle$.

Answer: The matrix representing $\hat{R}(\frac{\pi}{2}\boldsymbol{z}) = e^{-i\pi/4} \, |+\hat{\boldsymbol{z}}\rangle \langle +\hat{\boldsymbol{z}}| + e^{i\pi/4} \, |-\hat{\boldsymbol{z}}\rangle \langle -\hat{\boldsymbol{z}}|$ can be attained using:

$$|+\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and}$$
$$|-\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Similarly

$$\langle +\hat{\boldsymbol{z}}| \leftrightarrow (1 \;\; 0) \quad \text{and}$$
$$\langle -\hat{\boldsymbol{z}}| \leftrightarrow (0 \;\; 1).$$

Thus

$$|+\hat{\boldsymbol{z}}\rangle \langle +\hat{\boldsymbol{z}}| \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \;\; 0)$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$|-\hat{\boldsymbol{z}}\rangle \langle -\hat{\boldsymbol{z}}| \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \;\; 1)$$
$$= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

These give

$$\hat{R}(\tfrac{\pi}{2}\boldsymbol{z}) \leftrightarrow \begin{pmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{+i\pi/4} \end{pmatrix}.$$

It follows that

$$\begin{aligned} \hat{R}(\tfrac{\pi}{2}\boldsymbol{z})\,|+\hat{\boldsymbol{z}}\rangle &\leftrightarrow \begin{pmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{+i\pi/4} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= e^{-i\pi/4} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned}$$

and thus

$$\hat{R}(\tfrac{\pi}{2}\boldsymbol{z})\,|+\hat{\boldsymbol{z}}\rangle = e^{-i\pi/4}\,|+\hat{\boldsymbol{z}}\rangle\,.$$

This is identical to $|+\hat{\boldsymbol{z}}\rangle$ ignoring the global phase factor $e^{-i\pi/4}$. On the other hand, applying the rotation to $|+\hat{\boldsymbol{x}}\rangle$ gives

$$\begin{aligned} \hat{R}(\tfrac{\pi}{2}\boldsymbol{z})\,|+\hat{\boldsymbol{x}}\rangle &\leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{+i\pi/4} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\pi/4} \\ e^{+i\pi/4} \end{pmatrix} \\ &= e^{-i\pi/4}\,\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{+i\pi/2} \end{pmatrix} \\ &= e^{-i\pi/4}\,\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}. \end{aligned}$$

Thus

$$\hat{R}(\tfrac{\pi}{2}\boldsymbol{z})\,|+\hat{\boldsymbol{x}}\rangle = e^{-i\pi/4}\,|+\hat{\boldsymbol{y}}\rangle$$

which is identical to $|+\hat{\boldsymbol{y}}\rangle$ ignoring the global phase factor $e^{-i\pi/4}$.

4.5 Completeness relation

Any rotation operator must satisfy

$$\hat{R}(0\boldsymbol{n}) = \hat{I}$$

since a rotation through an angle of 0 should leave the physical state unaltered. This implies that for any basis states $\{|+\hat{\boldsymbol{n}}\rangle\,, |-\hat{\boldsymbol{n}}\rangle\}\,,$

$$|+\hat{\boldsymbol{n}}\rangle\,\langle+\hat{\boldsymbol{n}}| + |-\hat{\boldsymbol{n}}\rangle\,\langle-\hat{\boldsymbol{n}}| = \hat{I}.$$

However, we should be able to check independently that this is true. This is guaranteed by the following result.

Theorem (Completeness Relation): For any basis $\{|+\hat{\boldsymbol{n}}\rangle\,, |-\hat{\boldsymbol{n}}\rangle\}\,,$

$$\boxed{|+\hat{\boldsymbol{n}}\rangle\,\langle+\hat{\boldsymbol{n}}| + |-\hat{\boldsymbol{n}}\rangle\,\langle-\hat{\boldsymbol{n}}| = \hat{I}} \tag{II.4.19}$$

Proof: Consider an arbitrary state $|\Psi\rangle$. Then, since the states $\{|+\hat{\boldsymbol{n}}\rangle, |-\hat{\boldsymbol{n}}\rangle\}$ form a basis,

$$|\Psi\rangle = c_+ |+\hat{\boldsymbol{n}}\rangle + c_- |-\hat{\boldsymbol{n}}\rangle .$$

Then

$$\begin{aligned}\left[|+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}| + |-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}|\right]|\Psi\rangle &= \left[|+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}| + |-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}|\right]\left[c_+ |+\hat{\boldsymbol{n}}\rangle + c_- |-\hat{\boldsymbol{n}}\rangle\right] \\ &= c_+ |+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}|+\hat{\boldsymbol{n}}\rangle + c_- |+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}|-\hat{\boldsymbol{n}}\rangle \\ &\quad + c_+ |-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}|+\hat{\boldsymbol{n}}\rangle + c_- |-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}|-\hat{\boldsymbol{n}}\rangle \\ &= c_+ |+\hat{\boldsymbol{n}}\rangle + c_- |-\hat{\boldsymbol{n}}\rangle \\ &= |\Psi\rangle .\end{aligned}$$

Since this is true for any state $|\Psi\rangle$, it follows that $|+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}| + |-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}| = \hat{I}$. •

The completeness relation is useful in constructing representations of kets and operators. For example, for an arbitrary ket using the completeness relation with $\hat{\boldsymbol{n}} = \hat{\boldsymbol{z}}$,

$$\begin{aligned}|\Psi\rangle &= \hat{I}\,|\Psi\rangle \\ &= \left[|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\right]|\Psi\rangle \\ &= |+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}|\Psi\rangle + |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\Psi\rangle \\ &= \langle+\hat{\boldsymbol{z}}|\Psi\rangle\ |+\hat{\boldsymbol{z}}\rangle + \langle-\hat{\boldsymbol{z}}|\Psi\rangle\ |-\hat{\boldsymbol{z}}\rangle\end{aligned}$$

since both $\langle\pm\hat{\boldsymbol{z}}|\Psi\rangle$ are just complex numbers. Thus the representation of $|\Psi\rangle$ in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis is:

$$|\Psi\rangle \leftrightarrow \begin{pmatrix}\langle+\hat{\boldsymbol{z}}|\Psi\rangle \\ \langle-\hat{\boldsymbol{z}}|\Psi\rangle\end{pmatrix}. \tag{II.4.20}$$

This representation is nothing new as we have already encountered the rules for arriving at Eq. (II.4.20). However, using the completeness relation allows a purely algebraic way for reaching this result. In fact, this is often the most useful way in which to change the bases in which kets are represented.

Similarly the completeness relation is useful for representing operators. Here consider an operator $\hat{A}$ acting on $|\Psi\rangle$. Denote

$$|\Phi\rangle := \hat{A}\,|\Psi\rangle$$

Then the vector representing $|\Phi\rangle$ in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis is:

$$|\Phi\rangle \leftrightarrow \begin{pmatrix}\langle+\hat{\boldsymbol{z}}|\Phi\rangle \\ \langle-\hat{\boldsymbol{z}}|\Phi\rangle\end{pmatrix}. \tag{II.4.21}$$

To determine the two entries consider

$$\begin{aligned}\langle +\hat{\boldsymbol{z}}|\Phi\rangle &= \langle +\hat{\boldsymbol{z}}|\,\hat{A}\,|\Psi\rangle \\ &= \langle +\hat{\boldsymbol{z}}|\,\hat{A}\Big[|+\hat{\boldsymbol{z}}\rangle\,\langle +\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\,\langle -\hat{\boldsymbol{z}}|\Big]\,|\Psi\rangle \\ &= \langle +\hat{\boldsymbol{z}}|\,\hat{A}\Big[|+\hat{\boldsymbol{z}}\rangle\,\langle +\hat{\boldsymbol{z}}|\Psi\rangle + |-\hat{\boldsymbol{z}}\rangle\,\langle -\hat{\boldsymbol{z}}|\Psi\rangle\Big] \\ &= \langle +\hat{\boldsymbol{z}}|\,\hat{A}\,|+\hat{\boldsymbol{z}}\rangle\;\langle +\hat{\boldsymbol{z}}|\Psi\rangle + \langle +\hat{\boldsymbol{z}}|\,\hat{A}\,|-\hat{\boldsymbol{z}}\rangle\,\langle -\hat{\boldsymbol{z}}|\Psi\rangle\end{aligned}$$

Similarly

$$\begin{aligned}\langle -\hat{\boldsymbol{z}}|\Phi\rangle &= \langle -\hat{\boldsymbol{z}}|\,\hat{A}\,|\Psi\rangle \\ &= \langle -\hat{\boldsymbol{z}}|\,\hat{A}\Big[|+\hat{\boldsymbol{z}}\rangle\,\langle +\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\,\langle -\hat{\boldsymbol{z}}|\Big]\,|\Psi\rangle \\ &= \langle -\hat{\boldsymbol{z}}|\,\hat{A}\Big[|+\hat{\boldsymbol{z}}\rangle\,\langle +\hat{\boldsymbol{z}}|\Psi\rangle + |-\hat{\boldsymbol{z}}\rangle\,\langle -\hat{\boldsymbol{z}}|\Psi\rangle\Big] \\ &= \langle -\hat{\boldsymbol{z}}|\,\hat{A}\,|+\hat{\boldsymbol{z}}\rangle\;\langle +\hat{\boldsymbol{z}}|\Psi\rangle + \langle -\hat{\boldsymbol{z}}|\,\hat{A}\,|-\hat{\boldsymbol{z}}\rangle\,\langle -\hat{\boldsymbol{z}}|\Psi\rangle\end{aligned}$$

These can be collected to form

$$\begin{pmatrix}\langle +\hat{\boldsymbol{z}}|\Phi\rangle \\ \langle -\hat{\boldsymbol{z}}|\Phi\rangle\end{pmatrix} = \begin{pmatrix}\langle +\hat{\boldsymbol{z}}|\,\hat{A}\,|+\hat{\boldsymbol{z}}\rangle & \langle +\hat{\boldsymbol{z}}|\,\hat{A}\,|-\hat{\boldsymbol{z}}\rangle \\ \langle -\hat{\boldsymbol{z}}|\,\hat{A}\,|+\hat{\boldsymbol{z}}\rangle & \langle -\hat{\boldsymbol{z}}|\,\hat{A}\,|-\hat{\boldsymbol{z}}\rangle\end{pmatrix}\begin{pmatrix}\langle +\hat{\boldsymbol{z}}|\Psi\rangle \\ \langle -\hat{\boldsymbol{z}}|\Psi\rangle\end{pmatrix}$$

Thus the matrix representing $\hat{A}$ in the the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis is:

$$\hat{A} \leftrightarrow \begin{pmatrix}\langle +\hat{\boldsymbol{z}}|\,\hat{A}\,|+\hat{\boldsymbol{z}}\rangle & \langle +\hat{\boldsymbol{z}}|\,\hat{A}\,|-\hat{\boldsymbol{z}}\rangle \\ \langle -\hat{\boldsymbol{z}}|\,\hat{A}\,|+\hat{\boldsymbol{z}}\rangle & \langle -\hat{\boldsymbol{z}}|\,\hat{A}\,|-\hat{\boldsymbol{z}}\rangle\end{pmatrix}.$$

Again the completeness relation provides a purely algebraic method of reaching this result.

4.6 Rotations and unitary transformations for Spin-1/2 particles.

Every rotation operator for a spin-1/2 particle has been shown to be a unitary transformation. Remarkably the converse is nearly true.

Theorem: Let $\hat{U}$ be any unitary transformation on a spin-1/2 particle. Then there exist, an axis of rotation $\hat{\boldsymbol{n}}$, an angle of rotation φ and a real parameter α such that

$$\hat{U} = e^{i\alpha}\,\hat{R}(\varphi\boldsymbol{n}). \tag{II.4.22}$$

Proof: Later, ...•

The term $e^{i\alpha}$ merely contributes a global phase to any state which is produced by an operation of this unitary operator. From a physical perspective this is irrelevant. Thus

> Any unitary operation on a spin-1/2 system is equivalent to some rotation. Every evolution operator on a spin-1/2 system is equivalent to a rotation.

4.7 Generating rotations

For two rotations about the same axis, the definition of the rotation operator via Eq. (II.4.17) implies that

$$\hat{R}(\varphi_2 \boldsymbol{n})\hat{R}(\varphi_1 \boldsymbol{n}) = \hat{R}((\varphi_2 + \varphi_1)\boldsymbol{n}) \tag{II.4.23}$$

where φ_1 and φ_2 are any angles. One way to construct any rotation is as a sequence of rotations, each about the same infinitesimally small angle. The general technique that evolves from this will be essential for determining evolution operators in many physical situations.

The technique is best illustrated by considering the specific example of a rotation through φ about $\hat{\boldsymbol{z}}$. This is equivalent to N successive rotations, each through φ/N about $\hat{\boldsymbol{z}}$. Applying Eq. (II.4.23), gives

$$\hat{R}(\varphi \boldsymbol{z}) = \left[\hat{R}(\tfrac{\varphi}{N}\boldsymbol{z})\right]^N \tag{II.4.24}$$

For any N

$$\hat{R}(\tfrac{\varphi}{N}\boldsymbol{z}) = e^{-i\varphi/2N}\ |+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + e^{i\varphi/2N}\ |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\,.$$

Now consider $N \gg 1$. For convenience let $\Delta\varphi := \varphi/N$ and note that $\Delta\varphi \ll 1$. Then

$$e^{-i\varphi/2N} = 1 - i\frac{\Delta\varphi}{2} + \text{ terms of order } (\Delta\varphi)^2 \text{ or higher.}$$

Ignoring all terms that are of higher order than $\Delta\varphi$

$$\begin{aligned}\hat{R}(\tfrac{\varphi}{N}\boldsymbol{z}) &\approx \left(1 - i\frac{\Delta\varphi}{2}\right)|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + \left(1 + i\frac{\Delta\varphi}{2}\right)|-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| \\ &= |+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| + |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| - i\frac{\varphi}{2N}\left[|+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| - |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}|\right].\end{aligned}$$

Applying the completeness relation gives

$$\hat{R}(\tfrac{\varphi}{N}\boldsymbol{z}) \approx \hat{I} - i\frac{\varphi}{2N}\,\hat{\sigma}_z \tag{II.4.25}$$

where the Pauli z operator is defined to be

$$\hat{\sigma}_z := |+\hat{\boldsymbol{z}}\rangle\langle+\hat{\boldsymbol{z}}| - |-\hat{\boldsymbol{z}}\rangle\langle-\hat{\boldsymbol{z}}| \leftrightarrow \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}. \tag{II.4.26}$$

Thus

$$\hat{R}(\varphi \boldsymbol{z}) \approx \left[\hat{I} - i\frac{\varphi}{2N}\,\hat{\sigma}_z\right]^N. \tag{II.4.27}$$

Taking the limit as $N \to \infty$, this becomes exact

$$\hat{R}(\varphi \boldsymbol{z}) = \lim_{N\to\infty}\left[\hat{I} - i\frac{\varphi}{2N}\,\hat{\sigma}_z\right]^N. \tag{II.4.28}$$

The right hand side is reminiscent of

$$\lim_{N\to\infty}\left(1 - \frac{x}{N}\right)^N = e^{-x}$$

and this suggests that

$$\hat{R}(\varphi \boldsymbol{z}) = e^{-i\varphi\hat{\sigma}_z/2}. \tag{II.4.29}$$

It remains to provide a reasonable definition of the exponential of an operator.

4.7.1 Exponentiating operators

The Taylor series for the exponential of a real number is:

$$e^x = 1 + \frac{1}{1!}\,x + \frac{1}{2!}\,x^2 + \ldots = \sum_{n=0}^{\infty} \frac{1}{n!}\,x^n.$$

With this in mind the exponential of any operator, $\hat{A}$, is defined to be

$$\boxed{e^{\hat{A}} := \sum_{n=0}^{\infty} \frac{1}{n!}\,\hat{A}^n = \hat{I} + \frac{1}{1!}\,\hat{A} + \frac{1}{2!}\,\hat{A}^2 + \ldots.} \tag{II.4.30}$$

This definition in terms of an infinite series may not appear practical for computational purposes. However, as the following example illustrates, it is frequently possible to compute exponentials of operators for spin-1/2.

Example: Let $\hat{A} = i\alpha\hat{\sigma}_z$ where α is real. Determine $e^{\hat{A}}$.

Answer: By the definition, Eq. (II.4.30),

$$e^{\hat{A}} := \sum_{n=0}^{\infty} \frac{1}{n!}\,(\alpha\hat{\sigma}_z)^n = \sum_{n=0}^{\infty} \frac{1}{n!}\alpha^n\hat{\sigma}_z^n.$$

Now

$$\hat{\sigma}_z^2 \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \leftrightarrow \hat{I}.$$

Thus

$$\begin{aligned} e^{i\alpha\hat{\sigma}_z} &= \left[1 - \frac{1}{2!}\,\alpha^2 + \frac{1}{4!}\,\alpha^4 - \ldots\right]\hat{I} + i\left[\alpha - \frac{1}{3!}\,\alpha^3 + \frac{1}{5!}\,\alpha^5 - \ldots\right]\hat{\sigma}_z \\ &= \cos\alpha\,\hat{I} - i\sin\alpha\,\hat{\sigma}_z \\ &\leftrightarrow \cos\alpha\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i\sin\alpha\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\alpha + i\sin\alpha & 0 \\ 0 & \cos\alpha - i\sin\alpha \end{pmatrix} \\ &= \begin{pmatrix} e^{+i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix}. \end{aligned}$$

Exercise: Using the definition of Eq. (II.4.30) and the general definition of a rotation operator, Eq. (II.4.18) show that

$$\hat{R}(\varphi\boldsymbol{z}) = e^{-i\varphi\hat{\sigma}_z/2}.$$

The process of exponentiating operators satisfies several convenient relationships which are comparable to those satisfied by the exponential of numbers.

Theorem: Exponentiation satisfies

$$e^{\alpha\hat{A}}e^{\beta\hat{A}} = e^{(\alpha+\beta)\hat{A}} \tag{II.4.31a}$$

$$\left(e^{\alpha\hat{A}}\right)^{\beta} = e^{\alpha\beta\hat{A}} \tag{II.4.31b}$$

$$\frac{d}{dx}\, e^{x\hat{A}} = \hat{A}\, e^{x\hat{A}} \tag{II.4.31c}$$

for any complex α, β and real x.

Proof: Exercise.•

One word of caution is that for arbitrary operators $\hat{A}$ and $\hat{B}$, in general

$$e^{\hat{A}}e^{\hat{B}} \neq e^{\hat{A}+\hat{B}}$$

although it can be shown that the inequality is replaced by an equality whenever the two operators commute.

4.7.2 Generating arbitrary rotations

The Pauli z operator is said to *generate rotations* about the $\hat{\boldsymbol{z}}$ axis through the process of exponentiation. Thus

$$\hat{R}(\varphi \boldsymbol{z}) = e^{-i\varphi\hat{\sigma}_z/2}.$$

This result can be extended to rotations about arbitrary axes. In general

$$\boxed{\hat{R}(\varphi \boldsymbol{n}) = e^{-i\varphi\hat{\sigma}_n/2}} \tag{II.4.32}$$

where the definition of the Pauli operator for $\hat{\boldsymbol{n}}$ is

$$\boxed{\hat{\sigma}_n := |+\hat{\boldsymbol{n}}\rangle\langle+\hat{\boldsymbol{n}}| - |-\hat{\boldsymbol{n}}\rangle\langle-\hat{\boldsymbol{n}}|} \tag{II.4.33}$$

It would appear that an infinite collection of Pauli operators is required to generate all possible rotations. However, a remarkable fact is that this is not true.

Theorem: If $\hat{\boldsymbol{n}} = n_x\hat{\boldsymbol{x}} + n_y\hat{\boldsymbol{y}} + n_z\hat{\boldsymbol{z}}$ then

$$\boxed{\hat{\sigma}_n = n_x\hat{\sigma}_x + n_y\hat{\sigma}_y + n_z\hat{\sigma}_z} \tag{II.4.34}$$

and

$$\boxed{\hat{R}(\varphi \boldsymbol{n}) = e^{-i\varphi(n_x\hat{\sigma}_x + n_y\hat{\sigma}_y + n_z\hat{\sigma}_z)/2}} \tag{II.4.35}$$

Proof: Later...

Thus only the three Pauli operators, $\hat{\sigma}_x, \hat{\sigma}_y$ and $\hat{\sigma}_z$ are required to generate all rotations. Consequently their matrix representations are important.

The three operators corresponding $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ are represented in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis by

$$\begin{aligned} \hat{\sigma}_x &\leftrightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \hat{\sigma}_y &\leftrightarrow \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \hat{\sigma}_z &\leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{II.4.36}$$

4.8 Time dependent evolution of spin-1/2 particles

For a spinning charged particle, the spin angular momentum is proportional to the magnetic dipole moment. An external magnetic field interacts with a magnetic dipole and typically alter its orientation, thus changing the particle's spin angular momentum. At the quantum mechanical level this is true whether the particle is charged or not; the spin is an intrinsic quantity which refers both to angular momentum and the behavior of the particle in the presence of external magnetic fields. The aim now is to describe the effects of a given external magnetic field on a spin-1/2 particle.

Suppose that the particle enters the magnetic field at time $t = 0$ in a known state $|\Psi_i\rangle$. As the particle passes through the magnetic field, its state will evolve and, in general will depend on time, t. Thus the state will be denoted $|\Psi(t)\rangle$. Here t appears inside the ket label, which can be taken to change as time passes. In the case of spin-1/2 particle the state can always be labeled in terms of a unit vector $\hat{\boldsymbol{n}}$, and it is this unit vector label which changes with time. It follows that we will have to describe time-dependent kets. At any instant t the ket can still be expressed in terms of the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis and thus

$$\begin{aligned} |\Psi(t)\rangle &= c_+(t)\,|+\hat{\boldsymbol{z}}\rangle + c_-(t)\,|-\hat{\boldsymbol{z}}\rangle \\ &\leftrightarrow \begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix}. \end{aligned}$$

Note that the kets $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ are independent of time while the coefficients $c_+(t)$ and $c_-(t)$ do depend on time. In effect, this yields a time dependent column vector. This is consistent with the physical interpretation of the states $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ as those which yield definite outcomes for measurements of S_z. We do not regard the z direction as time dependent and thus the process of measuring S_z is time-independent. The two possible outcomes are not time dependent as they always remain $S_z = \pm\hbar/2$. However, the probability with

which either of these is attained does depend on the state and hence t and this is reflected in the time dependence of $c+$ and c_-. Note that the state at $t = 0$ is

$$|\Psi(0)\rangle = |\Psi_i\rangle .$$

The state at any later time is always related to that at an initial time via a unitary evolution operator and it follows that this must be time-dependent. Thus

$$|\Psi(t)\rangle = \hat{U}(t)\,|\Psi(0)\rangle$$

and, in terms of representations in the z basis,

$$\begin{pmatrix} c_+(t) \\ c_-(t) \end{pmatrix} = \begin{pmatrix} U_{++}(t) & U_{+-}(t) \\ U_{-+}(t) & U_{--}(t) \end{pmatrix} \begin{pmatrix} c_+(0) \\ c_-(0) \end{pmatrix}$$

where $U_{ij}(t)$ (with $i, j = \pm$)are complex valued functions of t.

Example: Consider a situation where the evolution operator is a rotation about the z axis through a time dependent angle ωt where ω is a constant. Determine the matrix that represents the operator $\hat{R}((\omega t)\boldsymbol{z})$. Suppose that the initial state of the spin-1/2 particle is $|+\hat{\boldsymbol{x}}\rangle$. Determine the state at time t and describe the direction of the component of spin whose measurement will yield $+\hbar/2$ with certainty

Answer: The matrix representation in the $\{|+\hat{\boldsymbol{z}}\rangle , |-\hat{\boldsymbol{z}}\rangle\}$ basis of a general rotation about the z axis is

$$\hat{R}(\varphi\boldsymbol{z}) \leftrightarrow \begin{pmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{pmatrix}$$

and thus

$$\hat{R}((\omega t)\boldsymbol{z}) \leftrightarrow \begin{pmatrix} e^{-i\omega t/2} & 0 \\ 0 & e^{i\omega t/2} \end{pmatrix} .$$

Now the state of the particle at $t = 0$ is

$$|\Psi(0)\rangle = |+\hat{\boldsymbol{x}}\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

Thus at a later time t the state is

$$\begin{aligned} |\Psi(t)\rangle &= \hat{R}((\omega t)\boldsymbol{z})\,|\Psi(0)\rangle \\ &\leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\omega t/2} & 0 \\ 0 & e^{i\omega t/2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\omega t/2} \\ e^{i\omega t/2} \end{pmatrix} \\ &\leftrightarrow \frac{1}{\sqrt{2}} e^{-i\omega t/2}\,|+\hat{\boldsymbol{z}}\rangle + \frac{1}{\sqrt{2}} e^{i\omega t/2}\,|-\hat{\boldsymbol{z}}\rangle . \end{aligned}$$

This gives

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} e^{-i\omega t/2} |+\hat{\boldsymbol{z}}\rangle \frac{1}{\sqrt{2}} e^{i\omega t/2} |-\hat{\boldsymbol{z}}\rangle$$
$$= e^{-i\omega t/2} \Big[\frac{1}{\sqrt{2}} |+\hat{\boldsymbol{z}}\rangle \frac{1}{\sqrt{2}} e^{i\omega t} |-\hat{\boldsymbol{z}}\rangle \Big]$$

and ignoring the global phase factor $e^{-i\omega t/2}$ the state of the particle at time t is equivalent to

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}} |+\hat{\boldsymbol{z}}\rangle \frac{1}{\sqrt{2}} e^{i\omega t} |-\hat{\boldsymbol{z}}\rangle .$$

This is the same as the state $|+\hat{\boldsymbol{n}}\rangle$ where $\hat{\boldsymbol{n}}$ has spherical coordinates $\theta = \pi/2$ and $\phi = \omega t$. The time dependence of the label $\hat{\boldsymbol{n}}$ is clear in this case. Note that this vector lies in the xy plane and precesses about the z axis with an angular frequency ω.

The remaining step is to relate the time evolution operator to observables associated with quantities such as external magnetic fields. The idea will be to consider derivatives of the state. Clearly the notion of differentiating a time-dependent ket will be crucial. This is defined via

$$|\Psi(t)\rangle = c_+(t) |+\hat{\boldsymbol{z}}\rangle + c_-(t) |-\hat{\boldsymbol{z}}\rangle \Rightarrow \frac{d}{dt} |\Psi(t)\rangle = \frac{dc_+}{dt} |+\hat{\boldsymbol{z}}\rangle + \frac{dc_-}{dt} |-\hat{\boldsymbol{z}}\rangle . \tag{II.4.37}$$

In terms of representation in the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis,

$$\frac{d}{dt} |\Psi(t)\rangle \leftrightarrow \begin{pmatrix} \dfrac{dc_+}{dt} \\ \dfrac{dc_-}{dt} \end{pmatrix} \equiv \frac{d}{dt} \begin{pmatrix} c_+ \\ c_- \end{pmatrix} . \tag{II.4.38}$$

Then

$$\frac{d}{dt} |\Psi(t)\rangle = \frac{d}{dt} \Big[\hat{U}(t) |\Psi(0)\rangle \Big]$$
$$= \left[\frac{d}{dt} \hat{U}(t) \right] |\Psi(0)\rangle$$
$$= \left[\frac{d}{dt} \hat{U}(t) \right] \hat{U}(t)^{-1} |\Psi(t)\rangle$$

where $\hat{U}(t)^{-1}$ is the inverse operator (or matrix) to $\hat{U}(t)$. It is straightforward to show that the operator $\hat{A} := \left[\frac{d}{dt} \hat{U}(t) \right] \hat{U}(t)^{-1}$ is anti-Hermitian, i.e. $\hat{A}^\dagger = -\hat{A}$. Thus the operator $i \left[\frac{d}{dt} \hat{U}(t) \right] \hat{U}(t)^{-1}$ is Hermitian. Thus means that

$$i \frac{d}{dt} |\Psi(t)\rangle = i \frac{d}{dt} \Big[\hat{U}(t) |\Psi(0)\rangle \Big]$$
$$= i \frac{d}{dt} \hat{U}(t) |\Psi(0)\rangle$$
$$= i \left[\frac{d}{dt} \hat{U}(t) \right] \hat{U}(t)^{-1} |\Psi(t)\rangle \tag{II.4.39}$$

and, since $i\left[\frac{d}{dt}\hat{U}(t)\right]\,\hat{U}(t)^{-1}$ is Hermitian and therefore an observable, it is reasonable to ask which observable this is. Considerations of energy conservation lead to the conclusion that a plausible choice is that this is related to the energy or equivalently that the observable is the Hamiltonian. This motivates the main rule which connects the evolution of quantum systems to factors such as magnetic fields that drive the evolution.

The time evolution of a quantum mechanical system that is isolated from other quantum mechanical systems is dictated by

$$i\hbar\frac{d}{dt}\left|\Psi(t)\right\rangle = \hat{H}\left|\Psi(t)\right\rangle \tag{II.4.40}$$

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

Eq. (II.4.40) is known as the **Schrödinger equation.** It's usefulnesses here is that it frequently straightforward to construct the Hamiltonian given a knowledge of the physical circumstances and then to solve the Schrödinger equation to determine the evolution operator, $\hat{U}(t)$. It is seldom straightforward to arrive at the evolution operator without first passing through this step.

Example: For a spin-1/2 particle of mass m, charge q and g-factor, g in a general, possibly time-dependent, magnetic field

$$\mathbf{B}(t) = B_x(t)\,\hat{\boldsymbol{x}} + B_y(t)\,\hat{\boldsymbol{y}} + B_z(t)\,\hat{\boldsymbol{z}}$$

the Hamiltonian is

$$\begin{aligned}\hat{H} &= -\frac{gq}{2m}\left[B_x(t)\,\hat{S}_x + B_z(t)\,\hat{S}_y + B_z(t)\,\hat{S}_z\right]\\ &= -\frac{gq}{2m}\,\frac{\hbar}{2}\left[B_x(t)\,\hat{\sigma}_x + B_y(t)\,\hat{\sigma}_y + B_z(t)\,\hat{\sigma}_z\right]\\ &\leftrightarrow -\frac{gq}{2m}\,\frac{\hbar}{2}\begin{pmatrix} B_z(t) & B_x(T) - iB_y(t)\\ B_x(T) + iB_y(t) & B_z(t)\end{pmatrix}.\end{aligned}$$

The Schrödinger equation is represented in terms of vectors and matrices by

$$i\hbar\frac{d}{dt}\begin{pmatrix}c_+\\ c_-\end{pmatrix} = -\frac{gq}{2m}\,\frac{\hbar}{2}\begin{pmatrix} B_z(t) & B_x(T) - iB_y(t)\\ B_x(T) + iB_y(t) & B_z(t)\end{pmatrix}\begin{pmatrix}c_+\\ c_-\end{pmatrix}.$$

Upon expansion this can be seen to yield two coupled first order differential equations for $c_\pm(t)$.

4.8.1 Solving the Schrödinger equation for time-independent Hamiltonians

In general the coupled differential equations yielded by the Schrödinger equation are too complicated to solve in closed form for the evolution operator. However, for time-independent Hamiltonians there is a general solution that is straightforward.

Theorem: If the Hamiltonian $\hat{H}$ is time-independent then,

$$\boxed{\hat{U}(t) = e^{-i\hat{H}t/\hbar}.} \tag{II.4.41}$$

Proof: By Eqs. (II.4.39) and (II.4.40), the evolution operator satisfies

$$i\hbar \left[\frac{d}{dt}\hat{U}(t)\right] \hat{U}(t)^{-1} = \hat{H},$$

or, equivalently, it satisfies the differential equation

$$\frac{d}{dt}\hat{U}(t) = \frac{-i}{\hbar}\hat{H}\hat{U}(t).$$

Additionally $\hat{U}(0) = \hat{I}$. Now consider the possible solution

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar}.$$

Then,

$$\begin{aligned}\frac{d}{dt}\hat{U}(t) &= \frac{d}{dt}e^{-i\hat{H}t/\hbar} \\ &= \left(\frac{-i}{\hbar}\,\hat{H}\right) e^{-i\hat{H}t/\hbar} \\ &= \frac{-i}{\hbar}\,\hat{H}\hat{U}(t).\end{aligned}$$

Additionally

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar} \Rightarrow \hat{U}(0) = \hat{I}.$$

Thus

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar}$$

satisfies the differential equation for the evolution operator and gives the correct initial value. There is a unique solution to this type of first order differential equation with a given initial value. •

Example: Suppose that a spin-1/2 particle of mass m, charge q and g-factor, g in the magnetic field

$$\mathbf{B} = B_0\hat{\boldsymbol{z}}$$

Determine the evolution operator and show that it is a rotation. Find the axis and angle of rotation.

Answer: First, the Hamiltonian is

$$\hat{H} = -\frac{gq}{2m}\frac{\hbar}{2}B\hat{\sigma}_z.$$

For convenience, define

$$\omega := -\frac{gq}{2m}B.$$

Then Eq. (II.4.41) implies

$$\begin{aligned}\hat{U}(t) &= e^{-i\hat{H}t/\hbar} \\ &= e^{-i\omega\hat{\sigma}_z\hbar t/\hbar 2} \\ &= e^{-i\omega\hat{\sigma}_z t/2}.\end{aligned}$$

This is a rotation about the $\hat{\boldsymbol{z}}$ axis through angle ωt.

This extends quite generally to any constant magnetic field. Consider a field of magnitude B oriented along $\hat{\boldsymbol{n}} = n_x\hat{\boldsymbol{x}} + n_y\hat{\boldsymbol{y}} + n_z\hat{\boldsymbol{z}}$. Then the field is $\mathbf{B} = B\left(n_x\hat{\boldsymbol{x}} + n_y\hat{\boldsymbol{y}} + n_z\hat{\boldsymbol{z}}\right)$. The Hamiltonian is

$$\begin{aligned}\hat{H} &= \frac{\hbar\omega}{2}\left[n_x\hat{\sigma}_x + n_y\hat{\sigma}_y + n_z\hat{\sigma}_z\right] \\ &= \frac{\hbar\omega}{2}\,\hat{\sigma}_n\end{aligned}$$

where $\omega = -gqB/2m$ for a particle of mass m, charge q and g-factor, g. Thus the evolution operator is

$$\hat{U}(t) = e^{-i\omega\hat{\sigma}_n t/2}$$

which is a rotation about axis $\hat{\boldsymbol{n}}$ through angle ωt. We have thus arrived at:

> For a spin-1/2 particle of mass m, charge q and g-factor, g in a constant magnetic field of strength B in direction $\hat{\boldsymbol{n}}$ the evolution operator is a rotation about $\hat{\boldsymbol{n}}$ through angle ωt where $\omega = -gqB/2m$.

Since every time-independent Hamiltonian has the form of that generated by a constant magnetic field plus a possible multiple of the identity operator (which is irrelevant for evolution) we have classified every possible type of evolution generated by time independent Hamiltonians in terms of rotations. In fact, this extends to neutral spin-1/2 particles where $\omega = \gamma B$ where γ is called a gyromagnetic ratio and is a property of a particle; neutrons are an example. These notions of evolution provide one of the basic concepts for nuclear magnetic resonance (NMR). In typical NMR scenarios there is an ensemble of particles in various quantum states. The simplest case is that where each particle in the ensemble contains a single spin-1/2 particle. It can be show that for this, the state of the ensemble is

equivalent to a smaller ensemble each of whose members (particles) is, at any stage in the same state, $|+\hat{\boldsymbol{n}}\rangle$, for some direction $\hat{\boldsymbol{n}}$. An NMR spectrometer's measurement outputs yield good approximations the expectation values $\langle S_x\rangle, \langle S_y\rangle$ and $\langle S_z\rangle$ and these are gathered into an ordinary three dimensional magnetization vector $\mathbf{M} := \langle S_x\rangle \hat{\boldsymbol{x}} + \langle S_y\rangle \hat{\boldsymbol{y}} + \langle S_z\rangle \hat{\boldsymbol{z}}$. It can then be shown that for an ensemble, each of whose members is in the state $|+\hat{\boldsymbol{n}}\rangle$, the magnetization vector is $\mathbf{M} := \alpha\hat{\boldsymbol{n}}$ where α is a constant. Thus if the spin-1/2 particle is placed in a constant magnetic field $\mathbf{B} = B\hat{\boldsymbol{k}}$ along some direction $\hat{\boldsymbol{k}}$, the state rotates about this axis at a constant rate with an angular frequency ω where $\omega = \gamma B$. It follows that the magnetization vector also rotates about this axis with an angular frequency ω. This essentially produces a time varying magnetic field which can induce currents in coils situated around the spin-1/2 particles. This is the quantum mechanical manifestation of *Larmor precession* and the frequency with which it occurs is called the *Larmor frequency.*

4.8.2 Evolution of energy eigenstates

The evolution of energy eigenstates occupies a special place in the dynamics of quantum systems, by virtue of the fact that their time evolution is particularly simple. Additionally this is of importance, since any state at an initial instant can be expressed as

$$|\Psi_i\rangle = c_1 |\phi_1\rangle + c_2 |\phi_2\rangle$$

where $|\phi_1\rangle$ and $|\phi_2\rangle$ are the energy eigenstates and c_1 and c_2 complex coefficients. Thus

$$\begin{aligned} |\Psi_f\rangle &= \hat{U} |\Psi_i\rangle \\ &= \hat{U}\Big[c_1 |\phi_1\rangle + c_2 |\phi_2\rangle\Big] \\ &= c_1\hat{U} |\phi_1\rangle + c_2\hat{U} |\phi_2\rangle \end{aligned} \tag{II.4.42}$$

since the evolution operator is linear. Thus the evolution of the state will be completely determined by the evolution of the energy eigenstates.

As an example consider a spin-1/2 particle in a magnetic field $\mathbf{B} = B_0\hat{\boldsymbol{z}}$. The Hamiltonian is

$$\hat{H} = \frac{\hbar\omega}{2}\,\hat{\sigma}_z.$$

where $\omega := -gqB_0/2m$. The energies and energy eigenstates of the Hamiltonian are given in Table II.4.1. Suppose that the particle is initially ($t = 0$) in the state $|\phi_1\rangle = |+\hat{\boldsymbol{z}}\rangle$. Then

$$|\Psi(0)\rangle = |+\hat{\boldsymbol{z}}\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The state at a later time t is

$$\begin{aligned} |\Psi(t)\rangle &= \hat{U} |\Psi(0)\rangle \\ &= \hat{R}((\omega t)\boldsymbol{z}) |\Psi(0)\rangle \\ &\leftrightarrow \begin{pmatrix} e^{-i\omega t/2} & 0 \\ 0 & e^{i\omega t/2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} e^{-i\omega t/2} \\ 0 \end{pmatrix} \\ &\leftrightarrow e^{-i\omega t/2} |+\hat{\boldsymbol{z}}\rangle. \end{aligned}$$

Outcome (E)	Associated State
$E_1 = \dfrac{\hbar\omega}{2}$	$\lvert\phi_1\rangle = \lvert+\hat{\boldsymbol{z}}\rangle$
$E_2 = -\dfrac{\hbar\omega}{2}$	$\lvert\phi_2\rangle = \lvert+\hat{\boldsymbol{z}}\rangle$

Table II.4.1: Table of measurement outcomes and states for measurement of the energy of a spin-1/2 particle with Hamiltonian $\hat{H} = \frac{\hbar\omega}{2}\hat{\sigma}_z$.

Thus the time evolution of this energy eigenstate is:

$$|\phi_1\rangle \;\rightarrow\; |\Psi(t)\rangle = e^{-iE_1 t/\hbar}\,|\phi_1\rangle\,.$$

Exercise: Show that the time evolution of the other energy eigenstate of a spin-1/2 particle with Hamiltonian $\hat{H} = \frac{\hbar\omega}{2}\hat{\sigma}_z$ is

$$|\phi_2\rangle \;\rightarrow\; |\Psi(t)\rangle = e^{-iE_2 t/\hbar}\,|\phi_2\rangle\,.$$

The results derived for this example are, in fact, completely general.

Theorem (Eigenstate Evolution): Suppose that $|\phi_i\rangle$ is an energy eigenstate of a time-independent Hamiltonian and the associated energy eigenvalue is E_i. Then the eigenstate evolves as

$$\boxed{|\phi_i\rangle \;\rightarrow\; |\Psi(t)\rangle = e^{-iE_i t/\hbar}\,|\phi_i\rangle\,.} \tag{II.4.43}$$

Proof: Exercise.

For a particle initially in an eigenstate, the state at a later time is identical except for a phase factor, $e^{-iE_i t/\hbar}$. In this case, this is a global phase and can be ignored. Thus a system which is initially in a given energy eigenstate is subsequently in the same energy eigenstate. In terms of measurement outcomes this means that the energy measurements will yield the same outcome (energy eigenvalue) regardless of the time at which the measurement is performed. Thus the energy is conserved, a reasonable criterion for any quantum system which is isolated from other quantum systems and only subjected to constant outside influences.

An immediate and general consequence of Eqs. (II.4.42) and (II.4.43) is that a general state evolves as

$$c_1\,|\phi_1\rangle + c_2\,|\phi_2\rangle \rightarrow c_1 e^{-iE_1 t/\hbar}\,|\phi_1\rangle + c_2 e^{-iE_2 t/\hbar}\,|\phi_2\rangle\,. \tag{II.4.44}$$

4.9 Evolution of Expectation Values

An ensemble of particles all initially in the same state initially ($t = 0$), $|\Psi(0)\rangle$ and subject to the same Hamiltonian, is such that each will, at a later time be in the same state

$$|\Psi(t)\rangle = \hat{U}(t)\ |\Psi(0)\rangle\,.$$

Suppose that the physical quantity A is measured at the same time on each of the particles. The sample average of the outcomes will, if the ensemble size is large enough, approximate the expectation value $\langle A\rangle$ well. However, both of these will depend on t and it is of interest to determine the evolution of this with respect to time. This is given by the following theorem.

Theorem (Ehrenfest's Theorem): Consider an ensemble of identical particles, each subject to the same Hamiltonian, $\hat{H}$. The expectation vale of any observable $\hat{A}$ satisfies

$$\boxed{\frac{d}{dt}\langle A\rangle = \langle\Psi(t)|\,\frac{\partial\hat{A}}{\partial t}\,|\Psi(t)\rangle + \frac{i}{\hbar}\,\langle\Psi(t)|\,\left[\hat{H},\hat{A}\right]|\Psi(t)\rangle} \qquad \text{(II.4.45)}$$

where $\dfrac{\partial\hat{A}}{\partial t}$ accounts for the fact that the observable may depend on time explicitly.

Proof: Starting with the left

$$\begin{aligned}\frac{d}{dt}\langle A\rangle &= \frac{d}{dt}\left(\langle\Psi(t)|\,\hat{A}\,|\Psi(t)\rangle\right)\\ &= \frac{d}{dt}\langle\Psi(t)|\left(\hat{A}\,|\Psi(t)\rangle\right) + \langle\Psi(t)|\,\frac{\partial\hat{A}}{\partial t}\,|\Psi(t)\rangle + \langle\Psi(t)|\,\hat{A}\frac{d}{dt}\,|\Psi(t)\rangle\,.\end{aligned}$$

However, the Schrödinger equation implies that

$$\frac{d}{dt}\,|\Psi(t)\rangle = -\frac{i}{\hbar}\,\hat{H}\,|\Psi(t)\rangle$$

and this gives

$$\frac{d}{dt}\,\langle\Psi(t)| = \frac{i}{\hbar}\,\langle\Psi(t)|\,\hat{H}.$$

Thus

$$\begin{aligned}\frac{d}{dt}\langle A\rangle &= \frac{d}{dt}\,\langle\Psi(t)|\left(\hat{A}\,|\Psi(t)\rangle\right) + \langle\Psi(t)|\,\frac{\partial\hat{A}}{\partial t}\,|\Psi(t)\rangle + \langle\Psi(t)|\,\hat{A}\frac{d}{dt}\,|\Psi(t)\rangle\\ &= \langle\Psi(t)|\,\frac{\partial\hat{A}}{\partial t}\,|\Psi(t)\rangle + \frac{i}{\hbar}\,\langle\Psi(t)|\,\hat{H}\hat{A}\,|\Psi(t)\rangle - \frac{i}{\hbar}\,\langle\Psi(t)|\,\hat{A}\hat{H}\,|\Psi(t)\rangle\\ &= \langle\Psi(t)|\,\frac{\partial\hat{A}}{\partial t}\,|\Psi(t)\rangle + \frac{i}{\hbar}\,\langle\Psi(t)|\,\left[\hat{H},\hat{A}\right]|\Psi(t)\rangle\,.\end{aligned}$$

and this completes the proof. •

Example: Consider an ensemble of spin-1/2 particles, each subject to the magnetic field $\mathbf{B} = B\hat{\boldsymbol{x}}$. Denote the resulting Hamiltonian

$$\hat{H} = \frac{\hbar\omega}{2}\,\hat{\sigma}_x.$$

Determine equations for the time evolution of $\langle S_x \rangle$ and $\langle S_z \rangle$.

Answer: First

$$\hat{S}_x = \frac{\hbar}{2}\,\hat{\sigma}_x$$

and clearly this is independent of time. By Ehrenfest's theorem,

$$\begin{aligned}
\frac{d}{dt}\langle S_x \rangle &= \frac{i}{\hbar}\,\langle \Psi(t)|\left[\hat{H}, \hat{S}_x\right]|\Psi(t)\rangle \\
&= \frac{i}{\hbar}\frac{\hbar}{2}\frac{\hbar\omega}{2}\,\langle \Psi(t)|\left[\hat{\sigma}_x, \hat{\sigma}_x\right]|\Psi(t)\rangle \\
&= 0.
\end{aligned}$$

Similarly

$$\begin{aligned}
\frac{d}{dt}\langle S_z \rangle &= \frac{i}{\hbar}\,\langle \Psi(t)|\left[\hat{H}, \hat{S}_z\right]|\Psi(t)\rangle \\
&= \frac{i}{\hbar}\frac{\hbar}{2}\frac{\hbar\omega}{2}\,\langle \Psi(t)|\left[\hat{\sigma}_x, \hat{\sigma}_z\right]|\Psi(t)\rangle \\
&= \frac{i\hbar\omega}{4}\,\langle \Psi(t)|\left[\hat{\sigma}_x, \hat{\sigma}_z\right]|\Psi(t)\rangle\,.
\end{aligned}$$

But

$$\left[\hat{\sigma}_x, \hat{\sigma}_z\right] = -2i\hat{\sigma}_y.$$

Thus

$$\begin{aligned}
\frac{d}{dt}\langle S_z \rangle &= -2i\frac{i\hbar\omega}{4}\,\langle \Psi(t)|\,\hat{\sigma}_y\,|\Psi(t)\rangle \\
&= \omega\,\langle \Psi(t)|\,\frac{\hbar}{2}\,\hat{\sigma}_y\,|\Psi(t)\rangle \\
&= \omega\,\langle \Psi(t)|\,\hat{S}_y\,|\Psi(t)\rangle \\
&= \omega\,\langle S_y \rangle\,.
\end{aligned}$$

It can be shown that the differential equations satisfied by the expectation values of the spin observables of a spin-1/2 particle are the same as those satisfied by the components of the spin of a classical dipole in a magnetic field. Ehrenfest's theorem frequently leads to such conclusions.

4.10 Application : Neutrino Oscillations

Neutrinos are subatomic particles that are involved in nuclear reactions driven by the weak interaction. Prominent examples include beta decay, in which a neutron decays into an

electron and an electron antineutrino,

$$n \rightarrow p + e^- + \overline{\nu_e}$$

or inverse beta decay, which involves an electron neutrino,

$$p \rightarrow n + e^+ + \nu_e.$$

Nuclear reactions of this type involving protons and neutrons invariably also involve electrons or positrons and the electron neutrino or antineutrino. However, there are other reactions driven by the weak interaction, which involve none of these. For example the decay of the pion π^+,

$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$

results in a neutrino that is not observed in weak interaction processes involving electrons. Originally this was regarded as a distinct type of neutrino. The two types of neutrinos could be referred to as electron and muon neutrinos respectively.

A variety of experimental evidence has suggested that these two types of neutrino may be two different manifestations of a single type of neutrino; these are called neutrino flavors. A neutrino with electron flavor will undergo nuclear reactions involving protons, neutrons and electrons while that with muon flavor will undergo nuclear reactions involving muons an pions. If the two flavors are associated with a single particle type it is not unreasonable to assume that a neutrino of one flavor can change into one of another flavor. A quantum mechanical description of neutrinos that describes possible changes of flavor commences with two states associated with the flavors: $|\nu_e\rangle$ and $|\nu_\mu\rangle$. The physical meaning of these states can be described in terms of a neutrino flavor measurement, for which a schematic diagram is provided in Fig. II.4.6. The state $|\nu_e\rangle$ means that a neutrino flavor measurement

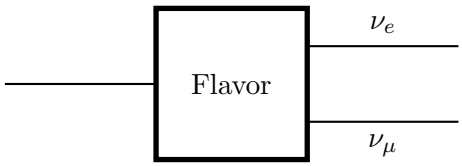


Figure II.4.6: Schematic diagram of a neutrino flavor measurement, in terms analogous to Stern-Gerlach measurements. Neutrinos are fired into the apparatus from the left. The trajectories on the right corresponding to each of the two flavors. Note that in real experiments the neutrinos generally will not persist after detection. However the apparatus can be arranged so that it will determine the type and eject a new neutrino of the same type in the correct output beam on the right.

will yield ν_e with certainty. Similarly $|\nu_\mu\rangle$ means that a neutrino flavor measurement will yield ν_μ with certainty. These measurement outcomes are mutually exclusive and thus the

two flavor states must be orthonormal. Thus

$$\langle \nu_e | \nu_e \rangle = 1 \tag{II.4.46a}$$
$$\langle \nu_e | \nu_\mu \rangle = 0 \tag{II.4.46b}$$
$$\langle \nu_\mu | \nu_e \rangle = 0 \tag{II.4.46c}$$
$$\langle \nu_\mu | \nu_\mu \rangle = 1. \tag{II.4.46d}$$

The general neutrino state can be expressed as a superposition

$$|\Psi\rangle = c_e \, |\nu_e\rangle + c_\mu \, |\nu_\mu\rangle \tag{II.4.47}$$

where c_e and c_μ satisfy

$$|c_e|^2 + |c_\mu|^2 = 1.$$

An interesting possibility is that there are two neutrino energy eigenstates, which are not necessarily the same as the flavor states. These will be denoted by the states $|\nu_I\rangle$ and $|\nu_{II}\rangle$; the corresponding energies are denoted E_I and E_{II}. It is always possible to express these in terms of the flavor states. A fairly general form relationship between the two types of states is

$$|\nu_I\rangle = \cos\theta \; |\nu_e\rangle - \sin\theta \; |\nu_\mu\rangle \tag{II.4.48a}$$
$$|\nu_{II}\rangle = \sin\theta \; |\nu_e\rangle + \cos\theta \; |\nu_\mu\rangle \, . \tag{II.4.48b}$$

Exercise: Show that the energy eigenstates given by Eqs. (II.4.48) are orthonormal.

Now consider the possibility that a neutrino is initially ($t = 0$) in the muon flavor state $|\Psi(0)\rangle = |\nu_\mu\rangle$. Then the standard rules for time evolution imply that at a later time t,

$$\begin{aligned} |\Psi(t)\rangle &= -\sin\theta \, e^{-iE_I t/\hbar} \; |\nu_I\rangle + \cos\theta \, e^{-iE_{II} t/\hbar} \; |\nu_{II}\rangle \\ &= \sin\theta\cos\theta \left(e^{-iE_{II} t/\hbar} - e^{-iE_I t/\hbar} \right) |\nu_e\rangle + \left(\sin^2\theta e^{-iE_I t/\hbar} + \cos^2\theta e^{-iE_{II} t/\hbar} \right) |\nu_\mu\rangle \, . \end{aligned} \tag{II.4.49}$$

Exercise: Prove Eq. (II.4.49).

It follows that the probability with which a flavor measurement gives ν_e is

$$\Pr(\nu_e) = \sin^2(2\theta) \sin^2[(E_{II} - E_I)t/2\hbar]. \tag{II.4.50}$$

In this sense, the flavor of the neutrino will oscillate provided that there are two distinct energy eigenstates. There is strong experimental evidence for the existence of such neutrino oscillations. A striking consequence of this is that the neutrino must have non-zero mass. This stems from the fact that the relativistic energy of a particle with mass m and momentum p is $E = \sqrt{m^2c^4 + p^2c^2}$ and that distinct energy levels indicate distinct masses. [1]

[1] A detailed derivation of this is provided in Sassaroli, Am. J. Phys, **67**, pg 869-875 (1999). Other useful review articles on neutrino oscillations are: Waltham, Am. J. Phys, **72**, pg 742-752 (2004) and Wolfenstein Rev. Mod. Phys, **71**, pg 140 (1999)

III. General Framework for Quantum Mechanics

1 Kinematics: States and Measurements

1.1 States and Measurements

The basic mathematical entity which represents the physical state of a quantum mechanical system is represented mathematically by a ket, generically denoted $|\Psi\rangle$.

> The possible states of a quantum mechanical system $\{|\Psi\rangle|$ all labels $\Psi\}$ are elements of a Hilbert space, i.e. a complex vector space which is equipped with an inner product.

The physical meaning of any state, $|\Psi\rangle$, is that there is physical quantity A which, when measured for a particle in this state, will yield just one particular outcome with certainty. A schematic diagram of this is provided in Fig. III.1.1. In general there is no single quantity

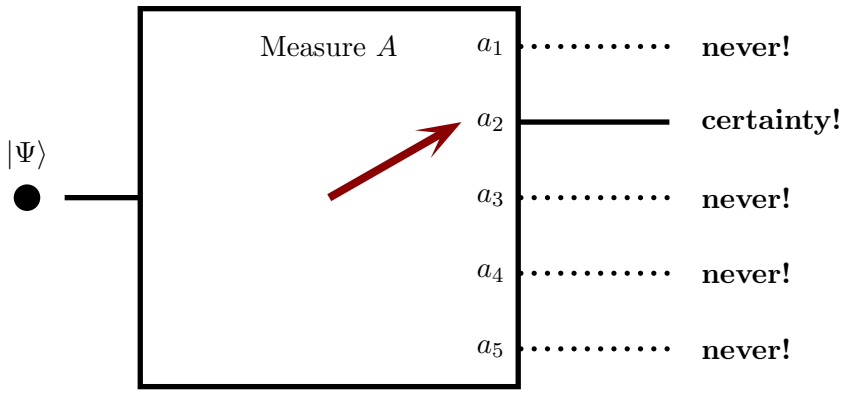


Figure III.1.1: Schematic of a generic measurement. The device measures the physical quantity A. The possible outcomes are $a_1, a_2, \ldots$. A particle in the state $|\Psi\rangle$ yields only one possible outcome upon measurement of A.

whose measurement will yield outcomes with certainty for all states. Thus given particles in states $|\Psi\rangle$ and $|\Phi\rangle$, measurement of one physical quantity A may yield one outcome with certainty for the particle in state $|\Psi\rangle$ but there will be more than one possible outcome for a particle in state $|\Phi\rangle$ for such a measurement. In this case there will be another, distinct quantity B which will yield one outcome with certainty for a particle in state $|\Phi\rangle$.

Example: For spin-1/2 particles, two possible states are $|\Psi\rangle := |+\hat{\boldsymbol{z}}\rangle$ and $|\Phi\rangle := |+\hat{\boldsymbol{x}}\rangle$. Thus measuring S_z yields one outcome with certainty ($S_z = +\hbar/2$) for a particle in the state $|\Psi\rangle$ but, for a particle in state $|\Phi\rangle$, both outcomes of a measurement of S_z are equally likely. However, a measurement of S_x on a particle in the state $|\Phi\rangle$ yields one outcome with certainty ($S_x = +\hbar/2$).

The structure of the vector space of kets is determined via the following edicts based on measurement outcomes. First, the dimension is determined by the following.

> The dimension of the vector space equals the number of distinct, mutually incompatible outcomes of any single maximal measurement, where a maximal measurement means one with the largest number of distinct outcomes.

A side note is that these measurements must be *repeatable*. Thus for a measurement of A which yields a_2 in one instance, an immediate repetition of the same measurement will again yield a_2.

Example: For a spin-1/2 system, it is an experimentally verified fact that measurements of any single component of spin S_n are repeatable but a combined measurement consisting of measuring S_n followed by S_m (where $\hat{\boldsymbol{m}}$ and $\hat{\boldsymbol{n}}$ are not parallel or opposite) is not repeatable. Additionally a measurement of S_n yields one of two possible outcomes. Thus for a spin-1/2 system, the dimension of the vector space is 2.

Similarly for a spin-1 system measurements of and component of spin S_n yield one of three possible outcomes and the dimension of the vector space is 3.

For repeatable measurements of a quantity A, if the measurement yields outcome a_j in one instance, then it will yield the same outcome subsequently provided that there is neither any interaction between the particle and its external environment nor that any other physical quantities are measured. The state of this particle will be denoted $|a_j\rangle$. The meaning of this state is twofold: this is the state of the particle following measurement provided that the outcome a_j is attained and also that this is the state which will yield a_j with certainty if A is measured. The state $|a_j\rangle$ described in this way is called the state associated with a_j. These meanings are illustrated schematically in Figs. III.1.2 and III.1.3.
The inner product of two states $|\Psi\rangle$ and $|\Phi\rangle$ is denoted $\langle\Phi|\Psi\rangle$ and aside from being linear in the righthand argument, satisfies

$$\langle\Psi|\Phi\rangle = \big(\langle\Phi|\Psi\rangle\big)^*. \tag{III.1.1}$$

The inner product gives rise to the notion of orthonormality and orthonormal bases. In terms of physical states these are fixed via the following requirement.

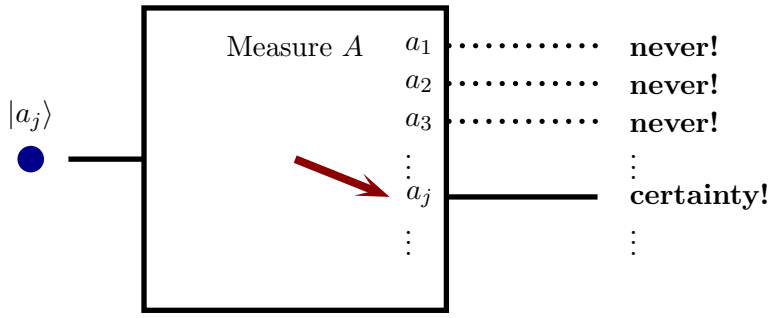


Figure III.1.2: Physical meaning of the state $|a_j\rangle$ that is associated with the outcome a_j in terms of outcomes of a subsequent measurement of A.

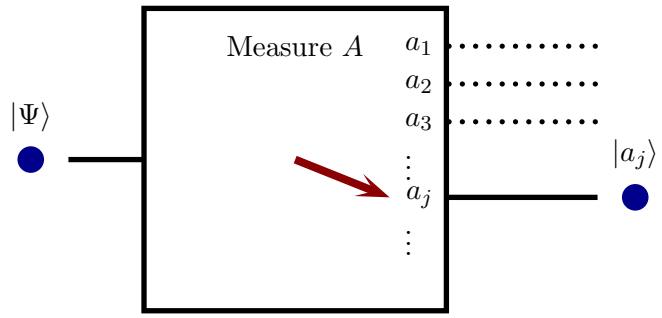


Figure III.1.3: Physical meaning of the state $|a_j\rangle$ that is associated with the outcome a_j given that outcome a_j did occur. None of the measurement outcomes is guaranteed with certainty, but the given that a_j did occur, the state after measurement is definitely $|a_j\rangle$.

The kets associated with mutually incompatible outcomes of any *one* measurement form an orthonormal basis. Conversely the kets that form any orthonormal basis are associated with the mutually incompatible measurement outcomes of one measurement.

For a quantum system and physical quantity A, it is convenient to list the measurement outcomes and states in the form of Table III.1.1.

Given an orthonormal basis $\{|a_1\rangle, |a_2\rangle, \ldots |a_N\rangle\}$, any state can be expressed as a linear combination or superposition,

$$|\Psi\rangle = \sum_{j=1}^{N} c_j |a_j\rangle \tag{III.1.2}$$

where c_j are unique complex numbers. These can be determined via

$$c_j = \langle a_j|\Psi\rangle . \tag{III.1.3}$$

Outcome	Associated State
a_1	$\|a_1\rangle$
a_2	$\|a_2\rangle$
a_3	$\|a_3\rangle$
$\vdots$	$\vdots$

Table III.1.1: Table of measurement outcomes and states for measurement of A. The possible measurement outcomes are denoted a_i and the associated states are denoted $|a_i\rangle$.

Any orthonormal basis vectors satisfy the *completeness relation*

$$\sum_j |a_j\rangle \langle a_j| = \hat{I} \tag{III.1.4}$$

where $\hat{I}$ is the identity operator.

The state of a system allows for calculation of probabilities of measurement outcomes for any measurable quantity A.

Given that a particle in the state $|\Psi\rangle$ is subjected to measurement of A, the probabilities of the various measurement outcomes are

$$\Pr(a_j) = |\langle a_j|\Psi\rangle|^2. \tag{III.1.5}$$

An essential requirement for this to make sense is that the state be normalized

$$\langle\Psi|\Psi\rangle = 1. \tag{III.1.6}$$

A convenient mathematical entity for calculating inner products is a bra, which is an element of the dual Hilbert space. For any ket $|\Psi\rangle$, there is an associated bra,

$$\langle\Psi| := |\Psi\rangle^\dagger \tag{III.1.7}$$

where, † denotes the complex conjugate transpose operation. A bra is a linear operator that maps kets to complex numbers. This is defined so that

$$\langle\Psi|\left(|\Phi\rangle\right) = \langle\Psi|\Phi\rangle \tag{III.1.8}$$

for any ket $|\Phi\rangle$. Additionally the complex conjugate transpose operation has the properties:

$$\left(\alpha\,|\Phi_1\rangle + \beta\,|\Phi_2\rangle\right)^\dagger = \alpha^*\,\langle\Phi_1| + \beta^*\,\langle\Phi_2| \tag{III.1.9a}$$

$$\langle\Psi|^\dagger = |\Psi\rangle\,. \tag{III.1.9b}$$

A measurable quantity, A, is frequently represented by an operator, $\hat{A}$ which is Hermitian, i.e.

$$\hat{A}^\dagger = \hat{A}. \tag{III.1.10}$$

Given the list of outcomes and associated states of Table III.1.1, the observable which represents the measurement A is

$$\hat{A} = \sum_j a_j\,|a_j\rangle\,\langle a_j|\,. \tag{III.1.11}$$

A mathematical result is that any Hermitian operator can be expressed in this form.

If $\hat{A}$ is an observable with eigenvalues and eigenstates given by

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

then the outcomes and states associated with the measurement of A are:

Outcome	Associated State
a_1	$\|a_1\rangle$
a_2	$\|a_2\rangle$
a_3	$\|a_3\rangle$
$\vdots$	$\vdots$

Observables are useful for calculating expectation values.

For an ensemble of particles, each in the state described by $|\Psi\rangle$ and subjected to measurement of A, for which the associated observable is $\hat{A}$, the expectation value of the outcomes is

$$\langle A \rangle = \langle \Psi | \hat{A} | \Psi \rangle .$$

(III.1.12)

2 Dynamics

The evolution of the state of a quantum mechanical system can be described via a linear, unitary evolution operator.

For a particle in state $|\Psi(0)\rangle$ at time $t = 0$ the state at any later time is given by

$$|\Psi(t)\rangle = \hat{U}(t) \; |\Psi(0)\rangle$$

where $\hat{U}(t)$ is a linear operator which is unitary

$$\hat{U}^\dagger \hat{U} = \hat{I}.$$

The operator $\hat{U}(t)$ is called the evolution operator.

(III.2.1)

The evolution of the state is also specified in terms the *Schrödinger* equation, a differential equation involving the Hamiltonian, $\hat{H}$, which is the energy observable.

The state of a quantum mechanical system satisfies

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

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

(III.2.2)

Whenever the Hamiltonian is time-dependent the following solution for the evolution

operator exists.

If the Hamiltonian $\hat{H}$ is time-independent then,

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar}. \tag{III.2.3}$$

The evolution of energy eigenstates is particularly simple.

Suppose that $|\phi_i\rangle$ is an energy eigenstate of a time-independent Hamiltonian and the associated energy eigenvalue is E_i. Then the eigenstate evolves as

$$|\phi_i\rangle \;\rightarrow\; |\Psi(t)\rangle = e^{-iE_it/\hbar}\,|\phi_i\rangle\,. \tag{III.2.4}$$

This is readily extended to superpositions of energy eigenstates by invoking the linearity of the evolution operator.

Expectation values evolve according to the Ehrenfest theorem.

Consider an ensemble of identical particles, each in the state $|\Psi(t)\rangle$ subject to the same Hamiltonian, $\hat{H}$. The expectation vale of any observable $\hat{A}$ satisfies

$$\frac{d}{dt}\langle A\rangle = \langle\Psi(t)|\,\frac{\partial\hat{A}}{\partial t}\,|\Psi(t)\rangle + \frac{i}{\hbar}\,\langle\Psi(t)|\left[\hat{H},\hat{A}\right]|\Psi(t)\rangle\,. \tag{III.2.5}$$

IV. Quantum Mechanics of Systems of Particles with Spatial Degrees of Freedom: One Dimensional Case

1 States and Measurements

For many physical systems, it is possible to measure position or momentum. Important examples include harmonic oscillators and atoms. Even in the case of physical systems restricted to one spatial dimension, not only are there are infinitely many possible outcomes for position measurements, but the range of outcomes is continuous. Although the basic language and rules for describing these systems are the same as those for systems in which any measurement can only yield one of finitely many possible outcomes, there are a variety of special techniques for dealing with the issues that stem from the existence of an infinite number of possible measurement outcomes.

Experience with classical descriptions of systems with one spatial degree of freedom would suggest that position and momentum measurements should be the cornerstone on which the physical description of a system is based. To a limited extent this is true for quantum mechanical systems; however notions of position and momentum exist here primarily to establish connections to classical mechanical systems. In fact, in quantum systems position and momentum produce problematical mathematical issues. In classical mechanics it is possible to consider a particle whose state is such that it will yield one precise value for a position measurement, e.g. $x = 3.5\,\mathrm{m}$. In contrast for a quantum mechanical system, it turns out that such a situation is accompanied by the fact that, were the particle's momentum to be measured, all values for momentum $-\infty < p < \infty$ are all *equally likely*. There is no way to reconcile this with standard probability theory.

However, it is possible to discuss the process of measuring energy without too many tricky mathematical complications. Fortunately too, energy measurements, or measurements of quantities closely related to energies are far more readily performed in typical laboratory situations. For example, much of the early evidence that supported the quantum mechanical description of the hydrogen atom was derived from spectroscopic experiments involving hydrogen. In these the frequencies of the light emitted by the atoms is easily related to differences in hydrogen atom energy levels. The key feature that provides for a reasonable discussion here is that the possible outcomes of an energy measurement are infinite in number but *not continuous*. A schematic of such a measurement is provided in Fig. IV.1.1. In such cases the energy measurements can be labeled with integers $n = 1, 2, \ldots$. The outcomes and the associated states can be listed as in Table IV.1.1. The fact that these states are all associated with mutually incompatible outcomes of a single measurement implies that they must be orthogonal. Also, it is standard to require that they are normalized. Thus

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \tag{IV.1.1}$$

where the *Kronecker delta symbol* is defined as

$$\delta_{ij} := \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if i = j.} \end{cases} \tag{IV.1.2}$$

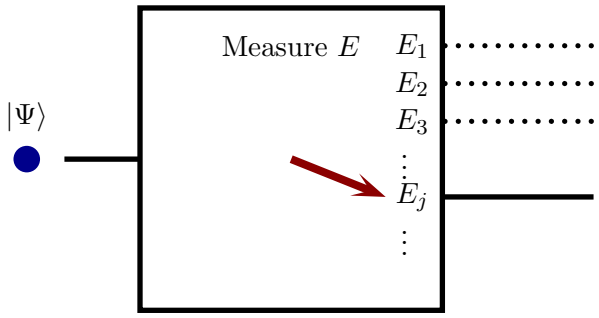


Figure IV.1.1: Cartoon of an energy measurement with an infinite, but discrete, number of outcomes.

Energy	Associated State
E_1	$\|\phi_1\rangle$
E_2	$\|\phi_2\rangle$
E_3	$\|\phi_3\rangle$
$\vdots$	$\vdots$

Table IV.1.1: Energy measurement outcomes and associated states.

For many quantum systems corresponding to particles moving in one dimension, there is only one energy state for a given value of energy; such systems are called *non-degenerate.* In such cases the energy states form a basis for the space of all kets. This means that a general ket can be expressed as

$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n \, |\phi_n\rangle \tag{IV.1.3}$$

where c_n are complex numbers. In terms of vectors, the ket can be represented in the basis $\{|\phi_n\rangle \, | n = 1, 2, \ldots\}$ via

$$\begin{aligned} |\Psi\rangle &= \sum_{n=1}^{\infty} c_n \, |\phi_n\rangle \\ &\leftrightarrow \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{pmatrix}. \end{aligned} \tag{IV.1.4}$$

Special cases are the energy states themselves:

$$|\phi_1\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix},$$

$$|\phi_1\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}, \text{etc,} \ldots$$

Clearly the infinite nature of these vectors will inhibit their effective use in calculations and it is better to work with the abstract format.

Bra vectors are defined in the usual manner, via $\langle\Psi| := |\Psi\rangle^\dagger$. If

$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n |\phi_n\rangle$$

then

$$\begin{aligned} \langle\Psi| &= |\Psi\rangle^\dagger \\ &= \left[\sum_{n=1}^{\infty} c_n |\phi_n\rangle\right]^\dagger \\ &= \sum_{n=1}^{\infty} c_n^* |\phi_n\rangle^\dagger \\ &= \sum_{n=1}^{\infty} c_n^* \langle\phi_n| . \end{aligned} \tag{IV.1.5}$$

In terms of vector representations, bras are represented by row vectors. Thus

$$\begin{aligned} \langle\Psi| &= \sum_{n=1}^{\infty} c_n^* \langle\phi_n| \\ &\leftrightarrow \begin{pmatrix} c_1^* & c_3^* & c_2^* & \ldots \end{pmatrix}. \end{aligned} \tag{IV.1.6}$$

The inner product of two vectors can be computed using these tools and extensions of the methods available for two dimensional vectors. Thus if,

$$|\Phi\rangle = \sum_{n=1}^{\infty} b_n |\phi_n\rangle$$

$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n |\phi_n\rangle .$$

then

$$\begin{aligned}\langle\Phi|\Psi\rangle &= \langle\Phi|\left(|\Psi\rangle\right)\\ &= \left[\sum_{n=1}^{\infty} b_n^* \langle\phi_m|\right]\left[\sum_{m=1}^{\infty} c_m |\phi_n\rangle\right]\\ &= \sum_{n=1}^{\infty} b_n^* \sum_{m=1}^{\infty} c_m \langle\phi_m|\phi_n\rangle\\ &= \sum_{n=1}^{\infty} b_n^* \sum_{m=1}^{\infty} c_m \delta_{nm}\\ &= \sum_{n=1}^{\infty} b_n^* c_n. \end{aligned} \tag{IV.1.7}$$

Note that in such calculations it is crucial to use distinct labels in each sum, since, in general

$$\sum_n \sum_n b_n c_n = b_1c_1 + b_2c_2 + b_3c_3 + \ldots$$

while

$$\sum_m \sum_n b_m c_n = b_1c_1 + b_1c_2 + b_1c_3 + \ldots$$

and thus

$$\sum_n \sum_n b_n c_n \neq \sum_m \sum_n b_m c_n.$$

Thus we have arrived at:

If

$$|\Phi\rangle = \sum_{n=1}^{\infty} b_n |\phi_n\rangle$$
$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n |\phi_n\rangle \tag{IV.1.8}$$

then

$$\langle\Phi|\Psi\rangle = \sum_{n=1}^{\infty} b_n^* c_n.$$

The outcomes of energy measurements can be predicted using the similar rule to that for spin-1/2 measurements. Thus for a particle in the state $|\Psi\rangle$,

$$\Pr(E_n) = |\langle\phi_n|\Psi\rangle|^2 \tag{IV.1.9}$$

Exercise: Show that if

$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n |\phi_n\rangle$$

then

$$\Pr(E_n) = |c_n|^2.$$

One consequence of this result is that

$$\sum_{n=1}^{\infty} |c_n|^2 = 1 \tag{IV.1.10}$$

and this is equivalent to the requirement that the state be normalized.

Any ket $|\Psi\rangle$ that represents the state of a physical system must be normalized:

$$\langle\Psi|\Psi\rangle = 1. \tag{IV.1.11}$$

Exercise: Show that the general normalization condition, Eq. (IV.1.11), implies Eq. (IV.1.10) for the state $|\Psi\rangle = \sum_{n=1}^{\infty} c_n |\phi_n\rangle$.

The fact that the energy states for a basis for all possible kets is equivalent to the *completeness relation*

$$\sum_{n=1}^{\infty} |\phi_n\rangle \langle\phi_n| = \hat{I}. \tag{IV.1.12}$$

In terms of vector representations

$$|\phi_1\rangle \langle\phi_1| \leftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & \dots \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \dots \\ 0 & 0 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

and repetition of this eventually yields Eq. (IV.1.12).

It should be apparent that typical states of these physical systems involve infinite sums of energy states (basis vectors). We seldom actually sum these so as to arrive at a single ket which represents the state of the system, as is possible for spin-1/2 systems. Mostly these infinite expansions are "placeholders" for later calculations. Even so, it may be

apparent that it will be necessary to perform infinite sums when calculating probabilities or expectation values of measurement outcomes. The following example indicates that this is frequently possible to actually do this.

Example: Suppose that for a quantum mechanical system the energies are

$$E_n = \frac{E_1\, 2^n}{2}$$

and that the system is in the state

$$|\Psi\rangle = \sum_{n=1}^{\infty} \frac{A}{3^n} |\phi_n\rangle$$

where $\{|\phi_n\rangle\}$ are the energy eigenstates. Determine A, the probability with which an energy measurement yield the outcome E_n for any n and the expectation value of the energy measurements.

Answer: The normalization condition gives:

$$\langle\Psi|\Psi\rangle = 1 \Rightarrow \sum_{n=1}^{\infty} \left|\frac{A}{3^n}\right|^2 = 1.$$

Thus

$$\begin{aligned} 1 &= |A|^2 \sum_{n=1}^{\infty} \frac{1}{3^{2n}} \\ &= |A|^2 \sum_{n=1}^{\infty} \frac{1}{9^n} \\ &= |A|^2\, \frac{1}{9} \left[1 + \frac{1}{9} + \frac{1}{9^2} + \ldots\right] \end{aligned}$$

The term in brackets is a geometric series and the following applies

$$1 + r + r^2 = \ldots = \frac{1}{1-r}$$

provided that $|r| < 1$. Thus

$$1 = |A|^2\, \frac{1}{9} \frac{1}{1 - 1/9} = |A|^2\, \frac{8}{9}$$

which leaves

$$|A| = \sqrt{\frac{9}{8}}$$

and the most convenient choice is

$$A = \sqrt{\frac{9}{8}} = \frac{3}{\sqrt{8}}.$$

Thus the normalized version of the state is

$$|\Psi\rangle = \sum_{n=1}^{\infty} \frac{1}{\sqrt{8}} \frac{1}{3^{n-1}} |\phi_n\rangle .$$

Now the probability of any given outcome for an energy measurement is

$$\Pr(E_n) = |\langle \phi_n | \Psi \rangle|^2$$

To this end

$$\begin{aligned}
\langle \phi_n | \Psi \rangle &= \langle \phi_n | \sum_{m=1}^{\infty} \frac{1}{\sqrt{8}} \frac{1}{3^{m-1}} |\phi_m\rangle \\
&= \sum_{m=1}^{\infty} \frac{1}{\sqrt{8}} \frac{1}{3^{m-1}} \langle \phi_n | \phi_m \rangle \\
&= \sum_{m=1}^{\infty} \frac{1}{\sqrt{8}} \frac{1}{3^{m-1}} \delta_{nm} \\
&= \frac{1}{\sqrt{8}} \frac{1}{3^{n-1}}.
\end{aligned}$$

Thus

$$\Pr(E_n) = \frac{1}{8} \frac{1}{9^{n-1}}.$$

The expectation value is

$$\begin{aligned}
\langle E \rangle &= \sum_{n=1}^{\infty} E_n \Pr(E_n) \\
&= \sum_{n=1}^{\infty} \frac{E_1 \, 2^n}{2} \frac{1}{8} \frac{1}{9^{n-1}} \\
&= \frac{E_1}{16 \cdot 9} \sum_{n=1}^{\infty} \left(\frac{2}{9} \right)^n \\
&= \frac{E_1}{16 \cdot 9} \frac{2}{9} \left[1 + \frac{2}{9} + \left(\frac{2}{9} \right)^2 + \ldots \right] \\
&= \frac{E_1}{8 \cdot 9 \cdot 9} \frac{1}{1 - 2/9} \\
&= \frac{E_1}{8 \cdot 9 \cdot 7}.
\end{aligned}$$

Note that the expectation value could also be computed using the energy observable,

i.e. the Hamiltonian

$$\hat{H} = \sum_{n=1}^{\infty} E_n \left|\phi_n\right\rangle \left\langle\phi_n\right| .$$

and

$$\langle E\rangle = \langle\Psi| \hat{H} |\Psi\rangle . \tag{IV.1.13}$$

In practice, calculating $\langle E\rangle$ using Eq. (IV.1.13) is no easier than first calculating the probabilities of the outcomes and then summing as in the previous example.

Time evolution of any state follows the general rules. Thus for a time-independent Hamiltonian, the evolution operator is

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar}.$$

Unlike the cases of spin-1/2 and other finite dimensional quantum systems, it is seldom possible to exponentiate this operator in closed form. However for a general initial state $|\Psi(0)\rangle = \sum_n c_n |\phi_n\rangle$, the state at a later time is

$$\begin{aligned} |\Psi(t)\rangle &= \hat{U}(t) |\Psi(0)\rangle \\ &= e^{-i\hat{H}t/\hbar} \sum_{n=1}^{\infty} c_n |\phi_n\rangle \\ &= \sum_{n=1}^{\infty} c_n e^{-i\hat{H}t/\hbar} |\phi_n\rangle \\ &= \sum_{n=1}^{\infty} c_n e^{-iE_n t/\hbar} |\phi_n\rangle . \end{aligned} \tag{IV.1.14}$$

Again, it is usually impossible to sum this to arrive at an expression involving just a single ket and coefficient.

1.1 Position Measurements: Discrete approach

The general description of position measurements in classical mechanics assumes that any *real number* $-\infty < x < \infty$ is a possible outcome of a position measurement. An equivalent quantum mechanical description of position measurements will surely mirror this with infinitely many possible outcomes possible. However, if this is so then the number of outcomes will not be *countable,* i.e. cannot be labeled by integers, and thus the mathematics of countably infinite sequences of vectors that was developed in the previous sections cannot be translated immediately to describe position measurement outcomes.

A first step is to consider position measurements with a finite precision. Thus suppose that the measuring device has a resolution of Δx. For example standard meter sticks are marked in steps of 1 mm and for these $\Delta x = 1\,\text{mm}$. We can break the range of all possible position measurement outcomes into sections, or *bins*, each of whose length is Δx. The actual outcome can be represented by the value of the x at the midpoint of each bin. This is illustrated in Fig. IV.1.2.

Each bin can be labeled with integers j according to the scheme of Fig. IV.1.2. Alternatively each bin can be labeled using the value at center of the j^{th} bin, i.e. $x_j = j\Delta x$. This

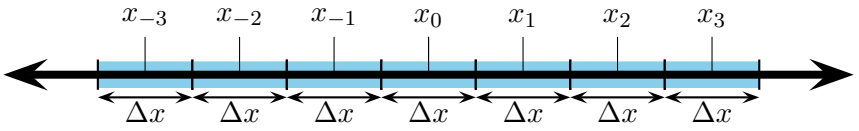


Figure IV.1.2: Discrete position measurement. The axis is divided into bins of length Δx. The bins can be labeled by counting from the centermost bin (centered at $x = 0$) outwards. Bins to the right will be labeled by positive integers, $j = 1, 2, \ldots$, and those to the left by negative integers, $j = -1, -2, \ldots$. Note that the center of the bin labeled j is at $x_j = j\Delta x$.

forms the basis for assigning position values for this type of discretized position measurement. If the particle is located in bin j during a position measurement, then it's position value is said to be x_j. Thus the possible outcomes for a classical position measurement are $\ldots x_{-1}, x_0, x_1 \ldots$.

A quantum mechanical description of this measurement begins with a description of the position measurement outcomes and the associated states. These are listed in Table IV.1.2. As with other types of measurements on quantum mechanical systems, the state $|x_j\rangle$ has

Position	Associated State
$\vdots$	$\vdots$
x_{-2}	$\|x_{-2}\rangle$
x_{-1}	$\|x_{-1}\rangle$
x_0	$\|x_0\rangle$
x_1	$\|x_1\rangle$
x_2	$\|x_2\rangle$
$\vdots$	$\vdots$

Table IV.1.2: Discretized position measurement outcomes and associated states.

the physical interpretation that if the position of the particle is measured then, the outcome will be that the particle is located in bin j with certainty. In terms of position values the outcome will fall in the range $(x_j - \Delta x/2, x_j + \Delta x/2)$ and we refer to this by saying that that the outcome of the position measurement will be x_j with certainty although this really just refers to the label at the center of the bin. These states are referred to as *position*

states.

Outcomes that lie within distinct bins will not be confused and thus the position states are orthonormal, i.e.

$$\langle x_i | x_j \rangle = \delta_{ij}. \tag{IV.1.15}$$

The position states form a basis for the vector space consisting of all states and this is indicated by the completeness relation

$$\sum_{j=-\infty}^{\infty} |x_j\rangle \langle x_j| = \hat{I}. \tag{IV.1.16}$$

For a general state $|\Psi\rangle$,

$$\begin{aligned} |\Psi\rangle &= \hat{I} |\Psi\rangle \\ &= \sum_{j=-\infty}^{\infty} |x_j\rangle \langle x_j| |\Psi\rangle \\ &= \sum_{j=-\infty}^{\infty} |x_j\rangle \langle x_j|\Psi\rangle . \end{aligned} \tag{IV.1.17}$$

The complex numbers

$$\Psi(x_j) := \langle x_j|\Psi\rangle \tag{IV.1.18}$$

form the components of $|\Psi\rangle$ in the basis $\{\ldots, |x_{-2}\rangle, |x_{-1}\rangle, |x_0\rangle, |x_1\rangle, |x_2\rangle, \ldots\}$, i.e.

$$|\Psi\rangle = \sum_{j=-\infty}^{\infty} \Psi(x_j) |x_j\rangle . \tag{IV.1.19}$$

Again the state $|\Psi\rangle$ dictates the probabilities for discretized position measurement outcomes. Thus

$$\Pr(x_j) = |\langle x_j|\Psi\rangle|^2. \tag{IV.1.20}$$

Exercise: Show that if

$$|\Psi\rangle = \sum_{j=-\infty}^{\infty} \Psi(x_j) |x_j\rangle$$

then

$$\Pr(x_j) = |\Psi(x_j)|^2. \tag{IV.1.21}$$

One consequence of Eq. (IV.1.21) is that

$$\sum_{j=-\infty}^{\infty} |\Psi(x_j)|^2 = 1$$

which is equivalent to the fact that the state must be normalized, i.e.

$$\langle \Psi|\Psi\rangle = 1. \tag{IV.1.22}$$

The observable associated with position measurements is

$$\hat{x} := \sum_{j=-\infty}^{\infty} x_j \, |x_j\rangle \langle x_j| \tag{IV.1.23}$$

and it can be shown that

$$\hat{x} \, |x_n\rangle := x_n \, |x_n\rangle \tag{IV.1.24}$$

Exercise: Show that

$$\hat{x} \, |x_n\rangle := x_n \, |x_n\rangle \, .$$

The expectation value of x for an ensemble of particles in the state $|\Psi\rangle$ is determined via

$$\begin{aligned}
\langle x \rangle &= \langle \Psi | \, \hat{x} \, | \Psi \rangle \\
&= \left[\sum_{j=-\infty}^{\infty} \Psi(x_j)^* \, \langle x_j | \right] \hat{x} \left[\sum_{k=-\infty}^{\infty} \Psi(x_k) \, |x_k\rangle \right] \\
&= \left[\sum_{j=-\infty}^{\infty} \Psi(x_j)^* \, \langle x_j | \right] \sum_{k=-\infty}^{\infty} \Psi(x_k) \hat{x} \, |x_k\rangle \\
&= \left[\sum_{j=-\infty}^{\infty} \Psi(x_j)^* \, \langle x_j | \right] \sum_{k=-\infty}^{\infty} \Psi(x_k) x_k \, |x_k\rangle \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \Psi(x_j)^* \Psi(x_k) x_k \, \langle x_j | x_k \rangle \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \Psi(x_j)^* \Psi(x_k) x_k \delta_{jk}
\end{aligned}$$

which gives

$$\langle x \rangle = \sum_{j=-\infty}^{\infty} x_j \, |\Psi(x_j)|^2 \, . \tag{IV.1.25}$$

Another important consequence is that the probability that the measurement outcome yields any value in the bins ranging from that labeled j_{low} to j_{high} is

$$\Pr\left(x_{j_{\text{low}}} < x < x_{j_{\text{high}}} \right) = \sum_{j=j_{\text{low}}}^{j_{\text{high}}} |\Psi(x_j)|^2 . \tag{IV.1.26}$$

Thus

$$\begin{aligned}\Pr\left(x_{j_{\text{low}}} < x < x_{j_{\text{high}}}\right) &= \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \Psi(x_j)^* \Psi(x_j) \\ &= \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \langle x_j|\Psi\rangle^* \langle x_j|\Psi\rangle \\ &= \sum_{j=j_{\text{low}}}^{j_{\text{high}}} \langle \Psi|x_j\rangle \langle x_j|\Psi\rangle \\ &= \langle\Psi| \left(\sum_{j=j_{\text{low}}}^{j_{\text{high}}} |x_j\rangle \langle x_j| \right) |\Psi\rangle \end{aligned} \tag{IV.1.27}$$

and in this sense the operator

$$\sum_{j=j_{\text{low}}}^{j_{\text{high}}} |x_j\rangle \langle x_j|$$

is associated with position measurements in the entire range $x_{j_{\text{low}}} < x < x_{j_{\text{high}}}$.

1.2 Position Measurements: Continuous approach

The resolution of position measurements can always be improved; this amounts to decreasing the size of the bins, Δx, in the discretized version. The mechanisms described for the discretized quantum description above, apply whenever $\Delta x \neq 0$. Is there an underlying description when $\Delta x = 0$? In a thorough mathematical treatment there is such a description but the mathematical tools used in the discretized version of position measurements have to be modified substantially. Nevertheless, the basic rules for calculation have many features similar to those of the discretized version. The set of all position states is

$$\big\{|x\rangle \,|\, \text{all real } x\big\}.$$

Note that although the labels inside the kets are real numbers, they are still just labels, and they are not explicitly involved in typical real number calculations. Thus

$$|2.5\rangle + |3.6\rangle \neq |6.1\rangle\,.$$

Additionally there is an important conceptual caution: it is impossible to prepare a particle in any of these states $|x\rangle$. Thus they do not have any reasonable physical interpretation in terms of the outcomes to position measurements.

The role of these states is to support superpositions, from which quantities pertaining to measurement outcomes can be calculated. In superpositions over states with continuous labels, the sums must be replaced by integrals. Thus the general state will have the form

$$\boxed{|\Psi\rangle = \int_{-\infty}^{\infty} \Psi(x)\,|x\rangle\,\mathrm{d}x} \tag{IV.1.28}$$

where $\Psi(x)$ is a complex valued function of x, called *the wavefunction representing* $|\Psi\rangle$ *in the position basis*, or colloquially, "the position wavefunction." In Eq. (IV.1.28), the position states $|x\rangle$ act as basis kets (analogous to the $\{|+\hat{\boldsymbol{z}}\rangle, |-\hat{\boldsymbol{z}}\rangle\}$ basis for spin-1/2 particles) and the position wavefunction $\Psi(x)$ acts as a collection of components, one for each value of x (analogous to c_+ and c_- in $|\Psi\rangle = c_+ |+\hat{\boldsymbol{z}}\rangle + c_- |-\hat{\boldsymbol{z}}\rangle$ for spin-1/2 particles). One could consider arranging the position wavefunction values at all x in a column vector but this would require a vector with (non-countable) infinitely many entries. It is more suitable to represent the state by the wavefunction

$$|\Psi\rangle \leftrightarrow \Psi(x)$$

where, on the right, the function is considered at all possible values of x. A graphical representation of the wavefunction can display information about it at all values of x although this is only really convenient of the wavefunction returns real values for all x. For example, if

$$|\Psi\rangle = \int_{-\infty}^{\infty} Ae^{-x^2/a} |x\rangle \, \mathrm{d}x$$

where $a > 0$ and A is real, then

$$|\Psi\rangle \leftrightarrow \Psi(x) = Ae^{-x^2/a}$$

and the graph of the position wavefunction is illustrated in Fig. IV.1.3.

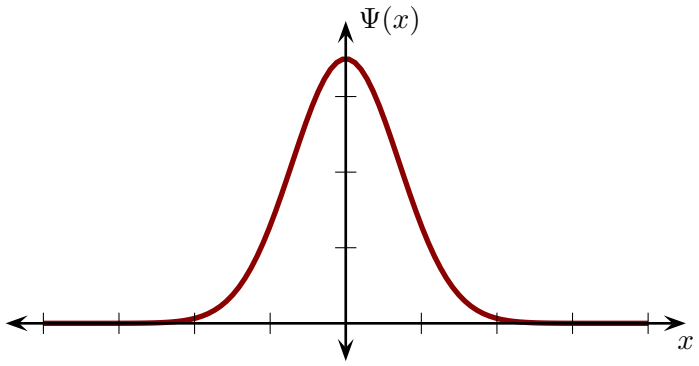


Figure IV.1.3: A possible wavefunction, $\Psi(x) = Ae^{-x^2/a}$ where $a > 0$ and A is real.

Note that the integral of Eq. (IV.1.28) or, for example that of $\int_{-\infty}^{\infty} Ae^{-x^2/a} |x\rangle \, \mathrm{d}x$, are never evaluated in a closed form. This contrasts with typical one dimensional integrals such as

$$\int_{-\infty}^{\infty} Ae^{-x^2/a} \mathrm{d}x = A\sqrt{a\pi}.$$

The integral of Eq. (IV.1.28) is only a representation of a superposition but it is one which will be useful in various later computations.

Associated with ket states, are the bra states. Formally

$$\langle x| := |x\rangle^\dagger .$$

Again the role of bra states will be to operate linearly on ket states to produce complex numbers. Thus for a single value of x, $\langle x|\Psi\rangle$ is a complex number. For a general state,

$$\begin{aligned}\langle x|\Psi\rangle &= \langle x| \int_{-\infty}^{\infty} \Psi(x') \left|x'\right\rangle \mathrm{d}x' \\ &= \int_{-\infty}^{\infty} \Psi(x') \left\langle x|x'\right\rangle \mathrm{d}x'. \end{aligned} \tag{IV.1.29}$$

The only obstacle to actually evaluating the integral of Eq (IV.1.29) is $\langle x|x'\rangle$ which must be a function of x and x'. It remains to specify this. By comparison with the discretized case, we require that

$$\left\langle x|x'\right\rangle = 0$$

whenever $x \neq x'$. The difficult is to define this when $x = x'$. Whatever definition is adopted, we require that it must satisfy

$$\Psi(x) = \langle x|\Psi\rangle .$$

1.2.1 Dirac delta function

The appropriate tool for defining $\langle x|x'\rangle$ lies in the Dirac delta function. The *Dirac delta function* is defined via:

$$\boxed{\delta(x) := \begin{cases} 0 & \text{if } x \neq 0 \\ \infty & \text{if } x = 0. \end{cases}} \tag{IV.1.30}$$

and so that it satisfies

$$\boxed{\int_{-\infty}^{\infty} f(x)\, \delta(x)\, \mathrm{d}x = f(0)} \tag{IV.1.31}$$

with the usual rules of integration applying. One representation of the Dirac delta function in terms of functions that are well-behaved and integrable in the usual sense is via:

$$\delta_\epsilon(x) := \begin{cases} 0 & \text{if } x < -\epsilon/2 \\ \dfrac{1}{\epsilon} & \text{if } -\epsilon/2 < x < \epsilon/2 \\ 0 & \text{if } x > \epsilon/2. \end{cases}$$

where the notion is that

$$\lim_{\epsilon \to 0} \delta_\epsilon(x) = \delta(x).$$

In a product of the form

$$f(x)\delta_\epsilon(x)$$

all values of $f(x)$ except those in the narrow range (of width ϵ) of $x = 0$ are irrelevant. In this sense the Dirac delta function isolates the value of f at $x = 0$, i.e. $f(0)$.

Example: Show that

$$\delta(x-x') := \begin{cases} 0 & \text{if } x \neq x' \\ \infty & \text{if } x = x'. \end{cases} \tag{IV.1.32}$$

and

$$\boxed{\int_{-\infty}^{\infty} f(x')\,\delta(x-x')\,\mathrm{d}x' = f(x).} \tag{IV.1.33}$$

Answer: First if $x \neq x'$ then $x - x' \neq 0$. Thus, in this case, $\delta(x-x') = 0$. The case where $x = x'$ follows by a similar reasoning.

More importantly,

$$\int_{-\infty}^{\infty} f(x')\,\delta(x-x')\,\mathrm{d}x'$$

can be cast into a standard form using $u := x - x'$. Thus $\mathrm{d}x' = -\mathrm{d}u$ and

$$\begin{aligned} \int_{-\infty}^{\infty} f(x')\,\delta(x-x')\,\mathrm{d}x' &= -\int_{\infty}^{-\infty} f(x-u)\,\delta(u)\,\mathrm{d}u \\ &= \int_{-\infty}^{\infty} f(x-u)\,\delta(u)\,\mathrm{d}u \\ &= f(x+0) = f(x). \end{aligned}$$

Exercise: Show that

$$\boxed{\int_{-\infty}^{\infty} f(x')\,\delta(x-x')\,\mathrm{d}x' = \int_{-\infty}^{\infty} f(x')\,\delta(x'-x)\,\mathrm{d}x'} \tag{IV.1.34}$$

and that therefore

$$\delta(x'-x) = \delta(x-x'). \tag{IV.1.35}$$

Exercise: Show that

$$\boxed{\int_{-\infty}^{\infty} \delta(x-x')\,\mathrm{d}x' = 1.} \tag{IV.1.36}$$

1.2.2 Inner product of position states

The inner product of position states is defined via

$$\boxed{\langle x|x'\rangle = \delta(x-x').} \tag{IV.1.37}$$

Applying this to Eq. (IV.1.29) yields

$$\boxed{\Psi(x) = \langle x|\Psi\rangle\,.} \tag{IV.1.38}$$

A related an formal result, can be derived using the formal properties of the inner product. Consider

$$\begin{aligned}\langle\Psi|x\rangle &= (\langle x|\Psi\rangle)^\dagger \\ &= (\langle x|\Psi\rangle)^* .\end{aligned}$$

Thus

$$\boxed{\langle\Psi|x\rangle = \Psi^*(x).} \tag{IV.1.39}$$

1.2.3 Completeness relation for position states

The fact that any state can be expressed in the form of Eq. (IV.1.28) requires that the position states satisfy a continuous version of the completeness relation. Thus

$$\boxed{\int_{-\infty}^{\infty} |x\rangle\langle x| \, \mathrm{d}x = \hat{I}} \tag{IV.1.40}$$

where $\hat{I}$ is the identity operator. Again the completeness relation is just a formal relationship whose purpose is to be used in various calculations; one will never actually carry out the integral of the completeness relation alone to arrive at an alternative closed form expression of it.

1.2.4 Calculations in terms of wavefunctions

The inner product definition opens the way for calculations using the wavefunctions representing states. First consider the inner product of two states; the following theorem provides an important result and illustrates the typical computational technique.

Theorem: Suppose that the states $|\Psi\rangle, |\Phi\rangle$ are represented by the position wavefunctions $\Psi(x), \Phi(x)$. Then

$$\boxed{\langle\Psi|\Phi\rangle = \int_{-\infty}^{\infty} \Psi^*(x)\Phi(x)\,\mathrm{d}x.} \tag{IV.1.41}$$

Proof: Using the completeness relation

$$\begin{aligned}\langle\Psi|\Phi\rangle &= \langle\Psi| \int_{-\infty}^{\infty} |x\rangle\langle x| \, \mathrm{d}x \, |\Phi\rangle \\ &= \int_{-\infty}^{\infty} \langle\Psi|x\rangle\langle x|\Phi\rangle \, \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \Psi^*(x)\Phi(x)\,\mathrm{d}x.\end{aligned}$$

One consequence of this is the normalization condition for states and wavefunctions:

$$\boxed{\langle\Psi|\Psi\rangle = 1 \Leftrightarrow \int_{-\infty}^{\infty} |\Psi(x)|^2 \,\mathrm{d}x = 1.} \tag{IV.1.42}$$

Example: Two possible states that describe a particle in one dimension are

$$|\Psi\rangle \leftrightarrow \Psi(x) = \left(\frac{1}{a\pi}\right)^{1/4} e^{-x^2/2a} \quad \text{and}$$
$$|\Phi\rangle \leftrightarrow \Phi(x) = \left(\frac{4}{a^3\pi}\right)^{1/4} x\, e^{-x^2/2a}$$

where $a > 0$. Show that the two states are normalized and compute the inner product $\langle\Psi|\Phi\rangle$.

Answer: For $|\Psi\rangle$,

$$\begin{aligned}\langle\Psi|\Psi\rangle &= \int_{-\infty}^{\infty} |\Psi(x)|^2\, \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \left| \left(\frac{1}{a\pi}\right)^{1/4} e^{-x^2/2a} \right|^2 \mathrm{d}x \\ &= \left(\frac{1}{a\pi}\right)^{1/2} \int_{-\infty}^{\infty} e^{-x^2/a} \mathrm{d}x \\ &= \sqrt{\frac{1}{a\pi}}\, \sqrt{a\pi} = 1.\end{aligned}$$

Similarly for $|\Phi\rangle$,

$$\begin{aligned}\langle\Phi|\Phi\rangle &= \int_{-\infty}^{\infty} |\Phi(x)|^2\, \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \left| \left(\frac{4}{a^3\pi}\right)^{1/4} x e^{-x^2/2a} \right|^2 \mathrm{d}x \\ &= \left(\frac{4}{a^3\pi}\right)^{1/2} \int_{-\infty}^{\infty} x^2 e^{-x^2/a} \mathrm{d}x \\ &= \sqrt{\frac{4}{a^3\pi}}\, \frac{\sqrt{a^3\pi}}{2} = 1.\end{aligned}$$

Thus the two states are normalized. The inner product is

$$\begin{aligned}\langle\Psi|\Phi\rangle &= \int_{-\infty}^{\infty} \Psi^*(x)\Phi(x) \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \left(\frac{1}{a\pi}\right)^{1/4} e^{-x^2/2a} \left(\frac{4}{a^3\pi}\right)^{1/4} x\, e^{-x^2/2a} \mathrm{d}x \\ &= \left(\frac{4}{a^4\pi}\right)^{1/4} \int_{-\infty}^{\infty} x\, e^{-x^2/a} \mathrm{d}x \\ &= 0.\end{aligned}$$

1.2.5 Probabilities of position measurement outcomes

The wavefunction can be used to determine the outcome of position measurements via a set of calculations analogous to those leading to Eq. (IV.1.27). Thus for a particle in the state $|\Psi\rangle$,

$$\begin{aligned}\Pr(a<x<b) &= \langle\Psi| \int_a^b |x\rangle\langle x|\,\mathrm{d}x\,|\Psi\rangle \\ &= \int_a^b \langle\Psi|x\rangle\langle x|\Psi\rangle\,\mathrm{d}x \\ &= \int_a^b \Psi^*(x)\Psi(x)\mathrm{d}x \\ &= \int_a^b |\Psi(x)|^2\,\mathrm{d}x. \end{aligned} \tag{IV.1.43}$$

For this reason the quantity

$$P(x) := |\Psi(x)|^2 \tag{IV.1.44}$$

is sometimes called the position probability density. This serves as a partial physical interpretation of the wavefunction in terms of position measurements. However, it is not complete since the two wavefunctions $\Psi(x)$ and $e^{i\theta(x)}\,\Psi(x)$, where $\theta(x)$ depends on x give identical position probability distributions and are not related by a global phase. They do represent distinct physical states of particles but this distinction is not born out in position measurements only (there will be a distinction in terms of other certain measurements).

Example: Determine the position probability density for the state

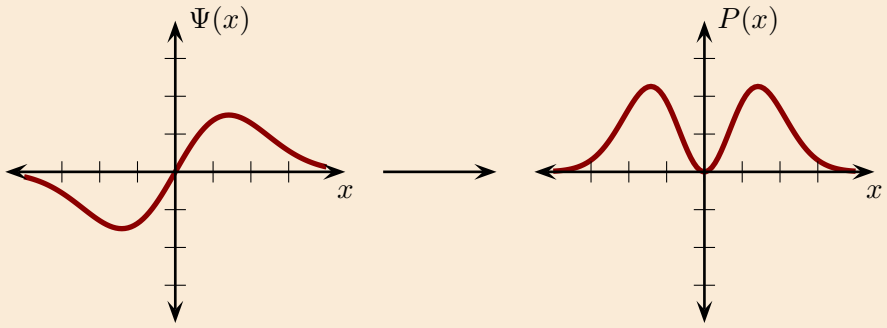
$$|\Psi\rangle \leftrightarrow \left(\frac{4}{a^3\pi}\right)^{1/4} xe^{-x^2/2a}$$

and graph this.

Answer: By Eq. (IV.1.44),

$$P(x) = |\Psi(x)|^2 = \sqrt{\frac{4}{a^3\pi}}\,x^2e^{-x^2/a}.$$

The graph is provided below



Regions where the position probability density is low correspond to regions for which position measurements are not likely or, where the probability of finding the particle is low.

1.3 Position and momentum observables

As with spin-1/2 and other discrete systems, the measurable physical quantities can be represented in terms of observables, i.e. Hermitian operators. The relationship between these, the measurement outcomes and associated states takes the same form as before. For position the observable is denoted $\hat{x}$ and is Hermitian. Thus

$$\hat{x}^\dagger = \hat{x}. \tag{IV.1.45}$$

The position observable operates on position states via

$$\boxed{\hat{x}\,|x\rangle = x\,|x\rangle} \tag{IV.1.46}$$

which mirrors the usual relationship between eigenstates, eigenvalues and measurement outcomes.

Example: Determine $\hat{x}\,|3.6\rangle$.

Answer: By Eq. (IV.1.46),

$$\hat{x}\,|3.6\rangle = 3.6\,|3.6\rangle\,.$$

Eq. (IV.1.49) determines the action of the position observable on general states. In this context the completeness relation is useful; thus

$$\begin{aligned}
\hat{x}\,|\Psi\rangle &= \hat{x}\hat{I}\,|\Psi\rangle \\
&= \hat{x}\left[\int_{-\infty}^{\infty} |x\rangle\,\langle x|\;\mathrm{d}x\right]|\Psi\rangle \\
&= \int_{-\infty}^{\infty} \hat{x}\,|x\rangle\,\langle x|\;|\Psi\rangle\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x\,|x\rangle\,\langle x|\Psi\rangle\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x\,|x\rangle\,\Psi(x)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x\Psi(x)\,|x\rangle\,\mathrm{d}x
\end{aligned} \tag{IV.1.47}$$

The quantity $\hat{x}\,|\Psi\rangle$ is a new ket and, according to this derivation, the corresponding wavefunction is $x\Psi(x)$. Alternatively,

$$\boxed{\langle x|\,\hat{x}\,|\Psi\rangle = x\Psi(x).} \tag{IV.1.48}$$

The process of operating with the position operator on either the ket or wavefunction can be diagrammed as follows:

$$\begin{array}{ccc} |\Psi\rangle & \xrightarrow{\hat{x}} & \hat{x}\,|\Psi\rangle \\ \| & & \| \\ \Psi(x) & \xrightarrow{\hat{x}} & x\Psi(x) \end{array} \tag{IV.1.49}$$

It is usually most convenient to work with the wavefunction representation of the particle's state.

Example: Suppose that

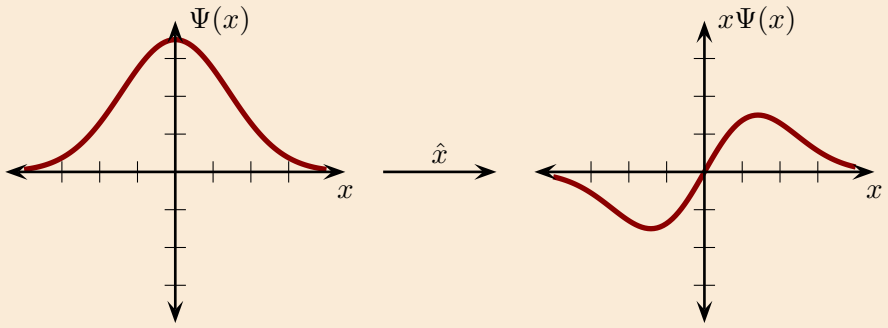
$$|\Psi\rangle \leftrightarrow \Psi(x) = Ae^{-x^2/a}.$$

Determine the wavefunction that represents $\hat{x}\,|\Psi\rangle$ and represent the process of performing this operation graphically.

Answer: Using Eq. (IV.1.49) gives

$$\hat{x}\,|\Psi\rangle \leftrightarrow x\Psi(x) = xAe^{-x^2/a}.$$

This is illustrated in the following diagram



More generally there will be many circumstances where a measurable quantity is constructed from position measurement outcomes. For example, the potential energy of a harmonic oscillator is $V = 1/2kx^2$ where k is a constant and x the displacement of the oscillator. In quantum theory it will be necessary to construct observables for such functions from the position operator. These are built from operators such as

$$\hat{x}^n = \hat{x}\hat{x}\ldots\hat{x}$$

where there are n factors on the right. The following result is useful.

Theorem: Consider the operator $\hat{x}^n$ where n is any integer. Then

$$\boxed{\hat{x}^n\,|x\rangle = x^n\,|x\rangle} \tag{IV.1.50}$$

and

$$\boxed{\langle x|\,\hat{x}^n = x^n\,\langle x|\,.} \tag{IV.1.51}$$

Proof: To follow. •

Now consider the measurable quantity

$$A(x) = a_0 + a_1 x + a_2 x^2 + \ldots = \sum_n a_n x^n.$$

The observable associated with this is defined to be

$$\hat{A} = a_0 + a_1\hat{x} + a_2\hat{x}^2 + \ldots = \sum_n a_n \hat{x} x^n. \tag{IV.1.52}$$

It follows that

$$\hat{A}\,|x\rangle = A(x)\,|x\rangle$$

for any position state. For a general state $|\Psi\rangle$, consider

$$|\Phi\rangle = \hat{A}\,|\Psi\rangle\,.$$

Then it is readily proved that

$$\Phi(x) = A(x)\Psi(x)$$

and this is summarized in the following diagram.

$$\begin{array}{ccc} |\Psi\rangle & \xrightarrow{\hat{A}} & \hat{A}\,|\Psi\rangle \\ \| & & \| \\ \Psi(x) & \xrightarrow{\hat{A}} & A(x)\Psi(x) \end{array} \tag{IV.1.53}$$

Example: Suppose that

$$|\Psi\rangle \leftrightarrow \Psi(x) = \left(\frac{1}{a\pi}\right)^{1/4} e^{-x^2/2a}$$

and that the potential energy for this particle is $V = 1/2kx^2$ where k is a constant. Construct the potential energy operator, $\hat{V}$, and determine the wavefunction that represents $\hat{V}\,|\Psi\rangle$. Represent the process of performing this operation graphically.

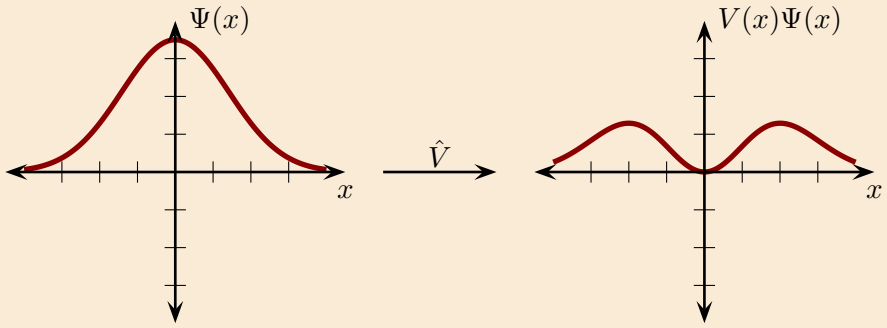
Answer: Using Eq. (IV.1.52) gives

$$\hat{V} = \frac{1}{2}\, k\hat{x}^2.$$

Then Eq. (IV.1.53) gives that the wavefunction transforms as

$$\hat{V}\,|\Psi\rangle \leftrightarrow V(x)\Psi(x) = \left(\frac{1}{a\pi}\right)^{1/4} \frac{1}{2}\, kx^2 e^{-x^2/2a}.$$

This is illustrated in the following diagram



It is also possible to measure momentum of a one dimensional quantum system. The observable which describes momentum measurements is denoted $\hat{p}$ and the momentum states are $|p\rangle$. The usual requirements are that

$$\hat{p}^\dagger = \hat{p} \tag{IV.1.54}$$

and

$$\boxed{\hat{p}\,|p\rangle = p\,|p\rangle} \tag{IV.1.55}$$

It remains to specify the action of the momentum operator on the position wavefunctions. There is no way to derive this from the basic structure of quantum mechanics; the rule must be inferred from experimental evidence, typically those involving applications of the deBroglie hypothesis. A variety of alternative formulations for the relationship between the momentum operator and position states exist. We adopt that given as follows.

$$\begin{array}{ccc} |\Psi\rangle & \xrightarrow{\hat{p}} & \hat{p}\,|\Psi\rangle \\ \| & & \| \\ \Psi(x) & \xrightarrow{\hat{p}} & -i\hbar\dfrac{\partial\Psi}{\partial x} \end{array} \tag{IV.1.56}$$

Alternatively,

$$\boxed{\langle x|\,\hat{p}\,|\Psi\rangle = -i\hbar\frac{\partial\Psi}{\partial x}.} \tag{IV.1.57}$$

Example: Suppose that

$$|\Psi\rangle \leftrightarrow \Psi(x) = \left(\frac{4}{a^3\pi}\right)^{1/4} xe^{-x^2/2a}.$$

Determine the wavefunction that represents $\hat{p}\,|\Psi\rangle$. Graph this transformation.

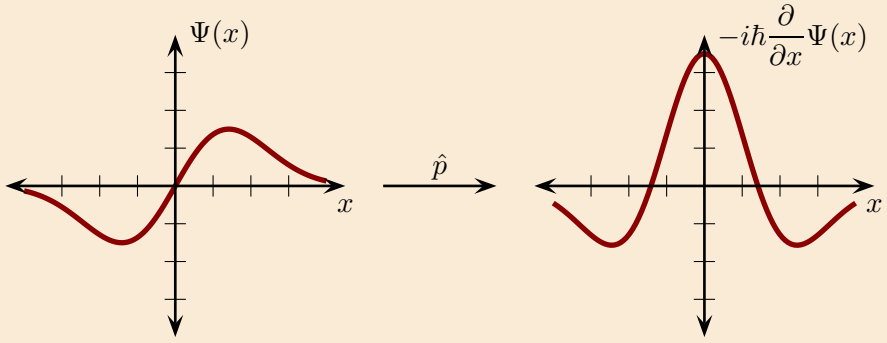
Answer: By Eq. (IV.1.56),

$$\hat{p}\,|\Psi\rangle \leftrightarrow -i\hbar\frac{\partial}{\partial x}\Psi(x).$$

Thus

$$\begin{aligned}\hat{p}\,|\Psi\rangle &\leftrightarrow -i\hbar\left(\frac{4}{a^3\pi}\right)^{1/4}\frac{\partial}{\partial x}\left(xe^{-x^2/2a}\right)\\ &= -i\hbar\left(\frac{4}{a^3\pi}\right)^{1/4}\left(1-\frac{x^2}{a}\right)e^{-x^2/2a}.\end{aligned}$$

Graphically this is represented as follows.



At this stage it is possible to determine the wavefunction for a state of definite momentum. Let

$$\Psi_p(x) = \langle x|p\rangle$$

Then

$$\hat{p}\,|p\rangle = p\,|p\rangle \Rightarrow \langle x|\,\hat{p}\,|p\rangle = p\,\langle x|p\rangle$$

which implies that

$$-i\hbar\frac{\partial\Psi_p(x)}{\partial x} = p\Psi_p(x).$$

The solution to this differential equation is

$$\Psi_p(x) = Ae^{ipx/\hbar}$$

where A is an arbitrary constant. Thus the position wavefunction for a momentum state is

$$\boxed{|p\rangle \leftrightarrow Ae^{ipx/\hbar}} \qquad \text{(IV.1.58)}$$

Exercise: Show that

$$\Psi_p(x) = Ae^{\alpha x}$$

satisfies

$$-i\hbar\frac{\partial\Psi_p(x)}{\partial x} = p\Psi_p(x)$$

provided that $\alpha = ip/\hbar$.

The wavefunction for a state with definite momentum is identical to the complex representation of a wave, e^{ikx} where $k = 2\pi/\lambda$ as the wavenumber. From this it follows that the wavefunction for a particle is exactly that of a wave where the momentum and wavelength are related by

$$\lambda = \frac{h}{p}. \tag{IV.1.59}$$

This is the deBroglie relationship.

It is often necessary to consider function of momentum such as the kinetic energy of a particle of mass m.

$$K = \frac{p^2}{2m}$$

In general these will be constructed in terms of a series such as

$$A(p) = a_0 + a_1 p + a_2 p^2 + \ldots = \sum_n a_n p^n.$$

The observable associated with this is defined to be

$$\hat{A} = a_0 + a_1\hat{p} + a_2\hat{p}^2 + \ldots = \sum_n a_n \hat{x} p^n. \tag{IV.1.60}$$

The following theorem is useful.

Theorem: Consider the operator $\hat{x}^n$ where n is any integer. Then

$$\boxed{\hat{p}^n\left|\Psi\right\rangle \leftrightarrow (-i\hbar)^n\frac{\partial^n\Psi}{\partial x^n}.} \tag{IV.1.61}$$

Proof: To follow. •

Thus

$$\hat{A}\left|\Psi\right\rangle \leftrightarrow \sum_n a_n\,(-i\hbar)^n\frac{\partial^n\Psi}{\partial x^n}.$$

2 Determining Energy Eigenstates

It is often useful to provide a quantum mechanical analog of a classical system. Examples include quantum mechanical descriptions of free particles and harmonic oscillators. The

task will be to use the classical description as a guide for determining the states and dynamics of the analogous quantum system.

In general, any quantum mechanical system can be described in terms of energy eigenstates; these are the states that return specific energy measurement outcomes with certainty. Energy eigenstates satisfy

$$\hat{H}\left|\phi_E\right\rangle = E\left|\phi_E\right\rangle \tag{IV.2.62}$$

where where $\hat{H}$ is the Hamiltonian operator, E denotes a possible energy measurement outcome and $|\phi_E\rangle$ is the state that gives outcome E with certainty when the energy of the system is measured. The importance of these lies in the fact that a general state can always be expressed as a superposition of energy eigenstates, i.e.

$$\left|\Psi\right\rangle = \sum_E c_E \left|\phi_E\right\rangle$$

where c_E are complex numbers. Additionally the time evolution of a system initially in an energy eigenstate, $|\Psi(0)\rangle = |\phi_E\rangle$ proceeds via,

$$\begin{aligned}\left|\Psi(t)\right\rangle &= \hat{U}(t)\left|\Psi(0)\right\rangle = \hat{U}(t)\left|\phi_E\right\rangle \\ &= e^{-i\hat{H}t/\hbar}\left|\phi_E\right\rangle \\ &= e^{-iEt/\hbar}\left|\phi_E\right\rangle .\end{aligned}$$

This can be extended to superpositions in the general fashion,

$$\left|\Psi(0)\right\rangle = \sum_E c_E\left|\phi_E\right\rangle = \sum_E c_E e^{-iEt/\hbar}\left|\phi_E\right\rangle .$$

Thus the key step in describing quantum systems in one dimension is to determine the Hamiltonian and to find its eigenstates. This is informed by the energy for the corresponding classical system. The total energy of a particle of mass m is

$$E = \frac{p^2}{2m} + V(x)$$

where p is the particle's momentum and $V(x)$ is the potential energy. As for spin-1/2 systems, the transition to the Hamiltonian is accomplished by replacing the physically measurable quantities x and p by observables, i.e. Hermitian operators, $\hat{x}$ and $\hat{p}$. Thus

$$\boxed{\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}).} \tag{IV.2.63}$$

Eq. (IV.2.62) gives an operator equation

$$\left[\frac{\hat{p}^2}{2m} + V(\hat{x})\right]\left|\phi_E\right\rangle = E\left|\phi_E\right\rangle .$$

In some circumstances, such as the quantum harmonic oscillator, it is possible to proceed directly from this algebraic equation and via a series of operator manipulations arrive at

expressions for energy eigenvalues and eigenstates. However, more often, the only route for solving the energy eigenvalue equation is to translate it into an equivalent equation involving wavefunctions. Let

$$|\phi_E\rangle \leftrightarrow \phi_E(x).$$

Then according to Eqs (IV.1.53) and (IV.1.61) the eigenvalue equation becomes

$$(-i\hbar)^2 \frac{1}{2m} \frac{\partial^2 \phi_E(x)}{\partial x^2} + V(x)\phi_E(x) = E\phi_E(x)$$

or equivalently

$$\boxed{-\frac{\hbar^2}{2m} \frac{\partial^2 \phi_E(x)}{\partial x^2} + V(x)\phi_E(x) = E\phi_E(x).} \tag{IV.2.64}$$

This is the *time-independent Schrödinger equation* (TISE). Thus the task of finding the energy eigenvalues and eigenstates has been reduced to that of solving a second order differential equation. It is important to note that, for any *single* solution, $\phi_E(x)$, the energy, E, is one *single real number which applies for all values of* x. However there are many possible energies, each corresponding to at least one wavefunction.

Any solution to the TISE must be normalized, i.e.

$$\int_{-\infty}^{\infty} |\phi_E(x)|^2 \mathrm{d}x = 1$$

and this typically constrains the possible solutions since it implies that

$$\phi_E(x) \to 0 \quad \text{as} \quad x \to \pm\infty. \tag{IV.2.65}$$

Additional constraints are that, in any regions where the potential is not infinite, both $\phi_E(x)$ and $\frac{\partial \phi_E(x)}{\partial x}$ must be continuous. If the potential is infinite at any point, it is possible that $\frac{\partial \phi_E(x)}{\partial x}$ may be discontinuous at that point.

2.0.1 Particle in an infinite well

The basic techniques for solving the TISE are well illustrated by considering a particle of mass m which is restricted to the region $0 \leqslant x \leqslant L$ and which moves freely inside this region. The potential is then

$$V(x) = \begin{cases} 0 & 0 \leqslant x \leqslant L \\ \infty & \text{otherwise.} \end{cases}$$

and, when graphed, resembles a well of infinite depth. The impossibility of a position measurement returning an outcome beyond the well implies that

$$\Psi(x) = 0 \quad \text{for} \quad x < 0 \quad \text{and} \quad x > L.$$

It follows that, for any wavefunction, at the boundaries of the well

$$\Psi(0) = \Psi(L) = 0.$$

Inside the well ($0 \leqslant x \leqslant L$), the time-independent Schrödinger equation gives

$$-\frac{\hbar^2}{2m}\frac{\partial^2\phi_E(x)}{\partial x^2} = E\phi_E(x).$$

and the corresponding *boundary conditions* are

$$\phi_E(0) = \phi_E(L) = 0.$$

Thus the task is to solve the differential equation

$$\frac{\partial^2\phi(x)}{\partial x^2} = -\frac{2mE}{\hbar^2}\phi(x)$$

subject to the given boundary conditions. The general solution is

$$\phi_E(x) = A\sin kx + B\cos kx$$

where A and B are complex constants and k satisfies

$$k^2 = \frac{2mE}{\hbar^2}. \tag{IV.2.66}$$

The first boundary condition implies

$$0 = \phi_E(0) = A\sin k0 + B\cos k0 = B$$

while the second gives

$$0 = \phi_E(L) = A\sin kL + B\cos kL = A\sin kL.$$

the only non-trivial solutions are that

$$k = \frac{n\pi}{L}$$

where $n = 1, 2, 3, \ldots$. Thus the possible energy values and states can be labeled by a *discrete* quantity, namely n. Substituting into Eq. (IV.2.66) gives the possible energy values

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2} \tag{IV.2.67}$$

where the index n labels the possibilities. This can be used to label the corresponding eigenstates, i.e. $|\phi_n\rangle$ which corresponds to wavefunction

$$\phi_n(x) = A\sin\left(\frac{n\pi x}{L}\right). \tag{IV.2.68}$$

The constant A may be determined by applying the normalization condition $\langle\phi_n|\phi_n\rangle = 1$.

Exercise: Show that the normalization condition for any infinite well energy eigenstate implies that

$$|A|^2 = \frac{2}{L}.$$

Energy	Associated State	Associated Wavefunction
$E_1 = \dfrac{\pi^2\hbar^2}{2mL^2}$	$\lvert\phi_1\rangle$	$\sqrt{\dfrac{2}{L}}\sin\left(\dfrac{\pi x}{L}\right)$
$E_2 = \dfrac{4\pi^2\hbar^2}{2mL^2}$	$\lvert\phi_2\rangle$	$\sqrt{\dfrac{2}{L}}\sin\left(\dfrac{2\pi x}{L}\right)$
$E_3 = \dfrac{9\pi^2\hbar^2}{2mL^2}$	$\lvert\phi_3\rangle$	$\sqrt{\dfrac{2}{L}}\sin\left(\dfrac{3\pi x}{L}\right)$
$\vdots$	$\vdots$	$\vdots$

Table IV.2.3: Energy measurement outcomes and associated states for the infinite well.

The most convenient choice is $A = \sqrt{2/L}$. Thus the wavefunctions that represent normalized energy eigenstates are

$$\phi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right). \tag{IV.2.69}$$

The energies, states, and their representative wavefunctions for a particle in the infinite well are listed in Table IV.2.3.

One striking feature about this system is that *the energies are discrete,* i.e. the energies do not assume a continuous range of values. This is depicted schematically in Fig. IV.2.4.

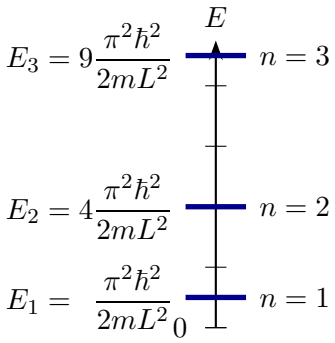


Figure IV.2.4: Energy levels for a particle in an infinite well. An energy measurement will yield a value corresponding to one of the levels. It will never yield a value between any of the levels.

The question arises as to why the energy levels are discrete. This follows from the relationship of Eq. (IV.2.66) between energies and the spatial frequency, k, in the solutions to the time-independent Schrödinger equation,

$$\phi_E(x) = A \sin(kx).$$

This particular solution already guarantees that the left boundary condition,

$$\phi_E(0) = 0$$

is satisfied. Functions of this type oscillate with a spatial frequency determined by k. This in turn will determine the value of $\phi_E(x)$ at $x = L$. For certain values of k, the rate of oscillation will be too small to satisfy $\phi_E(L) = 0$ and for others it will be too large. By this argument, almost all values of k can be ruled out as possibilities that satisfy the right boundary condition. An example of this situation is illustrated in Fig. IV.2.5. This type of graphical representation, alternating between too low and too high, can be used to repeatedly refine k until the wavefunction satisfies

$$\phi_E(L) = 0.$$

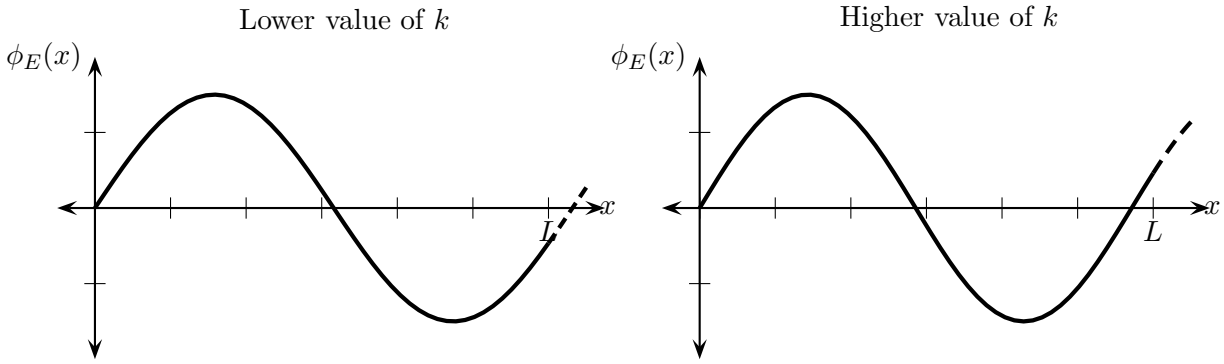


Figure IV.2.5: Wavefunctions for values of spatial frequency k that are slightly too high or slightly too low.

The wavefunctions for the lowest few energy eigenstates are plotted in Fig. IV.2.6. The wavefunctions oscillate, taking on multiple values along the width of the well. It may appear that this somehow describes the trajectory of the particle. However, the purpose of the wavefunction is as a tool for calculating the probabilities of measurement outcomes, using Eq. (IV.1.43). It is important to note that for each energy eigenstate, there is only one energy eigenvalue - a *single real number.*

In general when $E > V$ the wavefunction is an oscillating function of x, with a spatial frequency that increases as the energy increases. Thus wavefunctions for larger energies typically have more peaks and troughs than wavefunctions for lower energies.

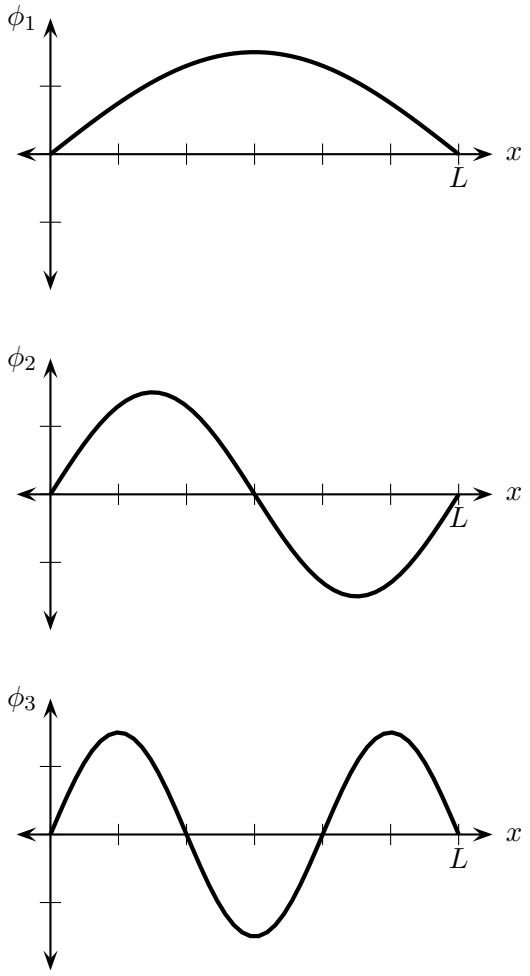


Figure IV.2.6: Energy eigenstate wavefunctions for the infinite well. The wavefunctions for the three lowest energy values are illustrated.

3 Time evolution of wavefunctions

In general the state of a quantum system depends on time and is denoted $|\Psi(t)\rangle$. In terms of wavefunctions, this implies that the wavefunction depends on both position and time. Thus

$$|\Psi(t)\rangle \leftrightarrow \Psi(x,t)$$

where, formally,

$$\Psi(x,t) = \langle x|\Psi(t)\rangle\,. \tag{IV.3.70}$$

The Schroödinger equation,

$$i\hbar\frac{d}{dt}\,|\Psi(t)\rangle = \hat{H}\;|\Psi(t)\rangle\,, \tag{IV.3.71}$$

dictates time evolution of a quantum system and it will be useful to represent this in terms

of wavefunctions. In this case

$$i\hbar\frac{d}{dt}\left|\Psi(t)\right\rangle \leftrightarrow i\hbar\frac{d\Psi(x,t)}{dt}$$

and

$$\left[\frac{\hat{p}^2}{2m}+V(\hat{x})\right]\left|\Psi(t)\right\rangle \leftrightarrow -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(t)}{\partial x^2}+V(x)\Psi(t)$$

imply the time dependent Schrödinger equation in one dimension,

$$\boxed{i\hbar\frac{d\Psi(x,t)}{dt}=-\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2}+V(x)\Psi(x,t).} \tag{IV.3.72}$$

Solving this equation with the appropriate boundary conditions and the initial wavefunction $\Psi(x,0)$ will yield the wavefunction at all subsequent times. Typically this is a very difficult problem, even for simple situations such as a particle in an infinite well.

One strategy would be to compute the evolution operator, $\hat{U}(t)$ which transforms the initial state as

$$\left|\Psi(t)\right\rangle = \hat{U}(t)\left|\Psi(0)\right\rangle .$$

If the Hamiltonian is time independent then it follows that

$$\hat{U}(t)=e^{-i\hat{H}t/\hbar},$$

giving

$$\left|\Psi(t)\right\rangle = e^{-i\hat{H}t/\hbar}\left|\Psi(0)\right\rangle .$$

This is also typically a difficult problem to solve when rewritten in terms of wavefunctions. However, if the initial state is an energy eigenstate, i.e. $\left|\Psi(0)\right\rangle = \left|\phi_E\right\rangle$, then

$$\begin{aligned}\left|\Psi(t)\right\rangle &= e^{-i\hat{H}t/\hbar}\left|\phi_E\right\rangle \\ &= e^{-iEt/\hbar}\left|\phi_E\right\rangle .\end{aligned}$$

This implies that if the initial wavefunction is that for an energy eigenstate, i.e. $\Psi(x,0)=\phi_E(x)$, then

$$\Psi(x,t)=e^{-iEt/\hbar}\,\phi_E(x). \tag{IV.3.73}$$

Note that in this case, $e^{-iEt/\hbar}$ is a global phase; this is the only part of the wavefunction which depends on time. Thus if the particle is initially in an energy eigenstate, probabilities that are computed for any measurement will all be independent of the time at which the measurement was carried out. For this reason, energy eigenstates are sometimes called *stationary states*. In particular the position probability density is independent of time.

Example: Consider a particle of mass m in an infinite well of width L. Suppose that the particle is initially in the state $\left|\phi_2\right\rangle$. Determine an expression for the wavefunction at all later times. Determine an expression for the position probability density, verify that it is time-independent and plot this.

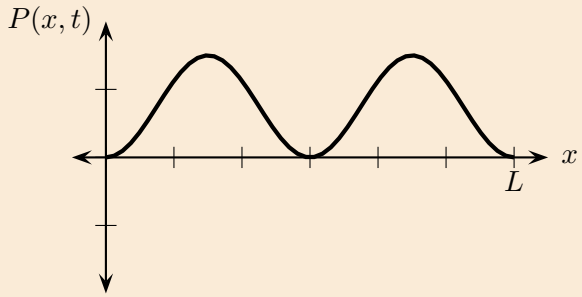
Answer: By Eq. (IV.3.73)

$$\Psi(x,t) = e^{-iE_2 t/\hbar}\,\phi_2(x) = e^{-iE_2 t/\hbar}\sqrt{\frac{2}{L}}\,\sin\left(\frac{2\pi x}{L}\right).$$

Thus the position probability density is

$$P(x,t) = |\Psi(x,t)|^2 = \frac{2}{L}\sin^2\left(\frac{2\pi x}{L}\right).$$

This is clearly time-independent. The plot is illustrated below.



The rule of Eq. (IV.3.73) does not apply to a particle in a general initial state $\Psi(x,t)$. In the general situation the wavefunction corresponds a state $|\Psi(t)\rangle$. However, any initial state can be expressed as a superposition of energy eigenstates,

$$|\Psi(0)\rangle = \sum_E c_E\,|\phi_E\rangle$$

where c_E are complex numbers. Here

$$c_E = \langle\phi_E|\Psi(0)\rangle$$

In terms of wavefunctions this is equivalent to

$$\Psi(x,0) = \sum_E c_E\phi_E(x) \tag{IV.3.74}$$

with

$$c_E = \langle\phi_E|\Psi(0)\rangle = \int_{-\infty}^{\infty}\phi_E^*(x)\Psi(x,0)\mathrm{d}x. \tag{IV.3.75}$$

In terms of states, the evolution proceeds as

$$|\Psi(t)\rangle = \sum_E e^{-iEt/\hbar}\,c_E\,|\phi_E\rangle \tag{IV.3.76}$$

and this implies that the wavefunction at a later time is

$$\Psi(x,t) = \sum_E e^{-iEt/\hbar}\,c_E\phi_E(x). \tag{IV.3.77}$$

Thus the strategy for determining the evolution of a wavefunction is to first express it as a superposition of energy eigenstates as Eq. (IV.3.74), using Eq. (IV.3.75) to determine the coefficients c_E. Then Eq. (IV.3.77) gives the wavefunction at later times.

Example: A particle in an infinite well is initially in the state

$$|\Psi(0)\rangle = \frac{1}{\sqrt{2}}\,|\phi_1\rangle + \frac{1}{\sqrt{2}}\,|\phi_2\rangle$$

where $|\phi_1\rangle$ are energy eigenstates. Determine the state at a later time and use this to determine an expression for the position probability density at a later time.

Answer: According to Eq. (IV.3.76),

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}}\,e^{-iE_1t/\hbar}\,|\phi_1\rangle + \frac{1}{\sqrt{2}}\,e^{-iE_2t/\hbar}\,|\phi_2\rangle$$

where E_i are the energy eigenvalues. The position probability density is given by

$$P(x,t) = |\Psi(x,t)|^2$$

and this requires an expression for $\Psi(x,t) \leftrightarrow |\Psi(t)\rangle$. Here representing the state in terms wavefunctions or using Eq. (IV.3.77)

$$\Psi(x,t) = \frac{1}{\sqrt{2}}\,e^{-iE_1t/\hbar}\,\phi_1(x) + \frac{1}{\sqrt{2}}\,e^{-iE_2t/\hbar}\,\phi_2(x).$$

This implies that

$$\begin{aligned}
P(x,t) = |\Psi(x,t)|^2 &= \Psi^*(x,t)\Psi(x,t) \\
&= \left[\frac{1}{\sqrt{2}}\,e^{-iE_1t/\hbar}\,\phi_1(x) + \frac{1}{\sqrt{2}}\,e^{-iE_2t/\hbar}\,\phi_2(x)\right]^* \\
&\quad \left[\frac{1}{\sqrt{2}}\,e^{-iE_1t/\hbar}\,\phi_1(x) + \frac{1}{\sqrt{2}}\,e^{-iE_2t/\hbar}\,\phi_2(x)\right] \\
&= \left[\frac{1}{\sqrt{2}}\,e^{iE_1t/\hbar}\,\phi_1^*(x) + \frac{1}{\sqrt{2}}\,e^{iE_2t/\hbar}\,\phi_2^*(x)\right] \\
&\quad \left[\frac{1}{\sqrt{2}}\,e^{-iE_1t/\hbar}\,\phi_1(x) + \frac{1}{\sqrt{2}}\,e^{-iE_2t/\hbar}\,\phi_2(x)\right] \\
&= \frac{1}{2}\left[|\phi_1(x)|^2 + |\phi_2(x)|^2 + \phi_1(x)\phi_2(x)\left(e^{i(E_2-E_1)t/\hbar} + e^{-i(E_2-E_1)t/\hbar}\right)\right]
\end{aligned}$$

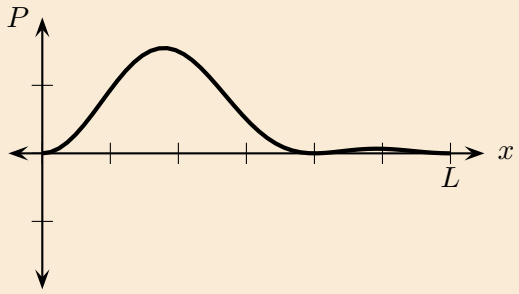
since the energy eigenstate wavefunctions for the infinite well are real. Thus

$$P(x,t) = \frac{1}{2}\left[|\phi_1(x)|^2 + |\phi_2(x)|^2 + 2\phi_1(x)\phi_2(x)\cos\left(e^{i(E_2-E_1)}t/\hbar\right).\right]$$

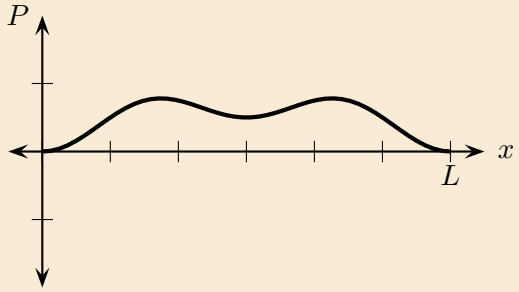
This clearly oscillates with time.

The position probability at various times is as illustrated.

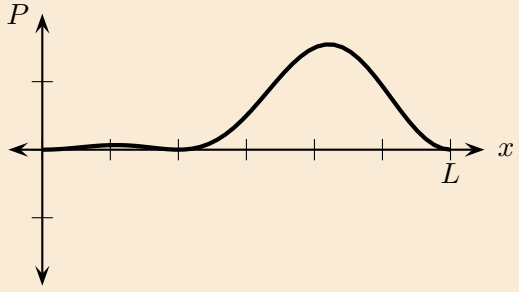
$P(x,t)$ at $t = 0$.



$P(x,t)$ at $t = t = \hbar\pi/2(E_2 - E_1)$.



$P(x,t)$ at $t = \hbar\pi/(E_2 - E_1)$.



The previous example illustrates a general situation in which the state of a particle is a superposition of two energy eigenstates, i.e.

$$|\Psi(0)\rangle = c_i\,|\phi_i\rangle + c_j\,|\phi_j\rangle\,.$$

Here it can be shown that the position probability density oscillates with angular frequency

$$\omega = \frac{E_j - E_i}{\hbar}.$$

This gives a version of the Einstein relationship

$$\Delta E = \hbar\omega.$$

4 Particles in piecewise constant potentials

A piecewise constant potential is one which takes constant values over finite ranges and changes discontinuously at a finite number of points. An example is illustrated in Fig. IV.4.7.

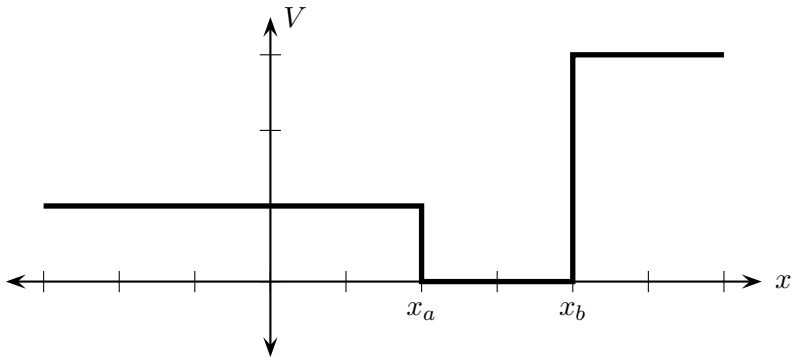


Figure IV.4.7: A piecewise constant potential.

In this example

$$V(x) = \begin{cases} V_1 & \text{for } x \leqslant x_a \\ V_2 = 0 & \text{for } x_a \leqslant x \leqslant x_a \\ V_3 & \text{for } x_b \leqslant x \end{cases}$$

where $V_1 < V_2$. The task is to determine the energy eigenstates, which satisfy the TISE. The strategy for solving this makes use of the fact that whenever V is constant then the TISE,

$$-\frac{\hbar^2}{2m}\frac{\partial^2\phi_E(x)}{\partial x^2} + V(x)\phi_E(x) = E\phi_E(x)$$

reduces to

$$\frac{\partial^2\phi_E(x)}{\partial x^2} = \frac{2m}{\hbar^2}\,(V-E)\,\phi_E(x). \tag{IV.4.78}$$

This is a second order differential equation with constant coefficients and the solutions to it are well known. It follows that it will be possible to obtain separate solutions in each region in which V is constant. For the example of Fig. IV.4.7, this will take the form,

$$\phi_E(x) = \begin{cases} \phi_{E1}(x) & \text{for } x \leqslant x_a \\ \phi_{E2}(x) & \text{for } x_a \leqslant x \leqslant x_a \\ \phi_{E3}(x) & \text{for } x_b \leqslant x \end{cases} \tag{IV.4.79}$$

where the subscript indicates the energy E which is the same for all pieces and a label to indicate the region in which each piece applies. Thus, for $x \leqslant x_a$,

$$\frac{\partial^2\phi_{E1}(x)}{\partial x^2} = \frac{2m}{\hbar^2}\,(V_1-E)\,\phi_{E1}(x),$$

and so on. Taking this approach, it is possible to find solutions for any possible energy E. However, the solutions are constrained as follows. At any boundary between the two regions of distinct constant potentials:

1. the wavefunction must be continuous and
2. the spatial derivative of the wavefunction must be continuous.

In the example of Fig. IV.4.7, this would imply that

$$\phi_{E1}(x_a) = \phi_{E2}(x_a)$$
$$\left.\frac{d\phi_{E1}}{dx}\right|_{x=x_a} = \left.\frac{d\phi_{E2}}{dx}\right|_{x=x_a}$$

with a similar condition at x_b. These conditions, together with the fact that $\phi_E(x) \to 0$ as $x \to \pm\infty$ will severely restrict the possible values of E.

The general form of the TISE for each region is

$$\frac{\partial^2 \phi_{Ej}}{\partial x^2} = \frac{2m}{\hbar^2}\,(V_j - E)\,\phi_{Ej} \tag{IV.4.80}$$

where j labels the region. There are two separate cases to consider: either $V_j > E$ or $V_j < E$ and each gives a distinct type of solution to the differential equation.

4.1 Solution for regions where $V_j < E$

If $V_j < E$ then Eq. (IV.4.80) takes the form

$$\frac{\partial^2 \phi_{Ej}}{\partial x^2} = -k^2 \phi_{Ej} \tag{IV.4.81}$$

where

$$k = \sqrt{\frac{2m(E - V_j)}{\hbar^2}}. \tag{IV.4.82}$$

The solutions to Eq. (IV.4.81) are

$$\phi_E(x) = A\cos kx + B\sin kx \tag{IV.4.83}$$

where A and B are constants. A key feature of this solution is that the spatial frequency k increases as the difference between E and V_j increases. Without considering behavior at the edges of the constant potential region, solutions of this type exist for any value E.

In the context of the example of Fig. IV.4.7 for any value of $E > V_3$, such solutions exist in each region. As an initial candidate, consider the function plotted in Fig. IV.4.8. In each region the solution is an oscillating function as required by Eq. (IV.4.83). The spatial frequency of oscillation is determined by Eq. (IV.4.82) and the plot demonstrates this qualitatively.

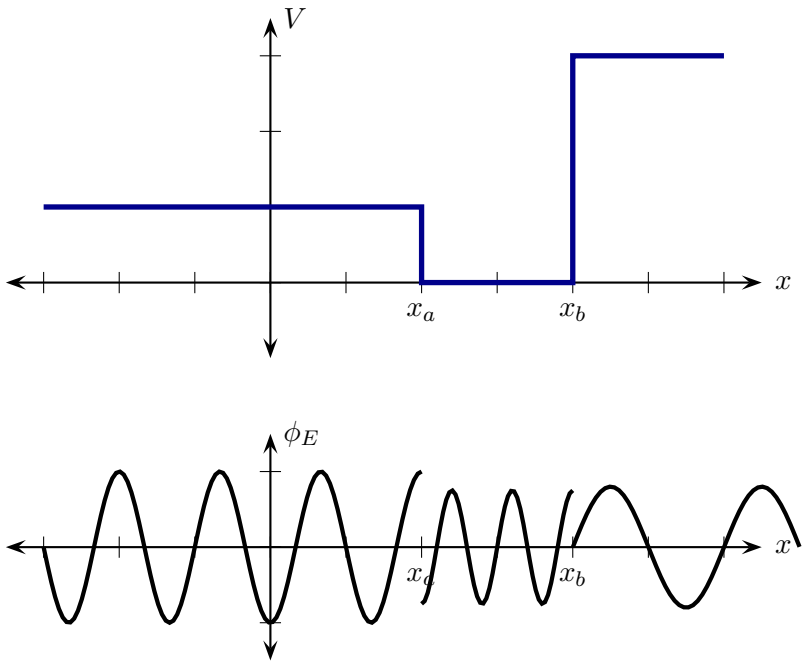


Figure IV.4.8: An initial trial for an eigenstate where $E > V_2$ for piecewise constant potential. The solution demonstrates the dependence of the spatial frequency on the difference between the total and potential energy in each region but does not satisfy the matching conditions at the edges of the regions.

This does not satisfy the matching conditions at the boundaries of the regions. This would require for the solutions of regions 1 and 2,

$$\phi_{E1}(x_a) = \phi_{E2}(x_a)$$
$$\left.\frac{d\phi_{E1}}{dx}\right|_{x=x_a} = \left.\frac{d\phi_{E2}}{dx}\right|_{x=x_a}.$$

To attain such a solution requires adjusting the amplitude and phase (but not the spatial frequency) of the the solution for region 2. It turns out that this fixes the phase and amplitude of the solution in region 2, relative to those of the solution in region 1. The solution of Fig IV.4.9 demonstrates this.

The matching conditions at x_b, require that

$$\phi_{E2}(x_b) = \phi_{E3}(x_b)$$
$$\left.\frac{d\phi_{E2}}{dx}\right|_{x=x_b} = \left.\frac{d\phi_{E3}}{dx}\right|_{x=x_b}.$$

This fixes the phase and amplitude of the solution in region 3 relative to those of region 2 and hence of region 1. This is illustrated in Fig. IV.4.10. A solution of this form can be

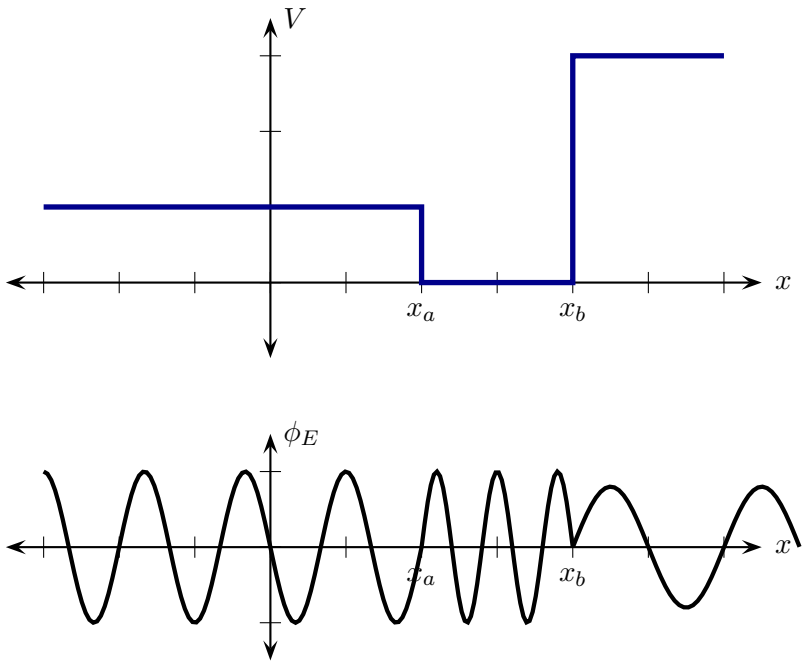


Figure IV.4.9: An improved trial for an eigenstate where $E > V_3$ for piecewise constant potential. The solution demonstrates the dependence of the spatial frequency on the difference between the total and potential energy in each region and satisfies the matching conditions at x_a but not those at x_b.

found for any $E > V_j$ for all regions. The matching conditions at the region boundaries merely determine the relative amplitudes of the solutions in each region.

4.2 Solution for regions where $V_j > E$

If $V_j > E$ then Eq. (IV.4.80) takes the form

$$\frac{\partial^2 \phi_{Ej}}{\partial x^2} = \kappa^2 \phi_{Ej} \tag{IV.4.84}$$

where

$$\kappa = \sqrt{\frac{2m(V_j - E)}{\hbar^2}}. \tag{IV.4.85}$$

The solutions to Eq. (IV.4.81) are

$$\phi_E(x) = Ae^{\kappa x} + Be^{-\kappa x} \tag{IV.4.86}$$

where A and B are constants. A key feature of this solution is that κ increases as the difference between E and V_j increases. Without considering behavior at the edges of the constant potential region, solutions of this type exist for any value E.

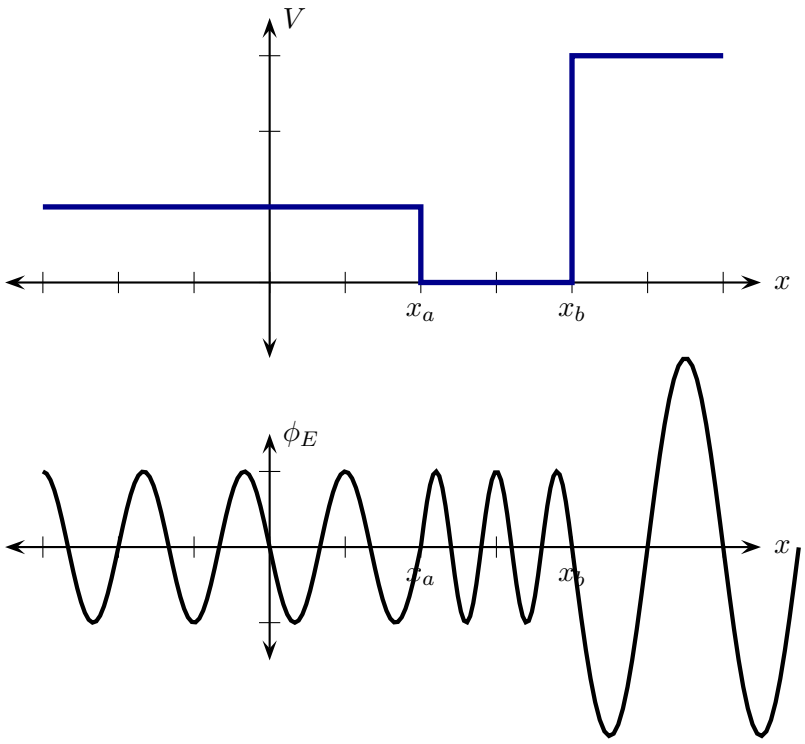


Figure IV.4.10: A solution for an eigenstate where $E > V_3$ for piecewise constant potential. The solution demonstrates the dependence of the spatial frequency on the difference between the total and potential energy in each region and satisfies the matching conditions at x_a and at x_b.

As an example, consider the potential of Fig. IV.4.7 where $V_1 < E < V_3$. Then in regions 1 and 2, the solutions will have an oscillating form and the matching conditions can be applied as before. Now in region 3, the solution must approach 0 as $x \to \infty$. This eliminates the term $Ae^{\kappa x}$ from Eq. (IV.4.86). Thus, in region 3,

$$\phi_E(x) = Be^{-\kappa x}$$

The best approach to constructing a solution here is to proceed from right to left. Again, consider any trial value for energy, E. This determines κ and k for each region. The solution in region 3 determines ϕ_{E2} and its spatial derivative at x_b; this, in turn, determine the amplitude and phase of ϕ_{E2}, which ultimately determines the amplitude and phase of ϕ_{E1} via matching conditions at x_a. This is illustrated in Fig. IV.4.11. Such solutions exist for any value of $E > V_1$. An important feature of this eigenstate is that the position probability density is non-zero for all values of $x > x_b$. This is the region where $V > E$ and in a classical mechanical system, the probability of locating the particle in this region is zero. In a quantum system the particle is able to *penetrate the barrier.*

A more interesting case arises when $0 < E < V_1$. Here the solution is an oscillating

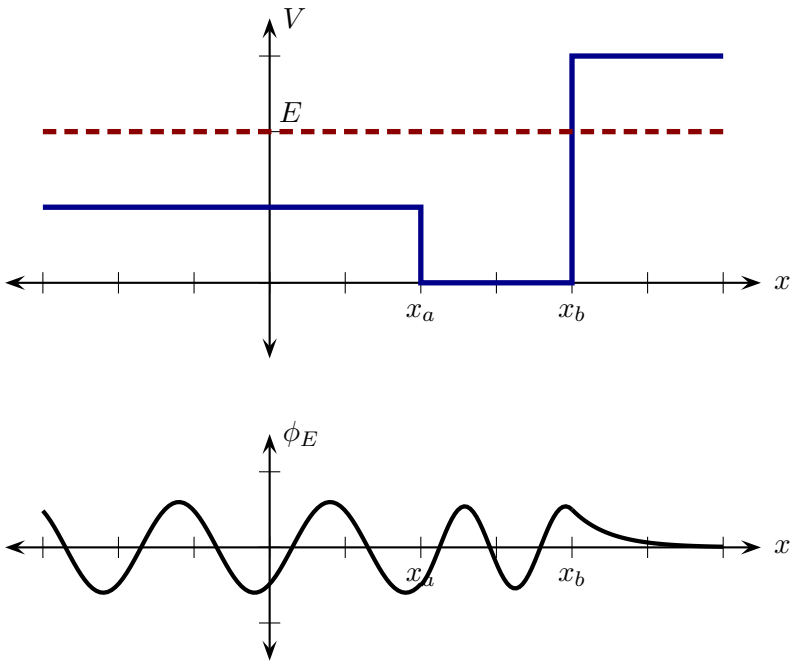


Figure IV.4.11: A solution for an eigenstate where $V_1 < E < V_3$ for piecewise constant potential. The solution demonstrates the dependence of the spatial frequency on the difference between the total and potential energy in each region and satisfies the matching conditions at x_a and at x_b.

function in region 2 but is exponentially decaying in regions 1 and 3. In region 1 $\phi_{E1}(x) \to 0$ as $x \to \infty$. This eliminates the term $Be^{-\kappa x}$ from Eq. (IV.4.86), giving

$$\phi_{E1}(x) = Ae^{\kappa_1 x}$$

in region 1 and similarly

$$\phi_{E3}(x) = Be^{-\kappa_3 x}$$

in region 3. Here

$$\kappa_j = \sqrt{\frac{2m(V_j - E)}{\hbar^2}}.$$

This is larger in region 3 than in region 1. Given a trial value of E, the values of the κ's and k are fixed. The wavefunction can be constructed from the right to the left by initially choosing an arbitrary value for A and applying the matching conditions at x_a. This will fix the amplitude and phase of the oscillating wavefunction in region 2. An example for an arbitrary choice of E is illustrated in Fig. IV.4.12. There is no value for B in

$$\phi_{E3}(x) = Be^{-\kappa_3 x}$$

in region 3 which can possibly satisfy the matching conditions at x_b. Thus there is no solution for this particular value of E.

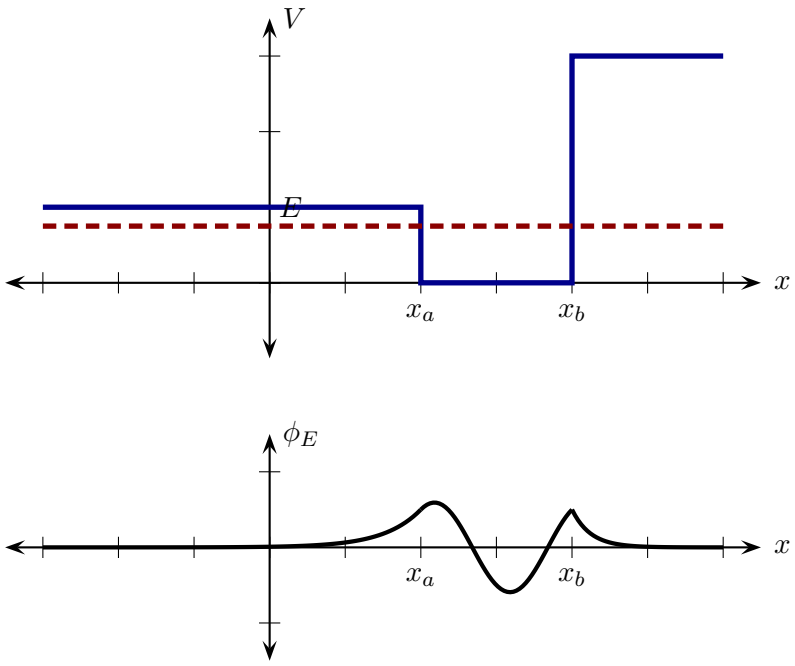


Figure IV.4.12: A trial solution for an eigenstate where $0 < E < V_1$ for a piecewise constant potential. The trial solution demonstrates the dependences of κ and the spatial frequency on the difference between the total and potential energy in each region and satisfies the matching conditions at x_a. However, it does not satisfy the matching conditions at x_b and is thus does not represent an energy eigenstate for this potential.

In such cases the matching conditions severely constrain the possible values of E; only certain very specific values will work. This leads to a discrete set of energy eigenvalues and energy eigenstates. This is quite general and typically if there is a region in which V is lower than for the surrounding regions, the set of energy eigenstates is discrete and often finite in number. On either side of this region, the wavefunction for energy eigenstates decays exponentially and the probability of locating it there is negligible compared to that of locating it within the energy "well." Such states are called *bound states.* A wavefunction for such an eigenstate is illustrated in Fig. IV.4.13.

5 Momentum representation of the particle state

The position wavefunction $\Psi(x,t)$ of a state $|\Psi(t)\rangle$ is a useful for describing position measurement outcomes via the position probability density

$$P(x,t) = |\Psi(x,t)|^2$$

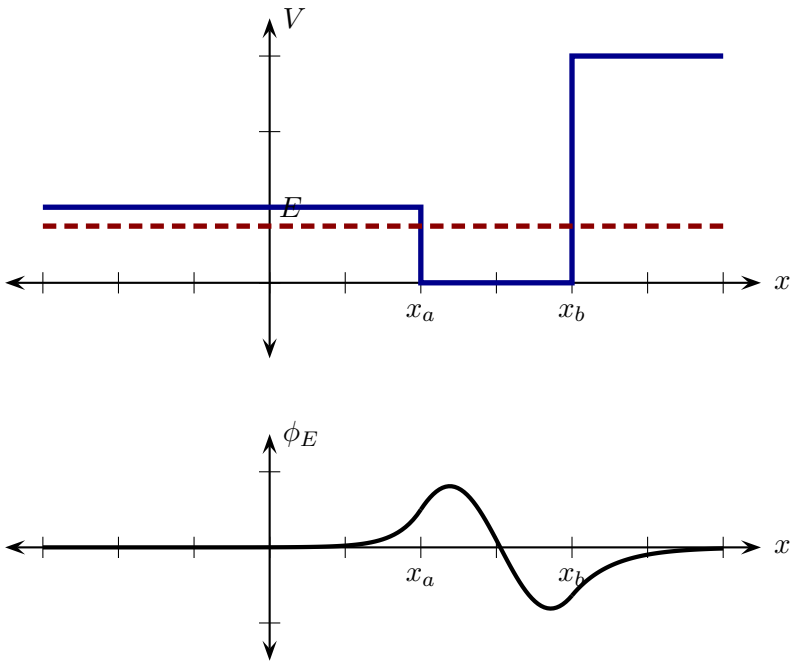


Figure IV.4.13: An eigenstate where $0 < E < V_1$ for a piecewise constant potential. This demonstrates the dependences of κ and the spatial frequency on the difference between the total and potential energy in each region and satisfies the matching conditions at x_a and x_b.

which can be used in probability calculations to give

$$\Pr\left(a \leqslant x \leqslant b\right) = \int_a^b P(x,t)\mathrm{d}x.$$

An equivalent *momentum probability density*, $\tilde{P}(p,t)$, which is possibly a very different type of function from the position probability density for the same state, would play the same role in determining the probability of momentum measurement outcomes via

$$\Pr\left(p_a \leqslant p \leqslant p_b\right) = \int_{p_a}^{p_b} \tilde{P}(p,t)\mathrm{d}p.$$

The issue is to determine an expression for the momentum probability density by determining a *momentum wavefunction*, $\tilde{\Psi}(p,t)$, and to provide a prescription for relating the position and momentum wavefunctions to each other. This is accomplished by considering

that for a particle in the state $|\Psi(t)\rangle$,

$$\begin{aligned}\Pr\left(p_a \leqslant p \leqslant p_b\right) &= \langle\Psi(t)|\left[\int_{p_a}^{p_b} |p\rangle\langle p|\,\mathrm{d}p\right]|\Psi(t)\rangle \\ &= \int_{p_a}^{p_b} \langle\Psi(t)|p\rangle\langle p|\Psi(t)\rangle\,\mathrm{d}p \\ &= \int_{p_a}^{p_b} |\langle p|\Psi(t)\rangle|^2\,\mathrm{d}p \\ &= \int_{p_a}^{p_b} \left|\tilde{\Psi}(p,t)\right|^2\mathrm{d}p\end{aligned}$$

where the momentum wavefunction is defined to be

$$\tilde{\Psi}(p,t) := \langle p|\Psi(t)\rangle\,. \tag{IV.5.87}$$

Thus the momentum probability density is

$$\tilde{P}(p,t) = \left|\tilde{\Psi}(p,t)\right|^2. \tag{IV.5.88}$$

The momentum wavefunction can be determined from the position wavefunction bu invoking the completeness relation for position states and the position wavefunction for momentum states, Eq. (IV.1.58). Thus

$$\begin{aligned}\tilde{\Psi}(p,t) &= \langle p|\Psi(t)\rangle \\ &= \langle p|\left[\int_{-\infty}^{\infty} |x\rangle\langle x|\,\mathrm{d}x\right]|\Psi(t)\rangle \\ &= \int_{-\infty}^{\infty} \langle p|x\rangle\langle x|\Psi(t)\rangle\,\mathrm{d}x \\ &= A\int_{-\infty}^{\infty} e^{-ipx/\hbar}\,\Psi(x,t)\mathrm{d}x\end{aligned}$$

It is conventional to use $A = 1/\sqrt{2\pi\hbar}$ giving,

$$\boxed{\tilde{\Psi}(p,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} e^{-ipx/\hbar}\,\Psi(x,t)\,\mathrm{d}x.} \tag{IV.5.89}$$

In mathematical terms the momentum wavefunction is the Fourier transform of the position wavefunction. It follows that the position wavefunction is the inverse Fourier transform of the position wavefunction,

$$\boxed{\Psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} e^{ipx/\hbar}\,\tilde{\Psi}(p,t)\,\mathrm{d}p.} \tag{IV.5.90}$$

Example: Suppose that the position wavefunction for a particle at $t = 0$ is

$$\Psi(x) = A\frac{1}{(x/x_0)^2 + a^2}$$

where x_0 is a constant with dimensions of position and $a > 0$ is a dimensionless constant. Determine the normalization constant A and the momentum wavefunction. Plot the position and momentum probability distributions.

Answer: Normalization is attained via

$$\langle\Psi|\Psi\rangle = 1$$

which implies that

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} |\Psi(x)|^2 \, \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \frac{|A|^2}{((x/x_0)^2 + a^2)^2} \mathrm{d}x \\ &= |A|^2 \frac{x_0\pi}{2a^3}. \end{aligned}$$

Thus

$$A = \sqrt{\frac{2a^3}{x_0\pi}}.$$

Eq. (IV.5.89) gives the momentum wavefunction. Thus, ignoring the t variable,

$$\begin{aligned} \tilde{\Psi}(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \Psi(x) \, \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi\hbar}} \sqrt{\frac{2a^3}{x_0\pi}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \frac{1}{(x/x_0)^2 + a^2} \, \mathrm{d}x \\ &= \begin{cases} \sqrt{\dfrac{ax_0}{\hbar}} \, e^{apx_0/\hbar} & \text{if } p < 0 \\ \sqrt{\dfrac{ax_0}{\hbar}} \, e^{-apx_0/\hbar} & \text{if } p > 0. \end{cases} \end{aligned}$$

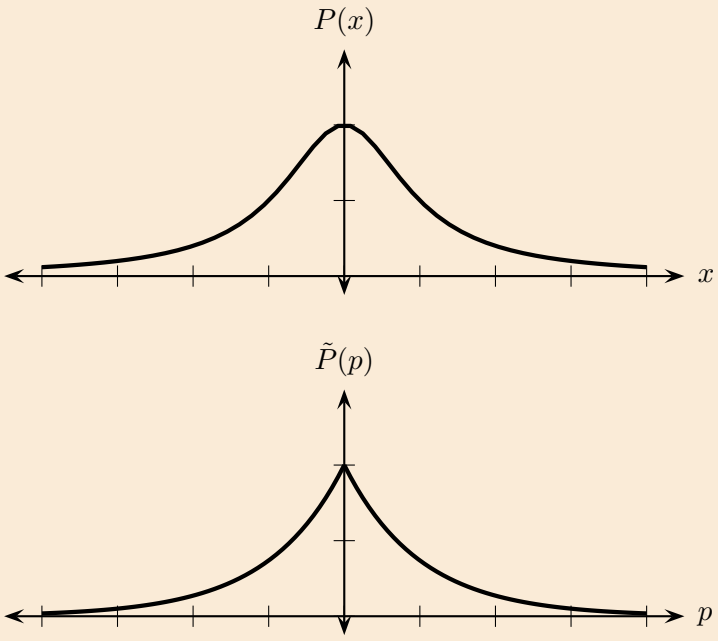
The position probability distribution is

$$P(x) = |\Psi(x)|^2 ,$$

with the momentum distribution given

$$\tilde{P}(p,t) = \left|\tilde{\Psi}(p,t)\right|^2 .$$

Plots of these are as illustrated.



A notable feature in the example above is that as a decreases, the position probability density narrows while the momentum probability density widens. Thus, for such wavefunctions, as the position measurement tends toward giving a single precise value, momentum measurements tend to give outcomes that are more and more evenly distributed over a wider range of values.

5.1 Momentum expectation values

Given the momentum probability distribution the expectation value of momentum can be computed via

$$\langle p \rangle = \int_{-\infty}^{\infty} p\tilde{P}(p,t)\mathrm{d}p. \tag{IV.5.91}$$

Alternatively, for an ensemble of particles, each in state $|\Psi\rangle$,

$$\begin{aligned}
\langle p \rangle = \langle \Psi | \hat{p} | \Psi \rangle &= \langle \Psi | \left[\int_{-\infty}^{\infty} |x\rangle \langle x| \, \mathrm{d}x \right] \hat{p} | \Psi \rangle \\
&= \int_{-\infty}^{\infty} \langle \Psi | x \rangle \langle x | \hat{p} | \Psi \rangle \, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \Psi^*(x) \left(-i\hbar \frac{\partial \Psi}{\partial x} \right) \mathrm{d}x \\
&= -i\hbar \int_{-\infty}^{\infty} \Psi^*(x) \frac{\partial \Psi}{\partial x} \, \mathrm{d}x
\end{aligned} \tag{IV.5.92}$$

where the second last line follows from Eq. (IV.1.57).

Example: Suppose that an ensemble of particles, each in the state represented by the wavefunction

$$\Psi(x) = \sqrt{\frac{2a^3}{x_0\pi}} e^{ip_0x/\hbar} \frac{1}{(x/x_0)^2 + a^2}$$

where p_0 is a constant with units of momentum, $x_0 > 0$ is a constant with units of position and a is a dimensionless constant. Determine the momentum wavefunction and use Eq. (IV.5.91) to determine the momentum expectation value. Verify that Eq. (IV.5.92) gives the same result.

Answer: First, Eq. (IV.5.89) gives

$$\begin{aligned}
\tilde{\Psi}(p) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \Psi(x)\, \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi\hbar}} \sqrt{\frac{2a^3}{x_0\pi}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} e^{ip_0x/\hbar} \frac{1}{(x/x_0)^2 + a^2}\, \mathrm{d}x \\
&= \begin{cases} \sqrt{\dfrac{ax_0}{\hbar}}\, e^{a(p-p_0)x_0/\hbar} & \text{if } p < p_0 \\ \sqrt{\dfrac{ax_0}{\hbar}}\, e^{-a(p-p_0)x_0/\hbar} & \text{if } p > p_0. \end{cases}
\end{aligned}$$

Then Eq. (IV.5.91) gives

$$\begin{aligned}
\langle p \rangle &= \int_{-\infty}^{\infty} p\tilde{P}(p,t)\mathrm{d}p \\
&= \int_{-\infty}^{p_0} p \left| \sqrt{\frac{ax_0}{\hbar}}\, e^{a(p-p_0)x_0/\hbar} \right|^2 \mathrm{d}p + \int_{p_0}^{\infty} p \left| \sqrt{\frac{ax_0}{\hbar}}\, e^{a(p-p_0)x_0/\hbar} \right|^2 \mathrm{d}p \\
&= p_0.
\end{aligned}$$

Alternatively, using Eq. (IV.5.92)

$$\begin{aligned}
\langle p \rangle &= -i\hbar \int_{-\infty}^{\infty} \Psi^*(x) \frac{\partial \Psi}{\partial x}\, \mathrm{d}x \\
&= -i\hbar \frac{2a^3}{x_0\pi} \int_{-\infty}^{\infty} e^{-ip_0x/\hbar} \frac{1}{(x/x_0)^2 + a^2} \frac{\partial}{\partial x} \left(e^{ip_0x/\hbar} \frac{1}{(x/x_0)^2 + a^2} \right) \mathrm{d}x \\
&= -i\hbar \frac{2a^3}{x_0^3\pi\hbar} \int_{-\infty}^{\infty} \frac{ip_0x^2 + ip_0a^2x_0^2 - 2x\hbar}{((x/x_0)^2 + a^2)^3} \mathrm{d}x \\
&= -i\hbar \frac{2a^3}{x_0^3\pi\hbar} \left[\frac{ip_0\pi x_0^3}{8a^3} + \frac{3ip_0\pi x_0^3}{8a^3} \right] \\
&= p_0.
\end{aligned}$$

The uncertainty in momentum satisfies

$$\Delta p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2}.$$

requires computation of $\langle p^2 \rangle$. If the momentum wavefunction is known, then

$$\langle p^2 \rangle = \int_{-\infty}^{\infty} p^2 \tilde{P}(p,t) \mathrm{d}p. \tag{IV.5.93}$$

On the other hand if the position wavefunction is known, then a derivation along the lines of that leading to Eq. (IV.5.92) gives

$$\langle p^2 \rangle = -\hbar^2 \int_{-\infty}^{\infty} \Psi^*(x) \frac{\partial^2 \Psi}{\partial^2 x} \, \mathrm{d}x. \tag{IV.5.94}$$

This is often expressed as

$$\langle p^2 \rangle = \int_{-\infty}^{\infty} \Psi^*(x) \left(i\hbar \frac{\partial}{\partial x} \right)^2 \Psi(x) \mathrm{d}x \tag{IV.5.95}$$

and, in general,

$$\langle p^n \rangle = \int_{-\infty}^{\infty} \Psi^*(x) \left(i\hbar \frac{\partial}{\partial x} \right)^n \Psi(x) \mathrm{d}x \tag{IV.5.96}$$

where

$$\left(i\hbar \frac{\partial}{\partial x} \right)^n \Psi(x) = (i\hbar)^n \frac{\partial^n \Psi}{\partial x^n}.$$

Exercise: Show that for an ensemble of particles, each in the state represented by the wavefunction

$$\Psi(x) = \sqrt{\frac{2a^3}{x_0 \pi}} \frac{1}{(x/x_0)^2 + a^2}$$

where $x_0 > 0$ is a constant with units of position and a is a dimensionless constant,

$$\Delta p = \frac{\hbar}{\sqrt{2} x_0 a}.$$

Show that

$$\Delta x = x_0 a.$$

5.2 Position and momentum uncertainty relation

The previous example indicates that, for a particle whose state is described by the wavefunction

$$\Psi(x) = \sqrt{\frac{2a^3}{x_0 \pi}} \frac{1}{(x/x_0)^2 + a^2},$$

neither a position nor a momentum measurement yields one outcome with certainty. Clearly, for $x_0 \to 0$ or $a \to 0$ the position uncertainty approaches zero but the momentum uncertainty becomes infinite. In the former limit, the position probability density distribution

approaches that of a delta function centered at $x = 0$, indicating that a position measurement outcome will yield 0 with certainty. However, the momentum wavefunction approaches a constant function and a momentum measurement can yield any one of a vast range of outcomes. This raises the question of whether there are any states of a quantum system for which both position and momentum measurements would *both* yield definite outcomes.

The answer to this requires consideration of the commutator of the position and momentum operators.

Theorem: For the position and momentum operators

$$\boxed{\left[\hat{x}, \hat{p}\right] = i\hbar\hat{I}.} \tag{IV.5.97}$$

Proof: In terms of position wavefunctions

$$\begin{aligned}
\left[\hat{x}, \hat{p}\right]|\Psi\rangle &= \hat{x}\hat{p}\,|\Psi\rangle - \hat{p}\hat{x}\,|\Psi\rangle \\
&\leftrightarrow x\left(-i\hbar\frac{\partial}{\partial x}\right)\Psi(x) - \left(-i\hbar\frac{\partial}{\partial x}\right)x\Psi(x) \\
&= -i\hbar x\frac{\partial\Psi}{\partial x} + i\hbar\left(\Psi(x) + x\frac{\partial\Psi}{\partial x}\right) \\
&= i\hbar\Psi(x) \\
&\leftrightarrow i\hbar\,|\Psi\rangle
\end{aligned}$$

where the product rule for differentiation was used in the third last line. Since this is true for any state $|\Psi\rangle$, the statement of the theorem follows. •

An immediate consquence of Eq. (IV.5.97) is that there is not a complete set of states which are simultaneously position and momentum eigenstates. In fact, applying the uncertainty relation, Eq. (II.3.76) shows that, for any state of the quantum system,

$$\Delta x\Delta p \geqslant \frac{\hbar}{2}. \tag{IV.5.98}$$

Thus there is no state which is such that either both a position eigenstate and a momentum eigenstate.

Exercise: Show that for a particle whose state is described by the wavefunction

$$\Psi(x) = \sqrt{\frac{2a^3}{x_0\pi}}\frac{1}{(x/x_0)^2 + a^2},$$

the uncertainties in position and momentum satisfy the position-momentum uncertainty relation. Show that they exceed the lower bound required by Eq. (IV.5.98).

6 Wavepackets

The position momentum uncertainty relation implies that it is impossible to construct a wavefunction which is represents a state in which the particle has both a precise position and momentum. In many cases, such as the energy eigenstates for a particle in an infinite well, the wavefunctions are such that the position probability distribution is substantially spread over a large region. For such a wavefunction there is no sense in trying to ascribe a specific location or even a narrow region of locations to the particle. However, it is possible to construct functions that do represent a situation in which a particle's location is restricted to a narrow region. An example is the wavefunction

$$\Psi(x) = \sqrt{\frac{2a^3}{x_0\pi}}\frac{1}{(x/x_0)^2 + a^2},$$

for which $\Delta x = x_0 a$. Choosing $x_0 = 1$, the uncertainty in position can be reduced by reducing a. This describes a particle whose location is increasingly precise and which can be regarded as *localized*. However the product of the position and the momentum uncertainties still exceeds the bound established by the uncertainty principle. As an alternative candidate consider a *Gaussian wavepacket*, for which the wavefunction, at one instant in time, has the form

$$\Psi(x) = Ae^{-(x-x_0)/2a^2} \tag{IV.6.99}$$

where x_0 and a are real constants and A is a normalization constant. Clearly $\Psi(x) > 0$ for all values of x. For any value of x, the exponent is less than or equal to zero and, as $x \to \pm\infty$, $\Psi(x) \to 0$. The maximum value of $\Psi(x)$ is attained when the exponent is minimized. This occurs when $x = x_0$. Graphs of Gaussian wavepackets are illustrated in Fig. IV.6.14.

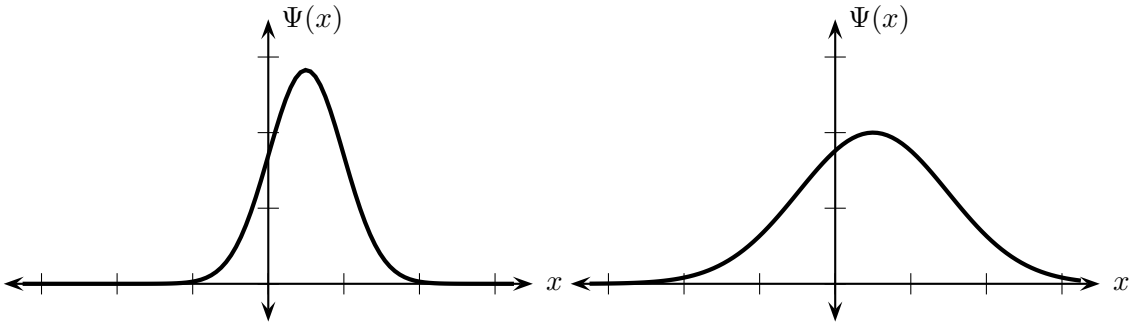


Figure IV.6.14: Gaussian wavepackets where $x_0 = 0.5$. The graph on the left has $a = 1/2$ and that on the right $a = 1$.

The plots illustrate the role of the constant a; as a increases the width of the wavepacket increases.

It will frequently be necessary to compute integrals of such Gaussian functions over all values of x. The following general result is useful in this context.

Theorem: If the real part of α is positive then

$$\boxed{\int_{-\infty}^{\infty} e^{-\alpha x^2+\beta x+\gamma}\mathrm{d}x = e^{\gamma}\, e^{\beta^2/4\alpha}\,\sqrt{\frac{\pi}{\alpha}}.} \tag{IV.6.100}$$

Proof: Later... •

Exercise: Show that for the state corresponding to

$$\Psi(x) = Ae^{-(x-x_0)/2a^2},$$

the normalization requirement implies that

$$A = \left(\frac{1}{\pi a^2}\right)^{1/4}.$$

Expectation values an uncertainties require computation of similar integrals, with the modification that the integrand contains an additional factor of x or x^2. Some integrals of this form can be evaluated by employing the symmetry of the integrand. However, generally useful results in this context are given by the following.

Theorem: If the real part of α is positive then,

$$\boxed{\int_{-\infty}^{\infty} xe^{-\alpha x^2+\beta x+\gamma}\mathrm{d}x = e^{\gamma}\, e^{\beta^2/4\alpha}\,\frac{\beta}{2}\,\sqrt{\frac{\pi}{\alpha^3}}} \tag{IV.6.101}$$

and

$$\boxed{\int_{-\infty}^{\infty} x^2e^{-\alpha x^2+\beta x+\gamma}\mathrm{d}x = e^{\gamma}\, e^{\beta^2/4\alpha}\,\frac{\beta^2+2\alpha}{4}\,\sqrt{\frac{\pi}{\alpha^5}}.} \tag{IV.6.102}$$

Proof: Later... •

Exercise: Show that, for the state corresponding to

$$\Psi(x) = \left(\frac{1}{\pi a^2}\right)^{1/4} e^{-(x-x_0)/2a^2},$$

the the expectation value of position measurements is $\langle x\rangle = x_0$ and the uncertainty in position measurement outcomes is $\Delta x = a/\sqrt{2}$.

It follows that the parameter a, which appears in the Gaussian wavepacket, quantifies the uncertainty in position measurement outcomes.

The probabilities of various momentum measurement outcomes can be assessed by determining the momentum wavefunction. A remarkable fact is that the momentum wavefunction has a very similar functional form to the position wavefunction. Here applying Eq. (IV.5.89) gives

$$\tilde{\Psi}(p) = \left(\frac{a^2}{\pi\hbar^2}\right)^{1/4} e^{-ipx_0/\hbar}\, e^{-p^2a^2/2\hbar}.$$

Exercise: Prove that, for a system whose position wavefunction is

$$\Psi(x) = \left(\frac{1}{\pi a^2}\right)^{1/4} e^{-(x-x_0)/2a^2},$$

the momentum wavefunction is given by

$$\tilde{\Psi}(p) = \left(\frac{a^2}{\pi\hbar^2}\right)^{1/4} e^{-ipx_0/\hbar}\, e^{-p^2a^2/2\hbar}.$$

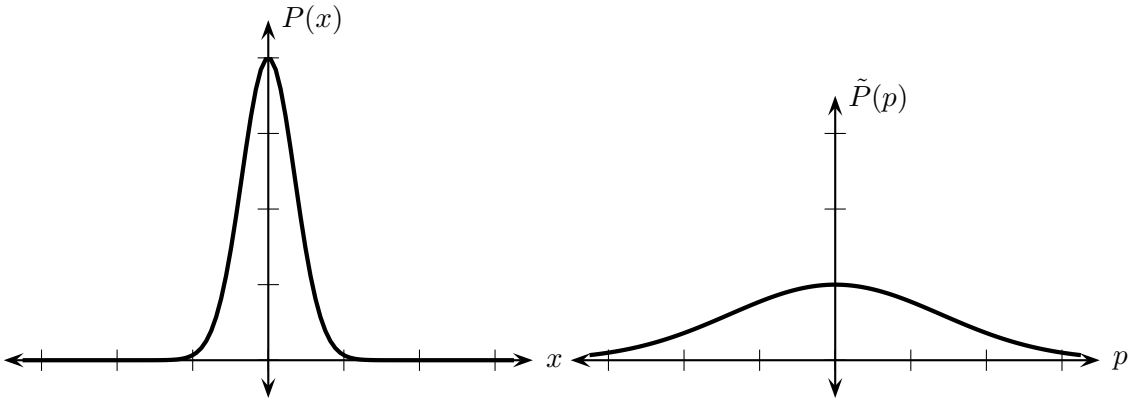
Show that $\langle p\rangle = 0$ and $\Delta p = \hbar/a\sqrt{2}$.

These results show that, as the parameter a decreases, the likelihood that position measurements return outcomes close to the position expectation value increases while the likelihood that momentum measurements return outcomes close to the momentum expectation value decreases. Thus as a decreases, the wavefcuntion decscribes a particle whose position is increasingly localized but whose momentum is decreasingly localized. This is required by the position-momentum uncertainty relation; in fact the product for such Gaussian wavepackets satisfies the lower bound, i.e.

$$\Delta x \Delta p = \frac{\hbar}{2}.$$

Graphs of the position and momentum probability densities for Gaussian wavepackets are provided in Fig. IV.6.15. These clearly illustrate the trade-off in uncertainties between position and momentum.

Probability densities for $a = 1/2$.



Probability densities for $a = 2$.

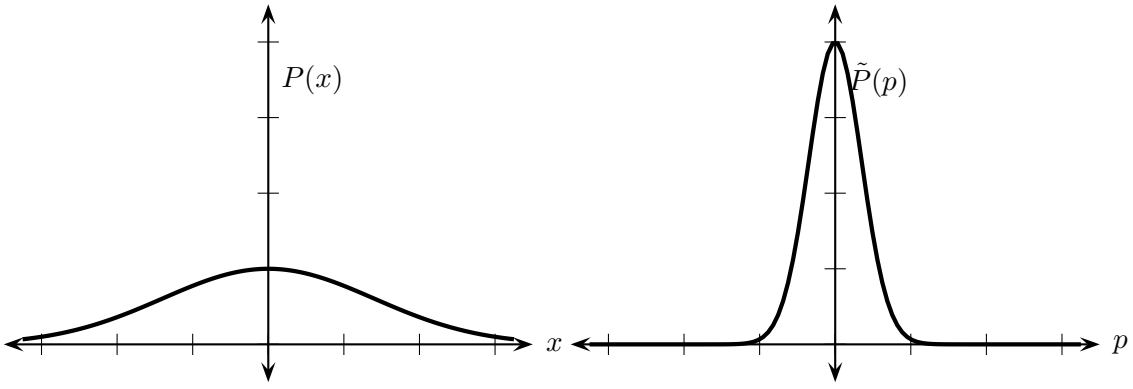


Figure IV.6.15: Probability densities for Gaussian wavepackets where $x_0 = 0$. The position probability distribution is given by $P(x) = |\Psi(x)|^2$ and the momentum probability distribution by $\tilde{P}(x) = |\tilde{\Psi}(x)|^2$.

V. Particles in Central Potentials and Angular Momentum

1 Angular Momentum

The three operators corresponding to components of orbital angular momentum operators, $\hat{L}_x, \hat{L}_y, \hat{L}_z$ satisfy the following commutation relations:

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z \tag{V.1.1a}$$

$$\left[\hat{L}_y, \hat{L}_z\right] = i\hbar\hat{L}_x \tag{V.1.1b}$$

$$\left[\hat{L}_z, \hat{L}_x\right] = i\hbar\hat{L}_y. \tag{V.1.1c}$$

Additionally each of these commutes with the angular momentum squared operator

$$\hat{\mathbf{L}}^2 := \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \tag{V.1.2}$$

These algebraic rules are identical to those that are satisfied by the spin operators, $\hat{S}_x, \hat{S}_y, \hat{S}_z$, for spin-1/2 particles and for this reason spin can be regarded as a type of angular momentum. Generally any set of three operators $\hat{J}_x, \hat{J}_y, \hat{J}_z$, that satisfy the commutation relations,

$$\left[\hat{J}_x, \hat{J}_y\right] = i\hbar\hat{J}_z \tag{V.1.3a}$$

$$\left[\hat{J}_y, \hat{J}_z\right] = i\hbar\hat{J}_x \tag{V.1.3b}$$

$$\left[\hat{J}_z, \hat{J}_x\right] = i\hbar\hat{J}_y, \tag{V.1.3c}$$

correspond to some form of angular momentum, which can be denoted $\mathbf{J}$. In order to determine the possible outcomes and statistics of these for any type of measurement of angular momentum, it is necessary to determine the eigenstates and eigenvalues of the operators $\hat{J}_x, \hat{J}_y$ and $\hat{J}_z$. The fact that the three operators do not commute implies that it is impossible to find a complete set eigenstates that are simultaneously eigenstates for more than one of these; thus it is necessary to focus of one the operators, typically $\hat{J}_z$. The aim is to determine the possible eigenstates and eigenvalues of this. Note that, by symmetry, the eigenvalues of $\hat{J}_x$ will be the same as those of $\hat{J}_z$. However, the eigenstates will differ for the two. The magnitude of the angular momentum is conveniently represented by the *magnitude of angular momentum squared observable,*

$$\hat{\mathbf{J}}^2 := \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2. \tag{V.1.4}$$

It follows from Eqs. (V.1.3) that

$$\left[\hat{\mathbf{J}}^2, \hat{J}_i\right] = 0 \tag{V.1.5}$$

for $i = x, y, z$.

Exercise: Show that Eqs. (V.1.3) imply that $\left[\hat{\mathbf{J}}^2, \hat{J}_z\right] = 0$.

One consequence of this is that it is possible to find simultaneous eigenstates of both $\hat{\mathbf{J}}^2$ and $\hat{J}_z$. These will be denoted $|j,m\rangle$ where j labels the outcome of a measurement of $\mathbf{J}^2$ and m an outcome of a measurement of J_z. Specifically the eigenstate satisfy

$$\hat{\mathbf{J}}^2 |j,m\rangle = \Lambda_j |j,m\rangle$$
$$\hat{J}_z |j,m\rangle = \lambda_m |j,m\rangle$$

where Λ_j and λ_m are the outcomes of measurements of $\mathbf{J}^2$ and J_z on a particle in the state $|j,m\rangle$. From a mathematical point of view it is convenient to rephrase the eigenvalue problem so as to involve dimensionless quantities; in this case it is important to realize that $\hbar$ has units of angular momentum. Thus define

$$A_j := \frac{\Lambda_j}{\hbar^2}$$

and

$$B_m := \frac{\lambda_m}{\hbar}$$

which gives the following eigenvalue equations.

$$\hat{\mathbf{J}}^2 |j,m\rangle = \hbar^2 A_j |j,m\rangle \tag{V.1.6}$$
$$\hat{J}_z |j,m\rangle = \hbar B_m |j,m\rangle . \tag{V.1.7}$$

The physical meaning of these state $|j,m\rangle$, is stated in terms of outcomes of measurements; this is illustrated in Fig. V.1.1. For such a state is it possible to state that $\mathbf{J}^2 = \hbar^2 A_j$ and $J_z = \hbar B_m$.

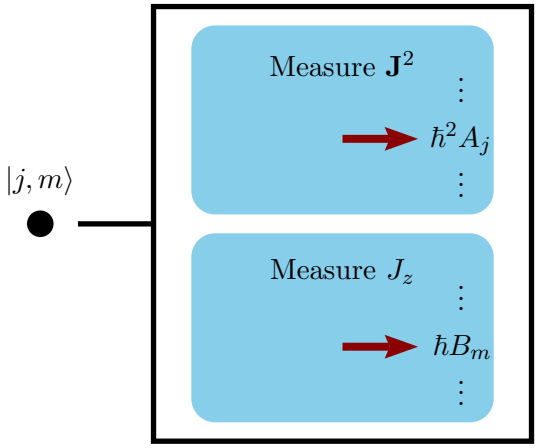


Figure V.1.1: Schematic illustration of the physical meaning of the state $|j,m\rangle$. The device measures the two physical quantities $\mathbf{J}^2$ and J_z. For each of these one and only one outcome is attained with certainty.

The process of determining values for A_j and B_m has three distinct parts:

1. Determine bounds on A_j. These give bounds on possible values for $\mathbf{J}^2$.

2. For a given value of A_j (and thus a given value of $\mathbf{J}^2$) determine bounds on B_m. These gives bounds on J_z for given possible values of $\mathbf{J}^2$.

3. Establish ladders of values for B_m for each possible value of A_j. These fix the possible values of J_z for a given value of $\mathbf{J}^2$.

1.1 Bounds on the Magnitude of Total Angular Momentum

In classical mechanics for any system the magnitude squared of the angular momentum must satisfy $\mathbf{J}^2 > 0$ since the dot product of any vector with itself is necessarily positive. The following theorem establishes an analogous result for quantum systems.

Theorem: For any eigenstate of $\hat{\mathbf{J}}^2$ the corresponding eigenvalue $\hbar^2 A_j$ satisfies

$$\boxed{A_j \geqslant 0.} \tag{V.1.8}$$

Proof: Consider any state $|\Psi\rangle$. Then

$$\langle\Psi|\,\hat{\mathbf{J}}^2\,|\Psi\rangle = \langle\Psi|\,\hat{J}_x^2\,|\Psi\rangle + \langle\Psi|\,\hat{J}_y^2\,|\Psi\rangle + \langle\Psi|\,\hat{J}_z^2\,|\Psi\rangle\,.$$

Now

$$\langle\Psi|\,\hat{J}_x^2\,|\Psi\rangle = \langle\Psi|\,\hat{J}_x^\dagger \hat{J}_x\,|\Psi\rangle$$

since $\hat{J}_x^\dagger = \hat{J}_x$. Thus $\langle\Psi|\,\hat{J}_x^2\,|\Psi\rangle$ is merely the inner product of the state $\hat{J}_x\,|\Psi\rangle$ with itself. This inner product is never negative. Thus

$$\langle\Psi|\,\hat{J}_x^2\,|\Psi\rangle \geqslant 0.$$

Similar results apply for the other components and thus

$$\langle\Psi|\,\hat{\mathbf{J}}^2\,|\Psi\rangle \geqslant 0$$

for any state $|\Psi\rangle$. For an eigenstate $|j,m\rangle$,

$$\langle j,m|\,\hat{\mathbf{J}}^2\,|j,m\rangle = \hbar^2 A_j\,\langle j,m|j,m\rangle = \hbar^2 A_j$$

and thus $\hbar^2 A_j \geqslant 0$. This proves that $A_j \geqslant 0$. •

It follows that for any quantum system in any state, a measurement of the magnitude squared of angular momentum gives $\mathbf{J}^2 \geqslant 0$.

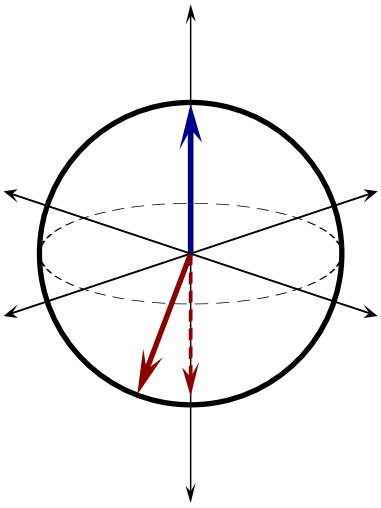


Figure V.1.2: Depiction of angular momentum vectors in classical mechanics. The radius of the sphere is equal to the magnitude of the angular momentum $|\mathbf{J}|$. The tip of the angular momentum vector, $\mathbf{J}$, indicated by a solid arrow, lies on the surface of the sphere. The z-component, J_z is indicated by a dashed arrow. It is possible for the angular momentum vector to lie on any point of the surface of the sphere. In classical mechanics it is possible that $J_z = |\mathbf{J}|$ as is indicated by the blue vector.

1.2 Bounds on the a Single Component of Angular Momentum

Consider a classical system with a given total angular momentum $|\mathbf{J}|$. Geometrically the tip of the angular momentum vector $\mathbf{J}$ must lie on a sphere of radius $|\mathbf{J}|$. In classical mechanics any such vector represents a possible angular momentum for this system. It follows that the z-component of angular momentum must satisfy $-|\mathbf{J}| \leqslant J_z \leqslant |\mathbf{J}|$ and that there are instances where $J_z = \pm|\mathbf{J}|$. This is illustrated geometrically in Fig. V.1.2.

An analogous statement for a quantum system would involve the eigenvalues $\hbar^2 A_j$ and $\hbar B_m$. Here $|\mathbf{J}| = \hbar\sqrt{A_j}$ and $J_z = \hbar B_m$. Does it follow that $-\hbar\sqrt{A_j} \leqslant \hbar B_m \leqslant \hbar\sqrt{A_j}$ and is it possible that $\hbar B_m = \pm\hbar\sqrt{A_j}$? Equivalently is $-\sqrt{A_j} \leqslant B_m \leqslant \sqrt{A_j}$ and is it possible that $B_m = \pm\sqrt{A_j}$? With regard to the latter question note that for a quantum system which has a definite value for J_z typically there is no definite value for J_x or J_y, since $\left[\hat{J}_z, \hat{J}_x\right] \neq 0$, etc,.... Note that if $J_z^2 = \mathbf{J}^2$, this would imply that $J_x = J_y = 0$ and thus suggest that these both have definite values along with that of J_z. This is a contradiction. A formal proof of this, using the algebraic properties of the angular momentum operators requires the following lemmas and intermediate results.

An initial step towards determining the relationship between $|\mathbf{J}|$ and J_z for quantum mechanical systems involves the expectation values of $\mathbf{J}^2$ and J_z for systems in the simultaneous eigenstate $|j, m\rangle$. Here,

$$\left\langle \mathbf{J}^2 \right\rangle = \hbar^2 A_j^2 = \mathbf{J}^2$$

since $|j,m\rangle$ is an eigenstate of $\hat{\mathbf{J}}^2$ with eigenvalue $\hbar^2 A_j^2$ and thus measuring $\mathbf{J}^2$ yields $\hbar^2 B_m^2$ with certainty. By a similar argument measuring J_z yields $\hbar B_m$ with certainty; thus measuring J_z^2 yields $\hbar^2 B_m^2$ with certainty. this implies that

$$\left\langle J_z^2 \right\rangle = \hbar^2 B_m^2 = J_z^2.$$

Then

$$\begin{aligned}\left\langle \mathbf{J}^2 \right\rangle &= \left\langle J_x^2 + J_y^2 + J_z^2 \right\rangle \\ &= \left\langle J_z^2 \right\rangle + \left\langle J_y^2 \right\rangle + \left\langle J_z^2 \right\rangle \\ &= \left\langle J_z^2 \right\rangle + \left\langle J_y^2 \right\rangle + J_z^2\end{aligned}$$

which implies that for an ensemble of systems in the state $|j,m\rangle$ the angular momentum measurement outcomes are related via

$$\mathbf{J}^2 = \left\langle J_z^2 \right\rangle + \left\langle J_y^2 \right\rangle + J_z^2. \tag{V.1.9}$$

Further insight into this relationship is provided by the following result.

Lemma: For any eigenstate of $\hat{J}_z$

$$\boxed{\langle J_x \rangle = \langle J_y \rangle = 0.} \tag{V.1.10}$$

Proof: Note that

$$\langle j,m| \left[\hat{J}_y, \hat{J}_z\right] |j,m\rangle = i\hbar \langle j,m| \hat{J}_x |j,m\rangle = i\hbar \langle J_x \rangle .$$

Thus

$$\langle j,m| \left(\hat{J}_y \hat{J}_z - \hat{J}_z \hat{J}_y\right) |j,m\rangle = i\hbar \langle J_x \rangle .$$

This gives

$$\langle j,m| \hat{J}_y \hat{J}_z |j,m\rangle - \langle j,m| \hat{J}_z \hat{J}_y |j,m\rangle = i\hbar \langle J_x \rangle .$$

Now

$$\hat{J}_z |j,m\rangle = \hbar B_m |j,m\rangle$$

and taking the complex conjugate transpose of this gives

$$\langle j,m| \hat{J}_z = \hbar B_m \langle j,m| .$$

Combining gives

$$\hbar B_m \langle j,m| \hat{J} |j,m\rangle - \hbar B_m \langle j,m| \hat{J} |j,m\rangle = i\hbar \langle J_x \rangle ,$$

which gives $\langle J_x \rangle = 0$. •

This result immediately simplifies calculations of uncertainties for angular momentum components. An exercise shows that for a system that yields a definite value for J_z,

$$\boxed{(\Delta J_x)^2 = \left\langle J_x^2 \right\rangle} \tag{V.1.11}$$

and

$$\boxed{(\Delta J_y)^2 = \left\langle J_y^2 \right\rangle .} \tag{V.1.12}$$

Exercise: Prove Eqs. (V.1.11) and (V.1.12).

Thus Eq. (V.1.9) gives

$$\mathbf{J}^2 = (\Delta J_x)^2 + (\Delta J_y)^2 + J_z^2. \tag{V.1.13}$$

Note that $(\Delta J_x)^2 + (\Delta J_y)^2 = \left\langle J_x^2 + J_y^2 \right\rangle$. This clearly implies that for a quantum system

$$-|\mathbf{J}| \leqslant J_z \leqslant |\mathbf{J}|$$

but that $J_z = \pm|\mathbf{J}|$ if and only if both the uncertainties in J_x and J_y are zero. The impossibility of this is determined by applying the uncertainty principle

$$\begin{aligned}
\Delta J_x \Delta J_y &\geqslant \frac{1}{2} \left| \langle j,m | \left[\hat{J}_x, \hat{J}_y \right] | j,m \rangle \right| \\
&= \frac{1}{2} \left| \langle j,m | \, i\hbar \hat{J}_z \, | j,m \rangle \right| \\
&= \frac{1}{2} \, |i\hbar J_z| \\
&= \frac{\hbar^2}{2} \, |J_z| \, .
\end{aligned}$$

An immediate consequence of this is that $J_z = \pm|\mathbf{J}|$ is only possible (but not necessarily guaranteed) if $J_z = 0$. In most cases both $(\Delta J_x)^2 \neq 0$ and $(\Delta J_y)^2 \neq 0$. For these, J_z is strictly less than the magnitude of the total angular momentum. Geometrically this also implies that the angular momentum state cannot be represented by a single vector (this would imply definite values for J_x and J_y). However it can be represented by a cone of vectors, whose tips lie on the surface of one sphere and all of which have the same z component. This is indicated in Fig. V.1.3.

In terms of eigenvalues Eq. (V.1.13) implies

$$\hbar^2 A_j = (\Delta J_x)^2 + (\Delta J_y)^2 + \hbar^2 B_m^2 \tag{V.1.14}$$

or, equivalently,

$$B_m^2 = A_j - \frac{1}{\hbar^2} \left[(\Delta J_x)^2 + (\Delta J_y)^2 \right] . \tag{V.1.15}$$

The issue is how close could J_z approach to $\pm|\mathbf{J}|$, which is equivalent to asking how close B_m could approach to $\pm\sqrt{A_j}$. Eq. (V.1.15) can be used to establish a bound by minimizing the term $(\Delta J_x)^2 + (\Delta J_y)^2$ subject to the constraint that $\Delta J_x \Delta J_y \geqslant \hbar^2 B_m/2$. This is easily done geometrically by envisioning a plane which represents all possible values

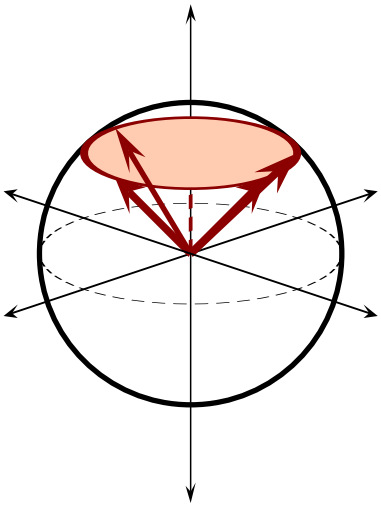


Figure V.1.3: Depiction of an angular momentum eigenstate $|j,m\rangle$ in quantum mechanics. The radius of the sphere is equal to the magnitude of the angular momentum $|\mathbf{J}|$. The radius of the disk at the top of the cone is $(\Delta J_x)^2 + (\Delta J_y)^2 = \langle J_x^2 + J_y^2 \rangle$.

of ΔJ_x and ΔJ_y. The constraint imposed by the uncertainty principle requires that allowed values of ΔJ_x and ΔJ_y lie beyond the hyperbolas traced by $\Delta J_x \Delta J_y = \hbar^2 B_m/2$. What minimizes $(\Delta J_x)^2 + (\Delta J_y)^2$ is the point in this region closest to the origin. This is attained when $\Delta J_x = \Delta J_y$ on the hyperbola, i.e. $\Delta J_x = \Delta J_y = \hbar\sqrt{|B_m|/2}$. This implies that $(\Delta J_x)^2 + (\Delta J_y)^2 \geqslant \hbar\,|B_m|$ and it follows that

$$B_m^2 \leqslant A_j - |B_m|\,.$$

An algebraic manipulation of this gives the bound

$$\boxed{\frac{1}{2} - \sqrt{A_j + \frac{1}{4}} \leqslant B_m \leqslant -\frac{1}{2} + \sqrt{A_j + \frac{1}{4}}.} \tag{V.1.16}$$

Exercise: Using $B_m^2 \leqslant A_j - |B_m|$ show that Eq. (V.1.16) is true.

In terms of outcomes of angular momentum measurements, multiplying every term in Eq. (V.1.16) by $\hbar$, implies that

$$\frac{\hbar}{2} - \sqrt{\mathbf{J}^2 + \frac{\hbar^2}{4}} \leqslant J_z \leqslant -\frac{\hbar}{2} + \sqrt{\mathbf{J}^2 + \frac{\hbar^2}{4}}. \tag{V.1.17}$$

Note that $\mathbf{J}^2 + \hbar^2/4 = (|\mathbf{J}| + \hbar/2)^2 - \hbar|\mathbf{J}|$. Thus

$$\frac{\hbar}{2} - \sqrt{\left(|\mathbf{J}| + \frac{\hbar}{2}\right)^2 - \hbar|\mathbf{J}|} \leqslant J_z \leqslant -\frac{\hbar}{2} + \sqrt{\left(|\mathbf{J}| + \frac{\hbar}{2}\right)^2 - \hbar|\mathbf{J}|}. \tag{V.1.18}$$

Again this shows that it is not possible that $J_z = \pm|\mathbf{J}|$ unless $|\mathbf{J}| = 0$.

The algebra required to find the eigenvalues is simplified by defining positive $a_j \geqslant 0$ such that

$$A_j = a_j(a_j + 1). \tag{V.1.19}$$

Then Eq. (V.1.16) implies that

$$\boxed{-a_j \leqslant B_m \leqslant a_j.} \tag{V.1.20}$$

Exercise: Show that Eq. (V.1.20) is true.

The process of finding values for a_j hinges on finding those for B_m. This is facilitated by the *angular momentum raising operator,*

$$\boxed{\hat{J}_+ := \hat{J}_x + i\hat{J}_y,} \tag{V.1.21}$$

and *lowering operator*

$$\boxed{\hat{J}_+ := \hat{J}_x - i\hat{J}_y.} \tag{V.1.22}$$

The important algebraic properties of these are provided by the following theorem.

Theorem: The angular momentum raising and lowering operators satisfy.

$$\hat{J}_\pm^\dagger = \hat{J}_\mp \tag{V.1.23}$$

$$\left[\hat{J}_z, \hat{J}_\pm\right] = \pm\hbar\hat{J}_\pm \tag{V.1.24}$$

$$\left[\hat{\mathbf{J}}^2, \hat{J}_\pm\right] = 0 \tag{V.1.25}$$

$$\left[\hat{J}_+, \hat{J}_-\right] = 2\hbar\hat{J}_z \tag{V.1.26}$$

$$\hat{\mathbf{J}}^2 = \hat{J}_+\hat{J}_- + \hat{J}_z^2 - \hbar\hat{J}_z \tag{V.1.27}$$

$$\hat{\mathbf{J}}^2 = \hat{J}_-\hat{J}_+ + \hat{J}_z^2 + \hbar\hat{J}_z. \tag{V.1.28}$$

Proof: Exercise. •

An immediate and important consequence of these captured by the following theorem.

Theorem: Suppose that $|j, m\rangle$ is a simultaneous eigenstate of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$, i.e.

$$\hat{\mathbf{J}}^2 |j, m\rangle = \hbar^2 a_j(a_j + 1) |j, m\rangle$$
$$\hat{J}_z |j, m\rangle = \hbar B_m |j, m\rangle .$$

Then $\hat{J}_+ |j, m\rangle$ is an eigenstate of $\hat{\mathbf{J}}^2$ with eigenvalue $\hbar^2 a_j(a_j+1)$ and of $\hat{J}_z$ with eigenvalue $\hbar(B_m+1)$. Similarly, $\hat{J}_- |j, m\rangle$ is an eigenstate of $\hat{\mathbf{J}}^2$ with eigenvalue $\hbar^2 a_j(a_j+1)$ and of $\hat{J}_z$ with eigenvalue $\hbar(B_m-1)$.

Proof: Consider

$$|\Phi\rangle := \hat{J}_+ |j,m\rangle .$$

Then Eq. (V.1.25) implies that

$$\hat{\mathbf{J}}^2 \hat{J}_+ |j,m\rangle - \hat{J}_+ \hat{\mathbf{J}}^2 |j,m\rangle = 0.$$

Thus

$$\hat{\mathbf{J}}^2 |\Phi\rangle = \hat{J}_+ \hbar^2 a_j (a_j + 1) |j,m\rangle$$

giving

$$\hat{\mathbf{J}}^2 |\Phi\rangle = \hbar^2 a_j (a_j + 1) |\Phi\rangle .$$

Thus $|\Phi\rangle = \hat{J}_+ |j,m\rangle$ is an eigenstate of $\hat{\mathbf{J}}^2$ with eigenvalue $\hbar^2 a_j(a_j+1)$. Now Eq. (V.1.24) implies that

$$\hat{J}_z \hat{J}_+ |j,m\rangle - \hat{J}_+ \hat{J}_z |j,m\rangle = \hbar \hat{J}_+ |j,m\rangle$$

which gives

$$\begin{aligned} \hat{J}_z |\Phi\rangle &= \hbar |\Phi\rangle + \hbar B_m |\Phi\rangle \\ &= \hbar (B_m + 1) |\Phi\rangle . \end{aligned}$$

This proves that $|\Phi\rangle = \hat{J}_+ |j,m\rangle$ is an eigenstate of $\hat{J}_z$ with eigenvalue $\hbar(B_m+1)$. A similar derivation yields the analogous results for the lowering operator. •

The angular momentum raising and lowering operators are used to generate multiple angular momentum eigenstates, starting from just one eigenstate. Suppose that j *and thus the magnitude squared of the angular momentum is fixed.* The given one angular momentum eigenstate $|j,m\rangle$, it is possible to repeatedly produce ladder of distinct angular momentum eigenstates via $\ldots, \hat{J}_-^2 |j,m\rangle , \hat{J}_- |j,m\rangle , |j,m\rangle , \hat{J}_+ |j,m\rangle \hat{J}_+^2 |j,m\rangle , \ldots$. This is depicted schematically in Fig. V.1.4. The next question is to determine the endpoints of each ladder for each value of j. Suppose that the ladder is finite. Thus there exists a state at the top of the ladder corresponding to $m = m_{\max}$ and a state at the bottom of the ladder corresponding to $m = m_{\min}$. The eigenvalues for these satisfy according to Eq. (V.1.20),

$$-a_j \leqslant B_{m_{\min}} \leqslant B_{m_{\max}} \leqslant -a_j.$$

Now if $B_{m_{\max}}$ is at the top of the ladder, then

$$a_j - 1 < B_{m_{\max}} \leqslant a_j,$$

otherwise applying the raising operator would give an eigenvalue of $\hat{J}_z$ that is still less than $\hbar a_j$. Similarly

$$-a_j \leqslant B_{m_{\min}} < -a_j + 1,$$

Additionally

$$\hat{J}_+ |j, m_{\max}\rangle = 0 \tag{V.1.29}$$

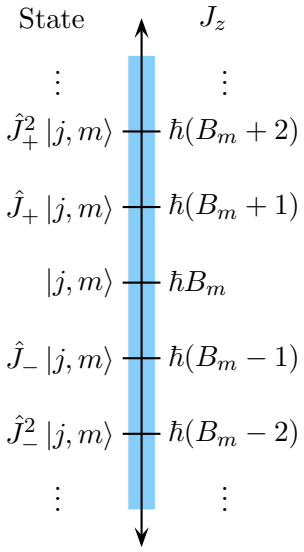


Figure V.1.4: A ladder of angular momentum eigenstates for fixed j.

since if this were not true, then the state on the left would give eigenvalue of $\hat{J}_z$ greater than $\hbar a_j$. A similar argument gives

$$\hat{J}_- |j, m_{\min}\rangle = 0. \tag{V.1.30}$$

The following theorem establishes a symmetrical arrangement of eigenvalues about $J_z = 0$.

Theorem: For angular momentum eigenstates

$$\begin{aligned} B_{m_{\min}} &= -a_j \\ B_{m_{\max}} &= a_j. \end{aligned}$$

Proof: First Eq. (V.1.28) implies

$$\begin{aligned} \hat{\mathbf{J}}^2 |j, m_{\max}\rangle &= \left(\hat{J}_-\hat{J}_+ + \hat{J}_z^2 + \hbar\hat{J}_z\right) |j, m_{\max}\rangle \\ &= \hat{J}_-\hat{J}_+ |j, m_{\max}\rangle + \hat{J}_z^2 |j, m_{\max}\rangle + \hbar\hat{J}_z |j, m_{\max}\rangle \\ &= \hbar^2 B_{m_{\max}}^2 |j, m_{\max}\rangle + \hbar^2 B_{m_{\max}} |j, m_{\max}\rangle \\ &= \hbar^2 B_{m_{\max}} (B_{m_{\max}} + 1) |j, m_{\max}\rangle \end{aligned}$$

This gives

$$\hbar^2 a_j(a_j+1) |j, m_{\max}\rangle = \hbar^2 B_{m_{\max}} (B_{m_{\max}} + 1) |j, m_{\max}\rangle$$

and acting on this with $\langle j, m_{\max}|$ gives

$$a_j(a_j+1) = B_{m_{\max}} (B_{m_{\max}} + 1). \tag{V.1.31}$$

Straightforward algebraic manipulations yield

$$B_{m_{\max}} = a_j \quad \text{or} \quad B_{m_{\max}} = -a_j - 1.$$

The latter is clearly impossible and thus

$$B_{m_{\max}} = a_j$$

A similar argument using Eq. (V.1.27) yields the result for $B_{m_{\min}}$. •

This theorem implies that every ladder of angular momentum eigenstates must terminate at the same maximum and minimum endpoints. It follows that every ladder gives the same set of eigenvalues for $\hat{J}_z$. Thus an integral number of applications, N, of the raising operator starting at the state with $\hbar B_{m_{\min}} = -\hbar a_j$ yields the state with maximum value $\hbar B_{m_{\max}} = \hbar a_j$. Thus

$$\hbar a_j = -\hbar a_j + \hbar N$$

which implies that

$$a_j = \frac{N}{2}.$$

Additionally the possible values for B_m are $-a_j, -a_j + 1, \ldots a_j - 1, a_j$. At this point it is customary to use the notation

$$a_j \to j$$

and

$$B_m \to m.$$

In these terms the angular momentum eigenstates and eigenvalues are described by.

The angular momentum eigenstates are $|j, m\rangle$ and satisfy

$$\hat{\mathbf{J}}^2 |j, m\rangle = \hbar^2 |j, m\rangle$$
$$\hat{J}_z |j, m\rangle = \hbar m |j, m\rangle$$

where the possible values of j are

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$$

and for each j the possible values of m are

$$m = -j, -j + 1, \ldots, j - 1, j.$$

(V.1.32)