Notes for Quantum Mechanics
Version 1.3 — April 14, 2013

Recent changes:First version posted.

1   Introduction

The material below provides a guide to some of the topics covered in class but not discussed in the textbook. Additional discussion will be added as needed *or requested*. We will focus on the mathematical description of quantum mechanics leaving the discussion of the experimental observation of quantum phenomena to that given in your textbook. While our discussion of complex vector spaces is intended to be self-contained, you may find it helpful to read portions of one of the books on linear algebra recommended on the course website.

1.1   State vectors

We begin by considering a spinning object with angular momentum. We assume that this has a fixed magnitude, represented by some fixed value of $\vec{J}^2$ and a z-component momentum J_z. Just as in classical physics we can often fix $\vec{J}^2$ by saying that we are working with a state of definite energy. The different values of J_z should arise from different orientations of the total angular momentum $\vec{J}$.

We begin by asserting that in quantum theory the z-component of angular momentum cannot take on values limited only by the magnitude of $\vec{J}$. Instead J_z is allowed only particular discrete values $\hbar m$. Here $\hbar$ is Planck's constant $h = 6.626\,10^{-27}$ erg-sec divided by 2π and m is an integer or an integer plus $1/2$, *i.e.* a half-integer. The "quantum number" m can take all integer-spaced values between $-j$ and $+j$ where, again, j is an integer or half-integer. The number of allowed values of m is thus $2j + 1$, the number of integers (or half-integers) lying between $-j$ and j including the end values $\pm j$ which are allowed. The simplest case of $j = 1/2$ thus corresponds to two allowed values $m = \pm 1/2$. For the case $j = 2$ five values are allowed: -2, -1, 0, 1 and 2.

Of even greater consequence than this peculiar "quantization" of the allowed values of J_z, is the hypothesis that to completely specify a particular physical situation for this spinning object we need not specify which value of m the system possesses. Instead we must describe a general situation in which there is some probability that J_z for this spinning object may have any one of these $2j+1$ values. These probabilities are specified by giving for each allowed value of m a complex amplitude ψ_m whose modulus squared, $|\psi_m|^2$ gives the probability that a measurement of J_z will return the value $\hbar m$. The collection of $2j+1$ complex numbers represent by the complex vector $(\psi_{-j}, \psi_{-j+1}, \psi_{-j+2}, \ldots, \psi_{j-1}, \psi_j)$ is said to precisely specify the state of the spinning object and is referred to as its complex "state vector".

The description matches well the amazing phenomena observed in the experiment of Stern and Gerlach described in sections 10.1-10.4 of the French and Taylor book. The Cesium atoms in that experiment can be viewed as spinning objects with known energy, likely the lowest energy possible for Cesium. They posses a magnetic moment $\vec{\mu}$ which is proportional to their angular momentum $\vec{J}$. When passing through a region of magnetic field aligned in the z-direction which increases in strength with increasing z (an inhomogenous field), those atoms with J_z positive will be drawn upward into the region of stronger magnetic field. After the beam of atoms passes through this region of inhomgenous field we expect the beam to be spread out: atoms with $J_z > 0$ will be deflected upward and those with $J_z < 0$ downward. For a classical system with continuous values of J_z this would be a smooth distribution centered at no deflection and with those atoms with the largest $|J_z|$ deflected the most. However, for Cesium, with $j = 1/2$, the experiment shows the atoms roughly deflected into two beams with $J_z = \pm\hbar/2$! If each Cesium atom passing through the Stern-Gerlach apparatus where prepared in the quantum state $(\psi_{-1/2}, \psi_{1/2})$ these two peaks would then have heights $|\psi_{\pm 1/2}|^2$ corresponding to the probabilities that the atoms were to be found with $J_z = \pm\hbar/2$.

Of course, for the probability description to make sense we must require that the total probability of any result being found is unity so that the sums of all probabilities must be one:

$$\sum_{m=-j}^{j} |\psi_m|^2 = 1. \tag{1}$$

One final postulate is needed to completely describe measurement. After a measurement is performed and a definite result obtained, for example a

specific half-integer, -1/2 or +1/2 for the case of the Cesium atom, the system changes to reflect this discovery. The state vector after measurement would have all $\psi_m = 0$ for values of m different from that measured and $\psi_m = 1$ for that m corresponding to the measure value. This is known as "the reduction of the wave function" and is a somewhat, unsatisfactory human-based concept of measurement since an observer is needed to decide what result is actually seen!

Thus, to develop the quantum theory of our spinning object we must deal with these $2j+1$-tuples of complex numbers. As suggested above, these make up a complex $2j + 1$-dimensional vector space in exactly the same way that the three spacial components, (x, y, z) are elements of a real, 3-dimensional vector space. We can add two of these $2j + 1$-tuples of complex numbers by simply adding their complex components. Likewise, such a vector can be multiplied by a complex number by multiplying each of the $2j+1$ amplitudes ψ_m by that complex number.

This is a standard example of a complex vector space which can be defined more abstractly as a natural generalization of the real vector spaces we have used in the first semester. In a complex vector space we can multiply a vector by a general complex number. Thus if A and B are complex vectors and c, c_1 and c_2 are complex numbers then $A+B = B+A$ and cC are also complex vectors and these operations obey:

$$
\begin{align}
(c_1 + c_2)A &= c_1 A + c_2 A \tag{2} \\
c(A + B) &= cA + cB \tag{3} \\
cA &= 0 \quad \text{if} \quad c = 0. \tag{4}
\end{align}
$$

where 0 is the zero vector obeying $A + 0 = A$.

Just as for our real vectors describing a particle's classical position, it is important that we can define a positive definite "dot" or "inner" product (A, B). For a complex vector space this dot product is linear in the right-hand argument but "anti-linear" in the left-hand argument obeying

$$
\begin{align}
(A, B + C) &= (A, B) + (A, C) \tag{5} \\
(A, cB) &= c(A, B) \tag{6} \\
(A, B) &= (B, A)^* \tag{7} \\
(A, A) &\geq 0 \tag{8} \\
(A, A) &= 0 \text{ implies that } A = 0, \tag{9}
\end{align}
$$

3

where the final two conditions make this a positive definite inner product. Note, the combination of Eq. 7 and Eq. 6 implies that the inner product is anti-linear in its left argument:

$$(cA, B) = c^*(A, B). \tag{10}$$

While the choice of left instead of right anti-linearity is one of convention, simple linearity in both arguments would be inconsistent with the important requirement that the "length squared" of A: (A, A), be real.

Often in quantum mechanics one uses a "bra-ket" notation:

$$(A, B) = \langle A|B \rangle \tag{11}$$

and the notation $|B\rangle$ for the state B.

Returning to our spinning object, it is conventional to use a more abstract and much more streamlined notation to represent the complete quantum state specified by the $2j+1$ complex numbers $(\psi_{-j}, \psi_{-j+1}, \psi_{-j+2}, \ldots, \psi_{j-1}, \psi_j)$. This is done by introducing $2j+1$ abstract vectors, $\{|m\rangle\}_{-j\leq m\leq j}$ which correspond to states with specific values of J_z and to use the probability ampliutudes ψ_m to construct a general state:

$$|\psi\rangle = \sum_{m=-j}^{j} \psi_m |m\rangle. \tag{12}$$

It is much easier to say "the state $|\psi\rangle$" than to say the "$2j + 1$-tuple of complex numbers $(\psi_{-j}, \psi_{-j+1}, \psi_{-j+2}, \ldots, \psi_{j-1}, \psi_j)$"!

Since the states with quite different values of J_z should be very independent, it is natural to assume that they are orthogonal and convenient to make them orthonormal. Thus, we will assume that these states obey:

$$\langle m'|m \rangle \equiv (|m'\rangle, |m\rangle) = \delta_{m'm} \equiv \begin{cases} 1 \text{ if } m' = m \\ 0 \text{ otherwise} \end{cases}. \tag{13}$$

Just as in the case of the coordinates (x, y, z) and the more abstract vector $\vec{r}$:

$$\vec{r} = x\hat{x} + y\hat{y} + z\hat{z} \tag{14}$$

the complex coordinates $\{\psi_m\}_{-j\leq m\leq j}$ and the vector $|\psi\rangle$ represent the same thing. The $2j + 1$ states $\{|m\rangle\}_{-j\leq m\leq j}$ can be viewed as the basis with which the coordinates $\{\psi_m\}_{-j\leq m\leq j}$ can be used to construct an abstract vector.

4

Finally, we might make the concept of "probability" and its relation to the magnitudes $|\psi_m|^2$ more clear by considering an example. If $j = 1$ then a general quantum state will have three components with

$$
\begin{align}
|\psi\rangle &= \psi_{-1}|-1\rangle + \psi_0|0\rangle + \psi_{+1}|+1\rangle \tag{15}\\
&= \sum_{m=-1,0,+1} \psi_m|m\rangle \tag{16}
\end{align}
$$

and a measurement of J_z can have the three possible values: $-\hbar$, 0 and $+\hbar$. The statement that the probability of finding the result $-\hbar$ for J_z is $|\psi_{-1}|^2$ has the following operational meaning. Assume that the state $|\psi\rangle$ given in Eq. 16 can be prepared exactly the same way N times so that the only uncertainties in our setup are the intrinsic probabilities in the quantum theory itself. Each time we start with precisely the same state $|\psi\rangle$. Then the number of times in which the value $-\hbar$ will be found when J_z is measured will be the product of the number of measurements N times the probability $|\psi_{-1}|^2$ of obtaining $-\hbar$: $N_{-1} = |\psi_{-1}|^2 N$.

If we wanted to compute the average result for J_z from these N measurements we would add the result $\hbar m_i$ for the ith measurement over all measurements and divide by N to get the average:

$$
\langle J_z \rangle = \frac{1}{N} \sum_{i=1}^{N} \hbar m_i. \tag{17}
$$

If we group together those terms with the same value for J_z, this can be written:

$$
\begin{align}
\langle J_z \rangle &= \frac{1}{N}\Big(N_{-1}(-\hbar) + N_0 0 + N_{+1}(+\hbar)\Big) \tag{18}\\
&= |\psi_{-1}|^2(-\hbar) + |\psi_0|^2 0 + |\psi_{+1}|^2(+\hbar). \tag{19}
\end{align}
$$

1.2   Operators

Just as it was awkward to continually refer to the $2j+1$-tuple $(\psi_{-j}, \psi_{-j+1}, \psi_{-j+2}, \ldots, \psi_{j-1}, \psi_j)$ to identify a state, so too it will become inconvenient when discussing the z-component of angular momentum J_z to continually refer to the set of allowed values $-j \le m \le +j$ and the set of vectors $\{|m\rangle\}_{-j \le m \le j}$ corresponding to these allowed values. As we will see, this is much more easily done by specifying a linear operator which acts on our vector space of states.

Just as when we were discussing rotations of our three dimensional coordinate system when considering particle motion in three dimensions or Lorentz transformations in space-time, it is useful in quantum mechanic to consider operators which act linearly on our complex vector space of states. Such operators can act on one state vector and will produce another. An operator O will be a linear transformation if this "mapping" of our complex vector space into itself satisfies the following conditions:

$$
\begin{align}
O\left(|A\rangle + |B\rangle\right) &= O|A\rangle + O|B\rangle & (20)\\
O(c|A\rangle) &= c\,O|A\rangle, & (21)
\end{align}
$$

where $|A\rangle$ and $|B\rangle$ are vectors and c a complex number.

Just as in the case of rotations or Lorentz transformations, a general linear transformation on a complex vector space corresponds to a matrix of complex numbers as soon as a specific basis is introduced. If we continue to use the language appropriate for the $2j + 1$-dimensional space of states for our spinning object, we can completely characterize a linear operator O by determining how it acts on each of the basis of states with definite values of J_z, $\{|m\rangle\}_{-j \leq m \leq j}$. This can be done by giving the result for $O|m\rangle$, again as a sum of the basis vectors:

$$
O|m\rangle = \sum_{m'=-j}^{j} O_{m'm}|m'\rangle. \tag{22}
$$

This equation shows how a $(2j + 1) \times (2j + 1)$ dimensional complex matrix $O_{m'm}$ can be used to construct the linear operator O. We can also use the orthogonality of the states $|m\rangle$ to determine the matrix $O_{m'm}$ from the linear operator O. All we need to do is to take the inner product of the states on the left- and right-hand sides of Eq. 22 with a specific state, $|m''\rangle$:

$$
\begin{align}
\langle m''|O|m\rangle &= \langle m''|\left(\sum_{m'=-j}^{j} O_{m'm}|m'\rangle\right) & (23)\\
&= \sum_{m'=-j}^{j} O_{m'm}\langle m''|m'\rangle & (24)\\
&= O_{m''m}. & (25)
\end{align}
$$

In the second line we use the linearity of the inner product to relate the inner product of the state $|m''\rangle$ with the sum of products of the states $O_{m'm}|m'\rangle$

to the sum of products of the inner product of $|m''\rangle$ with each state. In the third line we use the fact that $\langle m''|m'\rangle$ vanishes unless $m' = m''$ so only that term in the sum over m' survives.

Now we are ready to encode the set of allowed values for J_z and the states possessing those allowed values into a much more convenient linear operator which is typically also called J_z. Since this can be confusing, at least initially, we will add a superscript "op" to distinguish the operator J_z^{op}. The matrix corresponding to this operator for the basis of states $\{|m\rangle\}_{-j \leq m \leq j}$ is a simple diagonal matrix. The element on the diagonal corresponding to the state $|m\rangle$ is the value of J_z carried by that state: $\hbar m$. Thus we have:

$$(J_z)_{m'm} \equiv \langle m'|J_z|m\rangle = \hbar m \delta_{m'm}, \tag{26}$$

where the "Kronecker delta", $\delta_{m'm}$ is defined in Eq. 13. This can be also written as a large matrix

$$J_z^{\text{op}} \sim \begin{pmatrix} j\hbar & 0 & \ldots & 0 \\ 0 & (j-1)\hbar & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & -j\hbar \end{pmatrix}. \tag{27}$$

This linear operator enjoys a special relationship with the allowed values $\hbar m$ of J_z and the states $|m\rangle$ having those specific values. Both Eqs. 26 and 27 imply that

$$J_z^{\text{op}}|m\rangle = \hbar m|m\rangle. \tag{28}$$

The special relationship between J_z^{op} and the state $|m\rangle$ and the number $\hbar m$ can be described by identifying the vector $|m\rangle$ as an *eigenvector* of J_z and the number $\hbar m$ as the *eigenvalue* corresponding the eigenvector $|m\rangle$.

Not only can we use states $\{|m\rangle\}_{-j \leq m \leq j}$ and values $\hbar m$ to construct the operator J_z^{op}. We can also reverse the process. It is possible (although we have not yet discussed how) to start with the operator J_z^{op} and to find the basis of eigenvectors and corresponding eigenvalues of the operator. This basis of eigenvectors is said to "diagonalize" J_z^{op} as is implied by the diagonal character of the matrix found in Eq. 27. As in the case of quantum states, it is much easier to represent the measurement of angular momentum by referring to the operator J_z^{op} than to continually list its eigenvectors and eigenvalues.

2 Rotations

We have now done about as much as we can with this simple system of a single observable, J_z and the $2j+1$ values that it can have. While the experimental consequences of this discussion are remarkable, and the mathematical structure of complex vector spaces and linear operators elaborate, the underlying ideas may be a little simple-minded. The true character of quantum mechanics emerges when we consider the other two components of angular momentum, J_x and J_y. Instead of constructing independent complex vector spaces for each of these quantities which would classically be very independent, instead we will demand that they also correspond to operators on **the same** $2j + 1$-dimensional vector space as that on which J_z^{op} acts!

This situation is most easily recognized if we attempt to construct J_x^{op} and J_y^{op} from J_z^{op} by rotating the latter into the former. Thus, we are lead to consider rotating our quantum state $|\psi\rangle$. This should be something we can do: actively rotate $|\psi\rangle$ around a direction $\hat{n}$ through an angle θ. We can do this by applying a linear transformation to $|\psi\rangle$ just as we can rotate a vector $\vec{r}$ by applying a rotation matrix to its components. Let's call the linear operator which performs this rotation $R(\hat{n}, \theta)$. Note, these rotations simply turn our state so it has a new orientation. This is done without otherwise changing our state. For example, it is not caused to spin by such a rotation. Only sufficient torque is applied to change the direction of angular momentum, not its magnitude. (In fact, if you prefer, you could think of this as simply a change of coordinate system. However, this can be confusing because one must start to think about the coordinate system in our complex vector space which may be too unfamiliar for this "passive" approach to be helpful.)

2.1 Rotations about a fixed direction

The easiest place to start is with rotations about the z-axis since these should not change the value of J_z — it is only the other components which are being altered. Further consider a small rotation through $\delta\theta \ll 1$ and expand in $\delta\theta$:

$$R(\hat{z}, \delta\theta) = I - iK_z\delta\theta + O\left((\delta\theta)^2\right). \qquad (29)$$

where $O\left((\delta\theta)^2\right)$ represents a correction term that can be dropped if $\delta\theta$ is sufficiently small. When $\delta\theta = 0$ there is no rotation at all, so $R(\hat{z}, 0) = I$,

8

the identity operator. The next term in Eq. 39, $-iK_z\delta\theta$ is the first term in a Taylor expansion of $R(\hat{z}, \delta\theta)$ in $\delta\theta$. The factor of $-i$ is written to follow conventions and the operator K_z is called the "generator" of rotations about the z-direction. Since $R(\hat{z}, \delta\theta)$ and hence K_z cannot change the z component of angular momentum and since there is only one state with the value $m\hbar$ we can be sure:

$$K_z|m\rangle = k_m|m\rangle. \tag{30}$$

Thus, the states $|m\rangle$ are also eigenstates of K_z so that J_z and K_z have the same eigenstates! This is a very strong statement and suggests the possibility that J_z and K_z may be essentially the same operator, differing by only a multiplicative constant. (They do have different units.) This is in fact true and we can figure out what that constant is. Thus, let's assume $k_m = \alpha\hbar m$ and try to figure out what α might be.

This can be done by exploiting the very simple structure of the family of rotations about a fixed axis which obey the product rule:

$$R(\hat{z}, \theta_1)R(\hat{z}, \theta_2) = R(\hat{z}, \theta_1 + \theta_2). \tag{31}$$

This simply states the obvious fact that performing two rotations about a fixed axis in sequence should be equivalent to performing a single rotation through the sum of the angles of those individual rotations about the same axis. A very powerful method of exploiting this relation examines the special case where $\theta_1 = \delta\theta$, $\theta_2 = \theta$ and then expands in $\delta\theta$:

$$\begin{aligned} R(\hat{z}, \theta + \delta\theta) &= R(\hat{z}, \delta\theta)R(\hat{z}, \theta) & (32) \\ &= (I - iK_z\delta\theta)\, R(\hat{z}, \theta) & (33) \end{aligned}$$

Moving the "I" term to the left-hand-side:

$$R(\hat{z}, \theta + \delta\theta) - R(\hat{z}, \theta) = -iK_z\delta\theta R(\hat{z}, \theta) \tag{34}$$

and dividing by $\delta\theta$

$$\frac{R(\hat{z}, \theta + \delta\theta) - R(\hat{z}, \theta)}{\delta\theta} = -iK_z R(\hat{z}, \theta) \tag{35}$$

we can then take the limit $\delta\theta \to 0$ and obtain:

$$\frac{dR(\hat{z}, \theta)}{d\theta} = -iK_z R(\hat{z}, \theta). \tag{36}$$

9

This is our usual exponential equation with the easy solution:

$$R(\hat{z}, \theta) = e^{-iK_z\theta}. \tag{37}$$

This solution obeys the initial condition $R(\hat{z}, 0) = I$. It can be easily seen to solve the differential equation 36 by writing the exponential $\exp(-iK_z\theta)$ using its Taylor series and using the same properties of that series which ensured that this solution works for real and for complex numbers.

We can now figure out what the proportionality constant α is. Consider a general z-axis rotation applied to the state $|m\rangle$:

$$R(\hat{z}, \theta)|m\rangle = e^{-iK_z\theta}|m\rangle = e^{-i\alpha\hbar m\theta}|m\rangle. \tag{38}$$

For the last step we have written the exponential of the operator K_z out as a power series and replaced $(K_z)^n|m\rangle$ by $(k_m)^n|m\rangle$ since each operator K_z will act in turn on the state $|m\rangle$ and give the eigenvalue $k_m = \alpha\hbar m$ and then re-summed the series. We can learn two things from Eq. 38. First α must be a real number so that the overall factor $\exp(-i\alpha\hbar m\theta)$ has magnitude one and does not change the probability of having the J_z value $\hbar m$. Second if we consider a rotation through 2π and the case that m is an integer, this rotation will have no effect if α is an integer divided by $\hbar$ because in that case the phase $\exp(-i\alpha\hbar m 2\pi) = 1$. The simplest (and correct) assumption would be $\alpha = 1/\hbar$. For spin-1/2 and half-integral values of m this factor becomes -1 which, while strange (a 2π rotation changes our quantum state), this change is sufficiently innocuous that it is OK.

Thus, we conclude that $K_z = J_z/\hbar$ so that a rotation about the z-direction through a small angle $\delta\theta$ can be written:

$$R(\hat{z}, \delta\theta) = I - iJ_z\delta\theta/\hbar + O\left((\delta\theta)^2\right). \tag{39}$$

where we do not write the extra superscript op on J_z since it should be clear from the context that J_z is an operator.

2.2   Small rotations about perpendicular directions

Now we are ready to consider angular momentum operators different from J_z, specifically J_x and J_y. We will assume that they are related to small rotations about the x and y axes in the same way that we found J_z to be

related to rotations about the z axis. Writing all three similar equations we expect:

$$R(\hat{x}, \delta\theta_x) = I - \frac{i}{\hbar} J_x \delta\theta_x + O(\delta\theta_x^2) \tag{40}$$

$$R(\hat{y}, \delta\theta_y) = I - \frac{i}{\hbar} J_y \delta\theta_y + O(\delta\theta_y^2) \tag{41}$$

$$R(\hat{z}, \delta\theta_z) = I - \frac{i}{\hbar} J_z \delta\theta_z + O(\delta\theta_z^2) \tag{42}$$

We will now attempt to constrain the three operators J_x, J_y and J_z by demanding that the when the above rotations are combined the results agree with a similar combination of familiar rotations in three dimensions.

The essential rotation to consider is what is known as the "commutator" of two of these near-unity rotations, for example, the product:

$$R(\hat{x}, -\delta\theta_x) \cdot R(\hat{y}, -\delta\theta_y) \cdot R(\hat{x}, \delta\theta_x) \cdot R(\hat{y}, \delta\theta_y) \tag{43}$$

where the final two rotations, one through $-\delta\theta_y$ followed by a second through $-\delta\theta_x$, would neatly undo the effects of the first two rotations were they performed in the opposite order. As we will now work out, this combination of four rotations differs from the unit rotation by a single term of order $\delta\theta_x \delta\theta_y$ whose effect is a small rotation about the z axis. Making this identification then puts such a strong constraint on the operators J_x, J_y and J_z that we can actually determine which matrices can appear.

However, to know what should result from such a product of rotations we can do an "experiment" by examining the same product of rotations as they act on a standard 3-dimensional vector $\vec{r}$. Recall that a small rotation about the direction $\hat{n}$ through an angle $\delta\theta$ has the following effect on $\vec{r}$:

$$\vec{r} \to \vec{r} + \delta\theta\hat{n} \times \vec{r} \equiv (I + \delta\theta\hat{n}\times)\vec{r}. \tag{44}$$

The middle step in this sequence simply describes the effect of a small rotation as adding a small term proportional to the cross product of $\hat{n}$ and $\vec{r}$. The right step writes this as an "operation", written inside the curved brackets acting on $\vec{r}$. The operation of the "I" on $\vec{r}$ has no effect and simply returns $\vec{r}$. The second term is also simple, specifying that a cross product of $\hat{n}$ with $\vec{r}$ be evaluated and multiplied by $\delta\theta$. While this last step appears to obscure something that is already well expressed, this odd notation is very handy

if we want to describe the product of four rotations, analogous to those in Eq. 43:

$$(I - \delta\theta_x \hat{x}\times) \cdot (I - \delta\theta_y \hat{y}\times) \cdot (I + \delta\theta_x \hat{x}\times) \cdot (I + \delta\theta_y \hat{y}\times)\vec{r}. \qquad (45)$$

Written this way it is easy to work out the 16 terms that result from these four products to see concretely the effects of such a rotation on $\vec{r}$:

1. The easiest term is the product of the four I operations, giving $I\vec{r} = \vec{r}$, *i.e.* no rotation.

2. The next terms to examine are the four terms that are linear in $\delta\theta$. Two contain $\delta\theta_x$ and two $\delta\theta_y$. Since each pair includes terms with opposite signs, each pair cancels and no terms of order $\delta\theta$ remain in the product.

3. Moving on to terms of order $(\delta\theta)^2$ we can recognize one term of order $(\delta\theta_x)^2$ coming from product of pieces of the first and third term. A second term of order $(\delta\theta_y)^2$ appears from the product of terms in the second and fourth factors. While we are not interested in keeping track of these terms, you may be able to work out that they actually all cancel. However, this cancellation comes from four terms of order $(\delta\theta_x)^2$ and $(\delta\theta_y)^2$ that would appear in each of the rotation operations separately had we expressed them to one higher order in $(\delta\theta)^2$.

4. The final terms of order $\delta\theta_x \delta\theta_y$ are the ones of interest. There are four such terms which can be written as:

$$\delta\theta_x \delta\theta_y \Big(+\hat{x} \times (\hat{y} \times \vec{r}) - \hat{x} \times (\hat{y} \times \hat{x}) - \hat{y} \times (\hat{x} \times \vec{r}) + \hat{x}(\times\hat{y} \times \vec{r}) \Big). \quad (46)$$

Again, working from left to right these come from the product of the first and the second, the first and the fourth, the second and the third and the third and the fourth.

5. There are four additional terms of order $(\delta\theta)^3$ and one of $(\delta\theta)^4$ which we ignore.

The first two in the terms in Eq. 46 cancel and the next two terms can be simplified as:

$$- \hat{y} \times (\hat{x} \times \vec{r}) + \hat{x}(\times\hat{y} \times \vec{r}) \quad = \quad -(\hat{y} \cdot \vec{r})\hat{x} + (\hat{x} \cdot \vec{r})\hat{y} \qquad (47)$$
$$= \quad \hat{z} \times \vec{r}. \qquad (48)$$

12

Thus, keeping the unit and $\delta\theta_x\delta\theta_y$ terms, our product of four rotations acting on the vector $\vec{r}$ has the effect

$$\vec{r} \to \vec{r} + \delta\theta_x\delta\theta_y\hat{z} \times \vec{r}. \tag{49}$$

If we identify the same terms in Eq. 43, we find the effect of those rotations on a state vector $|\psi\rangle$ will be:

$$|\psi\rangle \to |\psi\rangle + \frac{\delta\theta_x\delta\theta_y}{\hbar}J_yJ_x|\psi\rangle - \frac{\delta\theta_x\delta\theta_y}{\hbar}J_xJ_y|\psi\rangle. \tag{50}$$

If we require that the final two terms in this expression amount to exactly the same rotation as discovered in our "experiment" and shown in Eq. 49 we learn

$$\frac{1}{\hbar}\left(J_yJ_x - J_xJ_y\right) = -iJ_z \tag{51}$$

which can be written:

$$J_xJ_y - J_yJ_x = i\hbar J_z. \tag{52}$$

This combination $J_xJ_y - J_yJ_x$ is also called the commutator of J_x with J_y and written $[J_x, J_y]$. Repeating what we have just learned and writing symmetrical equations for the other possible combinations of rotations we can summarize our results as:

$$[J_x, J_y] = J_xJ_y - J_yJ_x = i\hbar J_z \tag{53}$$
$$[J_y, J_z] = J_yJ_z - J_zJ_y = i\hbar J_x \tag{54}$$
$$[J_z, J_x] = J_zJ_x - J_xJ_z = i\hbar J_y. \tag{55}$$

Before we leave this general discussion of rotations, it will be useful to recognize how we can combine the three rotation generators J_x, J_y and J_z to write a formula for rotations about an arbitrary direction through both small and large angles. As before, we start with a rotation through a small angle $\delta\theta$, now about an arbitrary direction $\hat{n}$ and demonstrate, by applying this rotation to a vector $\vec{r}$, that such a rotation can be obtained by composing three separate rotations sequentially about the three directions $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$ through the three angles $n_1\delta\theta$, $n_2\delta\theta$, $n_3\delta\theta$, respectively.

$$\vec{r} \to (I + \delta\theta n_1\hat{e}_1\times)(I + \delta\theta n_2\hat{e}_2\times)(I + \delta\theta n_3\hat{e}_3\times)\vec{r} \tag{56}$$
$$= (I + \delta\theta n_1\hat{e}_1 \times + \delta\theta n_2\hat{e}_2 \times + \delta\theta n_3\hat{e}_3\times)\vec{r} \tag{57}$$
$$= (I + \delta\theta\hat{n}\times)\vec{r} \tag{58}$$
$$= \vec{r} + \delta\theta\hat{n} \times \vec{r}. \tag{59}$$

where we use our earlier notation of Eqs. 44 and 45.

Thus, to first order in $\delta\theta$ the effect of a single rotation through $\delta\theta$ about $\hat{n}$ is the same as that of sequentially performing three separate rotations about the three directions $\hat{e}_1$ $\hat{e}_2$ and $\hat{e}_3$ through the three angles $n_1\delta\theta$, $n_2\delta\theta$, $n_3\delta\theta$, respectively. This same property should be obeyed by the more general rotations acting on our space of quantum states. We therefore require:

$$
\begin{aligned}
R(\hat{n}, \delta\theta) &= R(\hat{e}_1, n_1\delta\theta)R(\hat{e}_2, n_2\delta\theta)R(\hat{e}_3, n_3\delta\theta) & (60) \\
&= (I - \frac{i}{\hbar}J_1 n_1\delta\theta)(I - \frac{i}{\hbar}J_2 n_2\delta\theta)(I - \frac{i}{\hbar}J_3 n_3\delta\theta) & (61) \\
&= I - \frac{i}{\hbar}J_1 n_1\delta\theta - \frac{i}{\hbar}J_2 n_2\delta\theta - \frac{i}{\hbar}J_3 n_3\delta\theta & (62) \\
&= I - \frac{i}{\hbar}\hat{n}\cdot\vec{J}\delta\theta. & (63)
\end{aligned}
$$

Thus, we have learned that the combination of operators $\hat{n}\cdot\vec{J} = J_1 n_1 + J_2 n_2 + J_3 n_3$ acts as a generator for rotations about the general direction $\hat{n}$.

Finally let's try to find a formula for a general rotation about the direction $\hat{n}$ using this new result for the generator of such rotations. Since we can combine rotations about any fixed axis $\hat{n}$ in the same way that we did for rotations about the $\hat{z}$ direction in Eq. 31, we can start with the relation:

$$
R(\hat{n}, \theta_1)R(\hat{n}, \theta_2) = R(\hat{n}, \theta_1 + \theta_2). \tag{64}
$$

We can then go through the same sequence of arguments as we did following Eq. 31 to show that a general rotation can be written as the exponential:

$$
R(\hat{n}, \theta) = e^{-i\hat{n}\cdot\vec{J}\theta/\hbar}. \tag{65}
$$

As in the case for rotations about the $\hat{z}$ direction, this formula has content because the exponential Taylor series with its $1/n!$ multiplying the nth term will converge quickly, at least for any finite dimensional matrix $\hat{n}\cdot\vec{J}$. We will now consider the simplest case of $j = 1/2$.

2.3   Spin-1/2

As is worked out in Appendix A, the three "commutation relations" in Eqs. 53-55 are very strong constraints and completely determine the form

of the matrices J_x, J_y and J_z. One learns that there is a single set of matrices that obey these relations given the dimension $2j + 1$ of the set of basis vector which can rotate into themselves.

However, here we would like a speedier start for our discussion of rotations. For the case $j = 1/2$ it is easy to simply write down the three 2×2 matrices which obey these commutation relations and explicitly check that they actually do. These three 2×2 matrices are traditionally written as $J_i = \frac{\hbar}{2}\sigma^i$ where the three Pauli matrices σ^i are given by:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{66}$$

It is not difficult to see that these matrices obey the commutation relations required of generators of rotations, for example we can explicitly multiply

$$\begin{aligned}
\frac{\hbar}{2}\sigma^1\frac{\hbar}{2}\sigma^2 - \frac{\hbar}{2}\sigma^2\frac{\hbar}{2}\sigma^1 &= \left(\frac{\hbar}{2}\right)^2 \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} \\
&= -i\hbar\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -i\hbar\frac{\hbar}{2}\sigma^3, \tag{67}
\end{aligned}$$

precisely as required by Eq. 53.

As we can see, these matrices are written in the standard basis of eigenfunctions of J_z since the matrix is σ^3 is diagonal. (Here we use the labels x, y, z and 1, 2, 3 interchangeably.) We will call these two basis vectors $|+\frac{1}{2}\rangle_z$ and $|-\frac{1}{2}\rangle_z$ where we have added the final subscript z anticipating that we may want to discuss eigenstates for the other angular momentum operators J_x and J_y. As we described above, the two states $|\pm\frac{1}{2}\rangle_z$ with the corresponding two allowed values for the z-component of angular momentum, $\pm\hbar\frac{1}{2}$ imply the diagonal form and diagonal elements of $J_z = \hbar\frac{1}{2}\sigma^z$. Conversely, given the operator J_z the states $|\pm\frac{1}{2}\rangle_z$ and $\pm\hbar\frac{1}{2}$ are its eigenvectors and eigenvalues.

We can get a better understanding of how this works by considering a new case: the operator J_y and its eigenvectors and eigenvalues. Can we make a similar analysis of angular momentum in the y-direction, finding the eigenstates and eigenvalues of J_y? Now we must seek a state, let's call it $|+\frac{1}{2}\rangle_y$, with the definite value $+\hbar/2$ for J_y. Thus, we require:

$$J_y|+\frac{1}{2}\rangle_y = \lambda|+\frac{1}{2}\rangle_y \tag{68}$$

15

where the constant $\lambda = \hbar/2$. If we write this state $|+\frac{1}{2}\rangle_y$ in terms of our basis:

$$|+\frac{1}{2}\rangle_y = a_{\frac{1}{2}}|\frac{1}{2}\rangle_z + a_{-\frac{1}{2}}|-\frac{1}{2}\rangle_z \tag{69}$$

then Eq. 68 can be written as a matrix equation:

$$\frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{\frac{1}{2}} \\ a_{-\frac{1}{2}} \end{pmatrix} = \lambda \begin{pmatrix} a_{\frac{1}{2}} \\ a_{-\frac{1}{2}} \end{pmatrix}. \tag{70}$$

This matrix equation can be written as two simple linear equations:

$$-i\frac{\hbar}{2}a_{-\frac{1}{2}} = \lambda a_{\frac{1}{2}} \tag{71}$$

$$i\frac{\hbar}{2}a_{\frac{1}{2}} = \lambda a_{-\frac{1}{2}}. \tag{72}$$

Substituting Eq. 71 into Eq. 72 we find $i\frac{\hbar}{2}a_{\frac{1}{2}} = i\lambda^2\frac{2}{\hbar}a_{\frac{1}{2}}$. This requires that $\lambda = \pm\frac{\hbar}{2}$, exactly what we should have found. The y- and z-directions can not be fundamentally different! The same values for angular momentum in each direction must be permitted. We can then find normalized values for $a_{\frac{1}{2}}$ and $a_{-\frac{1}{2}}$ for each of these cases and write the resulting two states as:

$$|+1/2\rangle_y = \frac{1}{\sqrt{2}}|\frac{1}{2}\rangle_z + \frac{i}{\sqrt{2}}|-\frac{1}{2}\rangle_z \quad \lambda = +\frac{\hbar}{2} \tag{73}$$

$$|-1/2\rangle_y = \frac{1}{\sqrt{2}}|\frac{1}{2}\rangle_z - \frac{i}{\sqrt{2}}|-\frac{1}{2}\rangle_z \quad \lambda = -\frac{\hbar}{2}. \tag{74}$$

When writing down this solution, the linear homogenous equations, Eqs. 71 and 72, only determine the ratio between the coefficients of $|+\frac{1}{2}\rangle_z$ and $|-\frac{1}{2}\rangle_z$. An over-all multiplication factor which rescales the entire state is not determined. Requiring that the states in Eqs. 73 and 74 have unit length fixes the magnitude of such a possible multiplicative factor but the phase remains arbitrary. Thus, the choice of the coefficients of $|+1/2\rangle_z$ in Eqs. 73 and 74 to be real and positive is arbitrary and any other phase would have been just as good.

Since we can imagine performing a Stern-Gerlach experiment with an inhomogenous magnetic field oriented in the y-direction instead of the z-direction, the determination of the J_y eigenstates $|\pm\frac{1}{2}\rangle$ above, opens the way for many problems and thought experiments. If we use such an apparatus to

identify a state with $J_y = \frac{\hbar}{2}$, then just as is the case for J_z the probability that a subsequent measurement of J_y will yield $+\frac{\hbar}{2}$ is 100% while the probability of finding the value $-\frac{\hbar}{2}$ for J_y is 0%. If instead this state, which was determined to have $J_y = \hbar/2$, passes into a new Stern-Gerlach apparatus set up to measure J_z, the probability that the value $+\frac{\hbar}{2}$ will be found is 50%, simply the square of the $1/\sqrt{2}$ coefficient of the state $|\frac{1}{2}\rangle_z$ in Eq. 73. Likewise, the probability of finding the value $-\frac{\hbar}{2}$ is such a subsequent measurement of J_z is 50%.

We can also use the relation between the angular momentum operators and rotations to study very concretely how these spin-1/2 state behave when rotated. Specifically to rotate $|\psi\rangle$ through θ around an axis $\hat{n}$ following Eq. 65, we would perform the following operation:

$$|\psi\rangle \rightarrow e^{-i\theta\hat{n}\cdot\vec{J}/\hbar}|\psi\rangle. \tag{75}$$

or, in terms of column vectors and the matrices σ^i:

$$\begin{pmatrix} a_{\frac{1}{2}} \\ a_{-\frac{1}{2}} \end{pmatrix} \rightarrow e^{-i\theta\hat{n}\cdot\vec{\sigma}/2} \begin{pmatrix} a_{\frac{1}{2}} \\ a_{-\frac{1}{2}} \end{pmatrix}. \tag{76}$$

While these equation may again seem to be a useless abstraction, the Taylor series will come to our rescue and, for this special 2-dimensional case, allow us to explicitly determine the effects of a general rotation as given by Eq. 76. The important point that simplifies this particular exponential is that for any 3-vector $\vec{A}$,

$$\begin{aligned} (\vec{A}\cdot\vec{\sigma})^2 &= (A_1\sigma^1 + A_2\sigma^2 + A_3\sigma^3)^2 & (77) \\ &= (A_1\sigma^1)^2 + (A_2\sigma^2)^2 + (A_3\sigma^3)^2 & (78) \\ &= A_1^2 + A_2^2 + A_3^3. & (79) \end{aligned}$$

The original square of the sum of three terms in the equation above will contain nine different terms. However, it is easy to verify by explicit multiplication that $\sigma^i\sigma^j = -\sigma^j\sigma^i$ if $i \neq j$ so that the six cross terms between the different pairs of terms will cancel leaving only the three simple squares. These are also easy to determine since, again by direct multiplication, we can verify that $(\sigma^1)^2 = (\sigma^2)^2 = (\sigma^3)^2 = 1$.

Thus, if we Taylor expand the rotation exponential in Eq. 76, all of the even terms will contain $(\hat{n}\cdot\vec{\sigma})^2 = \hat{n}^2 = 1$ since $\hat{n}$ is a unit vector. Just as for a simple exponent of a complex number, the Taylor series then becomes easy.

17

The even terms (which are those in the expansion of a cosine), will contain no matrices while the odd terms (which are those in the expansion of a sine) will contain one power of the quantity $-i\hat{n} \cdot \vec{\sigma}$. Thus, we have shown that:

$$e^{-i\theta\hat{n}\cdot\vec{\sigma}/2} = \cos(\theta/2) - i\hat{n} \cdot \vec{\sigma}\sin(\theta/s). \tag{80}$$

This relation might join a collection of results that you have assembled which separate the even and odd terms in the expansion of the exponential:

$$e^x = \cosh(x) + \sinh(x) \tag{81}$$
$$e^{ix} = \cos(x) + i\sin(x). \tag{82}$$

3 More detail about operators and measurement

There are general properties of linear operators acting on a complex vector space that underlie some of the physical interpretations we have made above for both the rotation operators and the association of measurements with operators and their eigenstates and eigenvalues. It may be helpful to discuss these general properties here before we consider further applications of these ideas to translations in time and additional physical measurements.

3.1 Unitary and Hermitian operators

An important property that a symmetry transformation such as a rotation operator should obey is preserving the inner product between two states. For example, in the case of rotations if A and B are two states and $R(\hat{n}, \theta)$ a rotation operator we should require that

$$(R(\hat{n}, \theta)A, R(\hat{n}, \theta)B) = (A, B) \tag{83}$$

so that rotations do not affect the fundamental aspects of our complex vector space such as the inner product. In order to characterize this equation obeyed by the operator $R(\hat{n}, \theta)$, it is useful to introduce two properties that an operator O might obey. First we need to define the *Hermitian conjugate* $O^\dagger$ of the operator O. We define $O^\dagger$ by applying it to an arbitrary vector $|\psi\rangle$

18

and then specifying the inner product of the resulting vector with a second arbitrary vector $|\psi'\rangle$:

$$\big(|\psi'\rangle, O^\dagger|\psi\rangle\big) = \big(O|\psi'\rangle, |\psi\rangle\big). \tag{84}$$

If we introduce a basis $|1\rangle, |2\rangle, \ldots |N\rangle$ then an operator O is complete characterized by the $N \times N$ matrix of complex numbers given by

$$O_{ij} = \big(|i\rangle, O|j\rangle\big) \tag{85}$$

which implies that Eq. 84 is sufficient to define $O^\dagger_{ij}$ for all our basis states $|i\rangle$ and $|j\rangle$ and hence defines the operator $O^\dagger$.

Equations 84 and 85 are also interesting because they tell us how to relate the elements of the complex matrices O_{ij} and $O^\dagger_{ij}$:

$$O^\dagger_{ij} = O^*_{ji}. \tag{86}$$

Thus, we can obtain the Hermitian conjugate of a matrix by reversing rows and columns and taking the complex conjugate.

Using Eq. 84 we can rewrite Eq. 83 as

$$\big(|\psi'\rangle, |\psi\rangle\big) = \big(|\psi'\rangle, R(\hat{n}, \theta)^\dagger R(\hat{n}, \theta)|\psi\rangle\big). \tag{87}$$

Since this is true for all $|\psi\rangle$ and $|\psi'\rangle$ we can conclude that

$$R(\hat{n}, \theta)^\dagger R(\hat{n}, \theta) = I \quad \text{or} \quad R(\hat{n}, \theta)^\dagger = R(\hat{n}, \theta)^{-1} \tag{88}$$

A matrix whose Hermitian conjugate is also its inverse is defined to be *unitary*.

The unitary character of the rotation matrix $R(\hat{z}, \theta)$ requires that its generator J_z also obey a related property. This property of J_z can be easily deduced if we know how to take the Hermitian conjugate of a product of operators. Let's consider two operators O_1 and O_2 and try to work out the Hermitian conjugate of their product $(O_1 O_2)^\dagger$:

$$
\begin{aligned}
(A, (O_1 O_2)^\dagger B) &= (O_1 O_2 A, B) \tag{89}\\
&= (O_2 A, O_1^\dagger B) \tag{90}\\
&= (A, O_2^\dagger O_1^\dagger B), \tag{91}
\end{aligned}
$$

19

where in each of these equations we have simply used the definition of Hermitian conjugate to relate the inner product with an operator acting on the left argument to the inner product with the Hermitian conjugate of that operator acting on the right argument. Thus, we have shown that

$$(O_1 O_2)^\dagger = O_2^\dagger O_1^\dagger. \tag{92}$$

We can simply take the Hermitian conjugate of each of the operators in the product but must reverse their order. With this information we can compute

$$R(\hat{z}, \theta)^\dagger = \left(e^{-iJ_z\theta/\hbar} \right)^\dagger \tag{93}$$

$$= \left(\sum_{n=0}^\infty \frac{1}{n!} (-iJ_z\theta/\hbar)^n \right)^\dagger \tag{94}$$

$$= \left(\sum_{n=0}^\infty \frac{1}{n!} (iJ_z^\dagger\theta/\hbar)^n \right) \tag{95}$$

$$= e^{+iJ_z^\dagger\theta/\hbar} \tag{96}$$

where in Eq. 95 we have used the relation 92 multiple times but ignored questions of the ordering of the factors of J_Z since they are all the same operator. Finally we can use the fact that the inverse of the operator $R(\hat{z}, \theta)$ is simply $R(\hat{z}, -\theta)$ to write the requirement that $R(\hat{z}, \theta)$ be unitary as the equation:

$$e^{-iJ_z(-\theta)/\hbar} = e^{+iJ_z^\dagger\theta/\hbar} \tag{97}$$

which will be obeyed if

$$J_z = J_z^\dagger, \tag{98}$$

i.e. the operator J_z and its Hermitian conjugate $J_z^\dagger$ are the same thing. Such an operator is said to be *Hermtian*. If you look back at our specific choice for the three sigma matrices used to define each of the generators J_x, J_y and J_z given in Eq. 66. you will recognize that in fact each is Hermitian — each is unchanged if we both reflect the the matrix in the diagonal and take the complex conjugate of the matrix element. For example:

$$\sigma_y^\dagger = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}^\dagger = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{99}$$

Thus, it is essential that the three angular momentum operators be Hermitian.

3.2 Connection between linear operators and measurement

When discussing the measurement of the z-component of angular momentum we have argued that only a discrete set of values are allowed, $\hbar m$ with $-j \leq m \leq j$ and that to each value of the integer (or half-integer) m there corresponds a normalized state $|m\rangle$ which is the unique quantum state which has that value of J_z. The states $|m\rangle$ describe all possible results and hence should from a basis for vector space of allowed states. We then defined the operator J_z which satisfied $J_z|m\rangle = \hbar m|m\rangle$ which we described by saying that $\hbar m$ was an eigenvalue of the operator J_z and that state $|m\rangle$ the corresponding eigenstate. This example is an excellent representative of the relation between the quantum mechanical measurement of a physical quantity and a corresponding operator. We can now accurately summarize the way measurements are described in quantum mechanics. For each quantity that can be measured there corresponds a linear operator O that acts on the quantum states in the complex vector space of states making up the quantum system.

In our earlier discussion of angular momentum J_y, it was important that the solutions of Eqs. 71 and 72, $\lambda = \pm\hbar/2$, were real since these were supposed to be the possible results of measurements of J_y. Complex values for J_y would not make physical sense. This property can be assured in the general case by requiring that the operator O be Hermitian: $O = O^\dagger$.

For Hermitian operators on finite dimensional vector spaces it is possible to show three important properties:

1. A solution to the equation:

$$O|\psi\rangle = \lambda|\psi\rangle \tag{100}$$

   determines an eigenvector $|\psi\rangle$ of O and the complex number λ, the corresponding eigenvalue. If O is Hermitian then all such eigenvalues are real.

2. If O is Hermitian then eigenvectors of O with different eigenvalues are orthogonal.

3. If O is Hermitian then the set of all eigenvectors of O form a basis for the vector space on which O acts. This set of basis vectors can be chosen to be orthonormal.

Let's prove statements 1. and 2. in turn. We can show that λ must be real if O is Hermitian by performing the following steps:

$$
\begin{align}
\lambda &= (|\lambda\rangle, O|\lambda\rangle) \tag{101}\\
&= (O^\dagger|\lambda\rangle, |\lambda\rangle) \tag{102}\\
&= (O|\lambda\rangle, |\lambda\rangle) \tag{103}\\
&= (\lambda|\lambda\rangle, |\lambda\rangle) \tag{104}\\
&= \lambda^*(|\lambda\rangle, |\lambda\rangle) \tag{105}\\
&= \lambda^*, \tag{106}
\end{align}
$$

where we have used $|\lambda\rangle$ as the eigenstate of O with eigenvalue λ.

Similarly, if $|\lambda\rangle$ and $|\lambda'\rangle$ have different eigenvalues λ and λ' and O is Hermitian then

$$
\begin{align}
\lambda(|\lambda'\rangle, |\lambda\rangle) &= (|\lambda'\rangle, O|\lambda\rangle) \tag{107}\\
&= (O|\lambda'\rangle, |\lambda\rangle) \tag{108}\\
&= (\lambda'|\lambda'\rangle, |\lambda\rangle) \tag{109}\\
&= \lambda'(|\lambda'\rangle, |\lambda\rangle) \tag{110}
\end{align}
$$

or subtracting the left hand side of Eq. 107 from the right hand side of Eq. 110 we find:

$$
(\lambda' - \lambda)(|\lambda', |\lambda\rangle) = 0 \tag{111}
$$

so if $\lambda' - \lambda \neq 0$ then $|\lambda'$ and $|\lambda\rangle$ must be orthogonal. Property 3. will be shown to be true in a later homework problem.

We can use these properties to give a complete description of the quantum theory of measurement. If O is a Hermitian operator corresponding to a physical quantity O (also called O for economy of notation), then a general quantum state $|\psi\rangle$ can be written in terms of a basis of eigenstates of O:

$$
\psi = \sum_{n=1}^{N} a_n|\lambda_n\rangle, \tag{112}
$$

where the eigenstates $|\lambda_n\rangle$ obey

$$
O|\lambda_n\rangle = \lambda_n|\lambda_n\rangle. \tag{113}
$$

When the quantity O is measured the only allowed result is one of the eigenvalues λ_n of O. If the state $|\psi\rangle$ is normalized to one, $(|\psi\rangle, |\psi\rangle) = 1$, then

$|a_n|^2$ is the probability of obtaining the value λ_n when measuring O. (In the case that more than one of the eigenvalues λ_n have the the same value then this probability is the sum of $|a_n|^2$ over all states ψ_n with this common value of λ_n.) After such a measurement is performed and a value λ_n found for O, the quantum state $|\psi\rangle$ changes and is replaced by the eigenstate $|\lambda_n\rangle$ which corresponds to the eigenvalue λ_n that was found. This "collapse" or "reduction" of the wave function insures the consistent situation that subsequent measurements of O will give the same result as was initially determined.

3.3 Expectation values

The state vector $|\psi\rangle$ corresponding to a specific physical situation is a complicated quantity. If we consider a measurable physical quantity such as the angular momentum about a given direction or the angular or linear position discussed later, such a vector is determined by a giving complex probability amplitude for each of the possible values for that physical quantity. This can be many complex amplitudes! This degree of detail makes it hard to describe what is going on as the time evolves or when other measurements act to change the state vector $|\psi\rangle$.

A convenient quantity, associated again with a specific measurable quantity or observable O, which summarizes some aspects of the state vector and may be easier to keep track of and understand is the expectation value of that observable:

$$\langle O \rangle = \langle \psi | O | \psi \rangle = (|\psi\rangle, O|\psi\rangle). \tag{114}$$

The expectation value has a simple physical meaning. If the same state $|\psi\rangle$ is prepared many times and the quantity O is measured each time, then the average result of all of those measurements will be the expectation value $\langle O \rangle$.

This statement is easy to demonstrate. Begin by introducing an orthonormal basis of states $|\lambda_n\rangle$ where $|\lambda_n\rangle$ is an eigenvector of O with eigenvalue λ_n:

$$O|\lambda_n\rangle = \lambda_n |\lambda_n\rangle. \tag{115}$$

We can then expand the state of interest, $|\psi\rangle$ in this basis:

$$|\psi\rangle = \sum_n \psi_n |\lambda_n\rangle \tag{116}$$

where ψ_n is the complex probability amplitude associated with the result λ_n of measuring the quantity O if the system is in the state $|\psi\rangle$. Thus, if a single

measurement of the quantity O is performed, the probability of finding the value λ_n is $|\psi_n|^2$. If the state $|\psi\rangle$ is prepared many times, the average result of measuring O on each of those identically prepared states will be the sum of each possible result, λ_n, weighted by the probability $|\psi_n|^2$ of that result occurring:

$$\langle O \rangle = \sum_n \lambda_n |\psi_n|^2. \tag{117}$$

The equivalence of Eqs. 114 and 117 is easy to see if we substitute the eigenvector expansion given in Eq. 116 into Eq. 114:

$$
\begin{aligned}
\langle O \rangle &= (|\psi\rangle, O|\psi\rangle) & (118)\\
&= \left(\sum_{n'} \psi_{n'}|\lambda_{n'}\rangle, O \sum_n \psi_n|\lambda_n\rangle \right) & (119)\\
&= \left(\sum_{n'} \psi_n|\lambda_{n'}\rangle, \sum_n \psi_n \lambda_n|\lambda_n\rangle \right) & (120)\\
&= \sum_n \left(\psi_n|\lambda_n\rangle, \psi_n \lambda_n|\lambda_n\rangle \right) & (121)\\
&= \sum_n \psi_n^* \psi_n \lambda_n & (122)\\
&= \sum_n |\psi_n|^2 \lambda_n & (123)
\end{aligned}
$$

which is the equality we set out to prove.

Of course, the expectation value of an observable O for a quantum state $|\psi\rangle$ takes on added meaning in a classical limit where we consider states for which the fluctuations in the value of O are much smaller than the average value of O. Then to a good approximation the expectation value of O is the value that will be found every time that O is measured and $\langle O \rangle$ takes on the meaning of the classical value of that observable and the other details of the state $|\psi\rangle$ become unimportant. We will return to discuss expectation values further in Sec. ?? after we have learned something about observables whose measured values are continuous rather than discrete.

4 Time translations and the Schrödinger equation

The description of rotations as linear operations which literally rotate the states in our Hilbert space and preserve the inner product is a description which applies to nearly all symmetry operations in quantum mechanics. Surely the first requirement of such a symmetry must be that it should not change the simplest aspect of our complex vector space, the inner product. The steps of considering small transformations as in Sec. 2.1, equating the operation giving (*i.e.* generating) such a small transformation with a derivative and then integrating the resulting first order differential equation to determine the operator which performs large transformations are in fact very general. The most important examples are translations in space (discussed in Sec. **??**) and translations in time (discussed here).

In a physical situation where things are changing with time, it is natural to describe such time variation by assuming that the state vector in Hilbert space which describes our system is itself a function of the time $|\psi\rangle \rightarrow |\psi(t)\rangle$. As in the case of rotations we might hypothesize that the effects of a small increase in the time $t \rightarrow t + \Delta t$ will be proportional to Δt and can be described by the action of a linear operator on the state at time t:

$$|\psi(t + \Delta t)\rangle = |\psi(t)\rangle - i\Delta t H/\hbar|\psi(t)\rangle. \tag{124}$$

Just as in the case of rotations, the requirement that:

$$\langle\psi(t + \Delta t)|\psi(t + \Delta t)\rangle = \langle\psi(t)|\psi(t)\rangle \tag{125}$$

implies that the operator H, called the Hamiltonian, is Hermitian.

There are profound connections between the symmetry (rotations, space translations, time translations, *etc.*) and the quantity which becomes a constant of the motion if the dynamics possesses that symmetry (angular momentum, linear momentum, energy, *etc.*). For example, if a physical system is unchanged by rotation about the z-direction, then it cannot experience a torque in the z-direction and the angular momentum in the z-direction will be a constant. This connection can be seen in Hamilton's formulation of classical mechanics (which you will likely study in a more advanced mechanics course) and is fundamental to quantum mechanics. This reasoning, aided by a familiarity with Hamiltonian classical mechanics, implies that the operator

H in Eq. 124 should be identified with the energy, called the Hamiltonian, in Hamiltonian mechanics.

If Eq. 124 is rewritten by moving the $|\psi(t)\rangle$ term from the right- to the left-hand side and dividing by $-i\Delta t/\hbar$, we learn:

$$\frac{|\psi(t+\Delta t))\rangle - |\psi(t)\rangle}{-i\Delta t/\hbar} = H|\psi(t)\rangle \quad \text{or} \tag{126}$$

$$i\hbar\frac{d}{dt}H|\psi(t)\rangle = H|\psi(t)\rangle, \tag{127}$$

where the final equation is known as Schrödinger's equation.

Just as for the case of rotations, we can solve this equations using exponentials:

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle. \tag{128}$$

The fact that the exponential of Eq. 128 solves Eq. 127 is not difficult to see if the exponential of the operator $-iHt/\hbar$ is defined through its power series. When $t = 0$ only the '1' term in the power series survives and the needed initial condition:

$$|\psi(t)\rangle|_{t=0} = |\psi(0)\rangle, \tag{129}$$

is automatically obeyed. Finally the same properties of the exponential series which would guarantee that Eq. 128 solves Eq. 127 for the case when the factor $-iH/\hbar$ in exponent is a real number will work here. The single operator $-iH/\hbar$ and its powers obey the same rules of algebra as for the case of a real number.

Since we have only begun our discussion of quantum mechanics, not many examples of the Schrödinger equation are yet available to us. However, one simple and important example is available and is very instructive. Imagine that we have a fixed particle with angular momentum $\vec{J}$ placed in a magnetic field $\vec{B}$. If the particle contains an electric charge density and that charge density is proportion to the particle's mass density, then just as the spinning mass will produce the angular momentum $\vec{J}$, the spinning charge will produce a magnetic moment $\vec{\mu}$ proportional to $\vec{J}$:

$$\mu = \gamma\vec{J}. \tag{130}$$

For the case of the simple proportionality of the change and mass densities, $\gamma = Q/2Mc$, where Q is the total charge and M the total mass. The quantity γ is called the gyromagnetic ratio.

For this problem the energy of the spinning particle in the magnetic field is simply $-\vec{\mu} \cdot \vec{B} = -\vec{\gamma} \vec{J} \cdot \vec{B}$. So far this discussion is entirely classical. However, this expression for the energy involves quantities which have a fine meanings in quantum mechanics so it is natural to use

$$H = -\vec{\gamma} \vec{J} \cdot \vec{B} \tag{131}$$

in quantum mechanics where H is now the energy operator and $\vec{J} = (J_x, J_y, J_z)$ the three angular momentum operators, *i.e.* the three generators of rotations. Now the Schrödinger equation and its solution become quite explicit:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = -\vec{\gamma} \vec{J} \cdot \vec{B} |\psi(t)\rangle \tag{132}$$

$$|\psi(t)\rangle = e^{+i\gamma \vec{J} \cdot \vec{B} t/\hbar} |\psi(0)\rangle. \tag{133}$$

In fact comparing with Eq. 65, we recognize that this exponential describes a rotation of our quantum state $|\psi(0)\rangle$ around the direction of $-\vec{B}$ through the angle $\theta(t) = \gamma Bt$. This is precisely the Larmor precession that a classical particle with angular momentum $\vec{J}$ and magnetic moment $\vec{\mu}$ related by $\vec{\mu} = \gamma \vec{J}$ undergoes when placed in a magnetic field! Thus, for this case the solution of the Schrödinger equation is easy to recognize and corresponds beautifully to the classical solution.

5 Combining two quantum degrees for freedom

An important construction in both classical and quantum mechanics is the creation a composite system made up of two parts which are separately sensible physical systems. For example, it is very standard after studying the classical motion of a single particle in three dimensions, described by the single 3-dimensional position vector $\vec{r}(t)$, to go on to consider a system of two such particles. In quantum mechanics after having understood the behavior of the system of a single particle with spin-1/2, we would naturally next consider a system of two such particles. It is valuable to discuss how this is done from a mathematical perspective since the construction of combined states is very different in quantum and classical mechanics.

Thus, let's begin by describing two different ways to combine two vector spaces, say an M-dimensional vector space D and an N-dimensional vector space E into a single larger vector space.

5.1 Vector space Cartesian products

The most familiar method to combine these two vector spaces D and E is the *Cartesian product, $D \oplus E$*. If $\{|d_i\rangle\}_{1 \le i \le M}$ and $\{|e_j\rangle\}_{1 \le j \le N}$ are sets of basis vectors for our two spaces, then $v_D = \sum_{i=1}^{M} a_i |d_i\rangle$ and $v_E = \sum_{j=1}^{N} b_j |e_j\rangle$ are general vectors in D and E respectively. To construct the Cartesian product, $D \oplus E$ of the two vector spaces D and E, we simply define the combined set of $M + N$ vectors $\{|d_i\rangle\}_{1 \le i \le M}$ and $\{|e_j\rangle\}_{1 \le j \le N}$ as the basis for the new space and declare that the combined sum $\sum_{j=1}^{M} a_j |d_j\rangle + \sum_{j=1}^{N} b_j |e_j\rangle$ describes a general vector in the new Cartesian product.

Thus, with a Cartesian product, we can take a vector v_D from D and v_E from E and simply add them to get $v_D + v_E$, a vector in $D \oplus E$. In fact, all vectors in $D \oplus E$ can be written as such a sum of a vector from D and a vector from E.

The dimension of this new space is the sum of the dimensions of its two components, $M + N$, and it contains non-zero vectors made up of a non-zero vector in one space and the zero vector in another. If we are studying the physics of two classical particles, it works very well to use this six dimensional vector space containing a general vector $\sum_{i=1}^{3} a_i \hat{d}_i + \sum_{j=1}^{3} b_j \hat{e}_j$ with the numbers (a_1, a_2, a_3) locating the position of the particle D and (b_1, b_2, b_3) locating particle E. This physical situation is nicely represented by writing these six numbers together as a six-component vector, $(a_1, a_2, a_3, b_1, b_2, b_3)$. The examples $(a_1, a_2, a_3, 0, 0, 0)$ and $(0, 0, 0, b_1, b_2, b_3)$ make good sense, describing the case were the particle E or the particle D happens to be located at the origin.

This last property (the non-zero character of the combination of a non-zero vector from one system and the zero vector from the other) makes this construction a poor choice for a two-particle quantum system. If we are locating a classical particle the zero vector does not have any special physical meaning and we certainly need to be able to describe the situation represented by the vector $(a_1, a_2, a_3, 0, 0, 0)$. However, this is is very different in quantum mechanics. In that case, the zero vector is special since it cannot be normalized to one and hence does not represent a valid quantum state. Combining a zero vector from one vector space with a non-zero vector from the other should give the zero vector for the combined system. Thus, to combine two quantum systems we need to define a different product.

5.2 Vector space tensor products

Given the two sets of basis vectors $\{|d_i\rangle\}_{1\leq i\leq M}$ and $\{|e_j\rangle\}_{1\leq j\leq N}$ for the vector spaces D and E, we define a set of basis vectors for the *tensor product* of these two spaces, $D\otimes E$ as the set of $M\cdot N$ vectors $\{|d_i\rangle\otimes|e_j\rangle\}_{1\leq i\leq M;1\leq j\leq N}$. Thus, the tensor product vector space, $D\otimes E$, has $M\cdot N$ basis vectors and hence dimension $M\cdot N$.

Given a vector in v_D in D and v_E in E, their product, $v_D\otimes v_E$ will be a vector in the tensor product vector space $D\otimes E$. In terms of components, this product can be defined by expanding out the product in the expected way:

$$v_D\otimes v_E = \sum_{i=1}^{M}\sum_{j=1}^{N} a_i b_j |d_i\rangle\otimes|e_j\rangle. \tag{134}$$

However, not all states in $D\otimes E$ can be written as a product of a state in D and a state in E. For example, the state

$$\psi = \hat{d}_1\otimes\hat{e}_1 + \hat{d}_2\otimes\hat{e}_2 \tag{135}$$

can't be written as such a product.

Now the product of the zero state in either vector space with an arbitrary vector in the other is the zero vector in the tensor product space. The resulting product space suits the needs of quantum mechanics very well.

5.2.1 Quantum interpretation of tensor product

It is very natural to interprete the product state $|d_i\rangle\otimes|e_j\rangle$ as that state in which particle D is in the state $|d_i\rangle$ and particle E is in the state $|e_j\rangle$. Similarly a general state

$$\psi = \sum_{i=1}^{M}\sum_{j=1}^{N} c_{i,j}\,|d_i\rangle\otimes|e_j\rangle \tag{136}$$

is a superposition of such simple product states and the coefficient $c_{i,j}$ should be interpreted as the probability amplitude for finding particle D in the state $|d_i\rangle$ and particle E in the state $|e_j\rangle$, with $|c_{i,j}|^2$ being the probability for that being true.

The interesting state ψ of Eq. 135 is a new and very important possibility in two-particle quantum theory. It describes a situation in which the wave functions of particles D and E are said to be *entangled* with information

about one of the particles implying knowledge about the other as well, even if no measurements are performed on the second particle. For example, if it is known that the system is in the entangled state ψ of Eq. 135 and a property of particle D is measured which implies that this particle must be in the state $|d_2\rangle$, then we know that particle E must be in the state $|e_2\rangle$ and there is no probability that it will be found in the state $|e_1\rangle$. This leads to all sorts of non-intuitive situations since these two particles could be very far separated and yet a measurement of the properties of one of the particles will instantaneously lead to predictive knowledge about the other!

6 Position and momentum

The example of spin-1/2 particles worked out in some detail in the previous section illuminates many very important parts of quantum theory. However, we appear far from discussing an actual particle moving in space with possible values for its position and momentum.

In fact, we are already surprisingly close to such a discussion. For the case of $j = 1/2$ or $j = 1$ it is very hard to see "what" is rotating to provide the angular momentum that these systems carry. However, this is because of what we will understand to be the uncertainty principle: "The angular momentum and angular position of a particle cannot be simultaneously known, each with arbitrary precision".

If we focus only on the z-component of angular momentum and attempt to interpret J_z as coming from something that is rotating, the fact that $J_z = \pm\hbar/2$ implies that the angular momentum is very well known (it's very small) so little can be said about the angular position of the quantity that is rotating.

6.1 Defining continuous angular position

This suggests that we should consider examples with larger values of j permitting greater uncertainty in J_z with $-j\hbar \leq J_z \leq +j\hbar$. Following that direction, let's consider a quantum space of $2j + 1$ states $|j, m\rangle_{-j \leq m \leq j}$ which are eigenstates of $\vec{J}^2$ and J_z as discussed above:

$$\vec{J}^2|j,m\rangle = \hbar^2 j(j+1)|j,m\rangle \quad \text{and} \quad J_z|j,m\rangle = m\hbar|j,m\rangle, \tag{137}$$

and consider the possibility that j is quite large and even consider the limit $j \to \infty$. (The relation between j and the eigenvalue of the operator $\vec{J}^2 = j_x^2 + J_y^2 + J_z^2$ is discussed in Appendix A and is not central to the present discussion and so can be ignored if you wish.)

Recall that under rotations through an angle ϕ around the z-axis, these states transform as:

$$|j, m\rangle \to e^{-iJ_z\phi/\hbar}|j, m\rangle = e^{-im\phi}|j, m\rangle. \tag{138}$$

Thus the states $|j, m\rangle$ hardly change at all when rotated since this simple change of phase has no physical effect. These states have, in some sense, no orientation!

Next let's examine some very different states that can be constructed from these:

$$|\theta_n\rangle = N_j \sum_{m=-j}^{+j} e^{-i\theta_n m}|j, m\rangle \quad \text{where} \quad \theta_n = \frac{2\pi}{2j+1}n \tag{139}$$

and $n = -j, -j+1, \ldots, j$. The constant N_j is a normalization factor that we will choose shortly. These states are very interesting because under rotation the change in phase of $|j, m\rangle$ acts to shift, or rotate, the label θ_n suggesting that these are states with an angular position θ_n:

$$|\theta_n\rangle \to e^{-iJ_z\phi/\hbar}|\theta_n\rangle = N_j \sum_{m=-j}^{+j} e^{-i(\theta_n+\phi)m}|j, m\rangle. \tag{140}$$

For the special case that $\phi = \phi_k = 2\pi k/(2j+1)$, this rotation actually shifts the state $|\theta_n\rangle$ to the state $|\theta_{n+k}\rangle$ just as if the angle θ_n were being rotated into the angle $\theta_{n+k} = \theta_n + \phi_k$.

We are trying to create states with a definite angular position out of our J_z eigenstates. The same quantum vector space will describe both angular position and angular momentum!

To get a better idea of what is going on, we need to understand our states $|\theta_n\rangle$ better. These are actually $2j + 1$ states which are orthogonal and form just as good a basis as the original eigenvectors $|j, m\rangle$. To see this (and determine the constant N_f) we need to compute the inner product between two of these new states:

$$(|\theta_{n'}\rangle, |\theta_n\rangle) = N_j^2 \left(\sum_{m'=-j}^{+j} e^{-i\theta_{n'}m'}|j, m'\rangle, \sum_{m=-j}^{+j} e^{-i\theta_n m}|j, m\rangle \right) \tag{141}$$

31

$$= N_j^2 \sum_{m=-j}^{+j} e^{i(\theta_{n'}-\theta_n)m} \tag{142}$$

$$= N_j^2 e^{-i(\theta_{n'}-\theta_n)j} \left(\sum_{m'=0}^{+2j} e^{i(\theta_{n'}-\theta_n)m} \right) \tag{143}$$

$$= N_j^2 e^{-i(\theta_{n'}-\theta_n)j} \left(\frac{1 - e^{i(\theta_{n'}-\theta_n)(2j+1)}}{1 - e^{i(\theta_{n'}-\theta_n)}} \right) \tag{144}$$

$$= N_j^2 \frac{e^{-i(\theta_{n'}-\theta_n)j} - e^{i(\theta_{n'}-\theta_n)(j+1)}}{1 - e^{i(\theta_{n'}-\theta_n)}} \tag{145}$$

$$= N_j^2 \frac{e^{i(\theta_{n'}-\theta_n)/2}}{e^{i(\theta_{n'}-\theta_n)/2}} \frac{e^{-i(\theta_{n'}-\theta_n)(j+1/2)} - e^{i(\theta_{n'}-\theta_n)(j+1/2)}}{e^{-i(\theta_{n'}-\theta_n)/2} - e^{i(\theta_{n'}-\theta_n)/2}} \tag{146}$$

$$= N_j^2 \frac{\sin\left((\theta_{n'} - \theta_n)(j + 1/2)\right)}{\sin\left((\theta_{n'} - \theta_n)/2\right)}. \tag{147}$$

Let's now go carefully through the steps in the above calculation. The first, Eq. 141, simply writes the inner product of the two states $|\theta_{n'}\rangle$ and $|\theta_n\rangle$ by expressing each as the appropriate sum over our orthonormal basis of states $|j, m\rangle$. The next line, Eq. 142, recognizes that only inner products between states on the left and right with identical J_z quantum numbers, $m' = m$, will survive (all other pairs of states are orthogonal) and writes the inner product as the sum over products of the complex conjugate of coefficients in the left factor $((e^{-i\theta_{n'}m})^*)$ with those on the right $(e^{-i\theta_n m})$.

In Eq. 143, we extract a factor of $e^{-i(\theta_{n'}-\theta_n)j}$ from the sum, leaving a sum which is a simple, finite geometrical series of the standard form:

$$\sum_{m=0}^{N} z^m = \frac{1 - z^{N+1}}{1 - z} \tag{148}$$

where $N = 2j$ and $z = e^{i(\theta_{n'}-\theta_n)}$. In the next equation, Eq. 144, we use this standard formula, Eq. 148, to evaluate the series in Eq. 143. In Eq. 145 the extra factor $e^{-i(\theta_{n'}-\theta_n)j}$ is multiplied back into the resulting sum, while in Eq. 146 a common factor of $e^{i(\theta_{n'}-\theta_n)/2}$ is removed from both the numerator and denominator. After canceling these identical factors and using the formula for the sine function we obtain the neat result given in Eq. 147.

We can use this result for two purposes. First because of our choice $\theta_n = 2\pi n/(2j + 1)$, the argument of the sine function in the numerator in Eq. 147 is $\pi(n' - n)$ making the numerator vanish. If $n' \neq n$, the denominator is non-zero which proves that the inner product $(|\theta_{n'}\rangle, |\theta_n\rangle) = 0$ for $n' \neq n$. Thus

our $2j + 1$ states $|\theta_n\rangle$ are orthogonal. The case $n' = n$ is difficult to evaluate using Eq. 147 since in that case both the numerator and the denominator vanish. However, in that case we can look at Eq. 142 and recognize that each term in the sum is $e^{i0} = 1$ so the sum gives $2j + 1$. We can thus make our states $|\theta_n\rangle$ orthonormal if we choose the factor $N_j = 1/\sqrt{2j + 1}$.

Thus, we have two equally good sets of basis vectors: the $|j, m\rangle$ with definite values of $J_z = m\hbar$ and the $|\theta_n\rangle$ with definite values of "position" $\theta = 2\pi n/(2j + 1)$. These are both bases with $2j + 1$ orthonormal elements.

We could define a Hermitian operator θ_{op} by specifying the basis $|\theta_n\rangle$ as its eigenvectors and the angles θ_n as its eigenvalues. A general state $|\psi\rangle$ could then be expanded in this basis:

$$|\psi\rangle = \sum_{n=0}^{2j} \psi_{\theta_n}|\theta_n\rangle \tag{149}$$

and we would interprete $|\psi_{\theta_n}|^2$ as the probability of finding the particle at the angular position θ_n. However, this is still a poor description of angular position which should be continuous and not have only $2j+1$ discrete values. **It would be contrary to present knowledge to think allowed angular positions were quantized.**

Never-the-less, we can see we are on the right track if we examine the effect of a rotation through a general angle. We can ask "If we start with a state $|\theta_n\rangle$ and rotate through ϕ what is the resulting distribution of positions?" This is easy to do. We first rotate the state $|\theta_n\rangle$:

$$|\theta_n\rangle \to e^{-iJ_z\phi/\hbar}|\theta_n\rangle = N_j \sum_{m=-j}^{+j} e^{-i(\theta_n + \phi)m}|j, m\rangle. \tag{150}$$

We can then compute the probability amplitude for this state having "position" $\theta_{n'}$ if we determine the component of this rotated state in the direction of the basis state $|\theta_{n'}\rangle$:

$$\psi_{\theta_{n'}} = (|\theta_{n'}\rangle, e^{-iJ_z\phi/\hbar}|\theta_n\rangle). \tag{151}$$

This exactly the inner product evaluated in Eqs. 141-147 except that the angle θ_n in those earlier equations is replaced by the sum $\theta_n + \phi$. Thus we get the result immediately from Eq. 147 by replacing $\theta_n \to \theta_n + \phi$:

$$\psi_{\theta_{n'}} = \frac{1}{2j + 1} \frac{\sin\left((\theta_{n'} - \theta_n - \phi)(j + 1/2)\right)}{\sin\left((\theta_{n'} - \theta_n - \phi)/2)\right)}. \tag{152}$$

As j becomes large this is a sharply peaked function of $\theta_{n'}$. It is largest for a very few terms that are of order one with $\theta_{n'} \approx \theta_n + \phi$ and the rest are of size $1/(2j+1)$ and hence very small as j grows.

Thus, our states $|\theta_n\rangle$ don't represent a really definite position but rather a smear of positions around θ_n with angles θ of the sort $\theta_{n-1} < \theta < \theta_{n+1}$. This is the uncertainty principle. Since J_z is known to lie between $-j\hbar$ and $+j\hbar$, the angular position must have this level of uncertainty. The largest possible uncertainty in J_z is $\Delta J_z \approx (2j+1)\hbar$ while the minimum uncertainty in angular position is $\Delta\theta \approx 2\pi/(2j+1)$ consistent with $\Delta J_z \Delta\theta \geq \hbar$, the uncertainty relation between angular momentum and angular position that we will discuss later.

Thus, to get the physical results we want, we must take the limit $j \to \infty$. As we will see our angular position-eigenstates $|\theta_n\rangle$ take poorly defined values in this limit so it is conventional to work with their coefficients ψ_{θ_n} which become the wavefunction of conventional quantum mechanics. Let's see how this works.

First we will write a general quantum state $|\psi\rangle$ in terms of our two bases:

$$|\psi\rangle \;=\; \sum_{m=-j}^{j} \widetilde{\psi}_m |j,m\rangle \tag{153}$$

$$|\psi\rangle \;=\; \sum_{n=-j}^{j} \psi_{\theta_n} |\theta_n\rangle \tag{154}$$

Next, we can relate the coefficients $\widetilde{\psi}_m$ and ψ_{θ_n} by substituting Eq. 139 into Eq. 154. We find

$$|\psi\rangle \;=\; \sum_{n=-j}^{j} \psi_{\theta_n} \Big(N_j \sum_{m=-j}^{+j} e^{-i\theta_n m} |j,m\rangle \Big) \tag{155}$$

$$=\; \sum_{m=-j}^{+j} \Big(N_j \sum_{n=-j}^{j} \psi_{\theta_n} e^{-i\theta_n m} \Big) |j,m\rangle. \tag{156}$$

$$\tag{157}$$

In passing from Eq. 155 to Eq. 156 we have simply rearranged the $(2j+1)^2$ term that are being added, performing the sum over m last. We can then compare Eq. 156 and Eq. 153 to pick out an equation for the coefficients $\widetilde{\psi}_m$

in terms of the ψ_{θ_n}:

$$\widetilde{\psi}_m = \frac{1}{\sqrt{2j+1}} \sum_{n=-j}^{j} \psi_{\theta_n} e^{-i\theta_n m}. \tag{158}$$

Using similar manipulations we can derive an equation going the other direction:

$$\psi_{\theta_n} = \frac{1}{\sqrt{2j+1}} \sum_{m=-j}^{+j} \widetilde{\psi}_m e^{+i\theta_n m}. \tag{159}$$

It is these two equations that can be easily modified so they have a good limit as $j \to \infty$. The idea is to rescale the amplitude ψ_{θ_n} so that Eq. 158 can be written as an integral over a continuous variable θ. This can be done by defining the "wavefunction" $\psi(\theta)$ as:

$$\psi(\theta_n) = \sqrt{\frac{2j+1}{2\pi}} \psi_{\theta_n}. \tag{160}$$

Equation 158 can then be written in the limit $j \to \infty$ as:

$$\widetilde{\psi}_m = \lim_{j \to \infty} \frac{1}{\sqrt{2\pi}} \sum_{n=j}^{j} \frac{2\pi}{2j+1} \psi(\theta_n) e^{-i\theta_n m} \tag{161}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} d\theta \psi(\theta) e^{-i\theta m}. \tag{162}$$

In taking the limit to go from Eq. 161 to Eq. 162 we have recognized $\Delta\theta = 2\pi/(2j+1)$ as the interval between θ_{n+1} and θ_n so this is the standard definition of an integral over θ. The second equation, Eq. 159 has an easier limit. This remains a sum, now with an infinite number of terms and the factor $\sqrt{2j+1}$ need only be moved to the other side to translate ψ_{θ_n} into $\psi(\theta_n)$:

$$\psi(\theta) = \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{+\infty} \widetilde{\psi}_m e^{+i\theta_n m}. \tag{163}$$

Since the probability of finding a particle at the position θ_n is $|\psi_{\theta_n}|^2 = |\psi(\theta_n)|^2 2\pi/(2j+1) = |\psi(\theta_n)|^2 \Delta\theta$, it is natural to interprete the modulus square of our new wave function $|\psi(\theta)|^2$ as the probability density per unit angle of finding our particle at the angular position θ.

The inner product between two states $|\psi\rangle$ and $|\psi'\rangle$ can easily be written in terms of their coefficients and does have a fine limit at $j \to \infty$. The easiest

relation is found for the coefficients $\widetilde{\psi}_m$ which are not changed when we try to take $j \to \infty$:

$$(|\psi'\rangle, |\psi\rangle) = \sum_{m=-\infty}^{+\infty} \widetilde{\psi'}^*_m \widetilde{\psi}_m. \tag{164}$$

For the case of the coefficients ψ_{θ_n} we must do a little more work but things still work out well:

$$(|\psi'\rangle, |\psi\rangle) = \lim_{j \to \infty} \sum_{n=-j}^{j} (\psi'_{\theta_n})^* \psi_{\theta_n} \tag{165}$$

$$= \lim_{j \to \infty} \sum_{n=0}^{2j} \psi'(\theta_n)^* \psi(\theta_n) \frac{2\pi}{2j+1} \tag{166}$$

$$= \int_{-\pi}^{\pi} d\theta \psi'(\theta)^* \psi(\theta). \tag{167}$$

Since it is much more convenient to work with the vector space formed from the coefficients $\psi(\theta)_{0 \le \theta \le 2\pi}$ or $(\psi_m)_{-\infty < m < \infty}$, we need to find out how the operators θ_{op} and J_z change these coefficients. For $\psi(\theta)$ the operator θ_{op} is easy. Since our original states $|\theta_n\rangle$ were each multiplied by θ_n this is the same a multiplying their coefficients by θ_n. Hence:

$$(\theta_{op}\psi)(\theta) = \theta\psi(\theta). \tag{168}$$

We simply multiply the function $\psi(\theta)$ by its argument.

The operator J_z is a little more complicated. Since it acts easily on the states $|j, m\rangle$, multiplying them by $\hbar m$, we can deduce:

$$(J_z\widetilde{\psi})_m = \hbar m \widetilde{\psi}_m. \tag{169}$$

We can then use Eq. 163 to determine $J_z\psi(\theta)$:

$$(J_z\psi)(\theta) = \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{+\infty} \hbar m \widetilde{\psi}_m e^{+i\theta m} \tag{170}$$

$$= -i\hbar \frac{d}{d\theta} \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{+\infty} \widetilde{\psi}_m e^{+i\theta m} \tag{171}$$

$$= -i\hbar \frac{d\psi(\theta)}{d\theta}, \tag{172}$$

a very simple equation for the effects of J_z:

$$(J_z\psi)(\theta) = -i\hbar \frac{d\psi(\theta)}{d\theta}. \tag{173}$$

36

6.2 Linear motion

We can now easily formulate the quantum theory of a single particle moving in a finite region $-L/2 \leq x \leq L/2$. We need only think of the angle θ introduced above as describing a location on a circle of radius R and introduce a position x as the distance moved along the circumference of that circle from the location as $\theta = 0$. Thus, we introduce the position $x = R\theta$. The range $-\pi \leq \theta < \pi$ then translates directly into the region $-L/2 \leq x \leq +L/2$ provided we choose $R = L/2\pi$.

Thinking of the a particle moving on a circle of radius R, it is natural to define a probability amplitude for linear motion:

$$\psi_l(x) = \psi_a(\theta = \frac{x}{R})/\sqrt{R}. \tag{174}$$

Here the extra subscripts a and l represent "linear" and "angular" and distinguish the two closely related functions $\psi_l(x)$ and $\psi_a(\theta)$. The function $\psi_a(\theta)$ is identical to the $\psi(\theta)$ discussed above. The factor $1/\sqrt{R} = \sqrt{2\pi/L}$ has been introduced to change the probability per unit angle, $|\psi_a(\theta)|^2$, to a probability per unit distance, $|\psi_l(x)|^2$. We can see by a simple change of variables that if the original wave function was normalized the new one will be too:

$$\int_{-L/2}^{+L/2} |\psi_l(x)|^2 dx = \int_{-\pi}^{\pi} |\psi_a(\theta)|^2 d\theta = 1. \tag{175}$$

We can also translate our angular momentum operator $J_z = -i\hbar\partial/\partial\theta$ into the linear momentum appropriate for the present case. We can use the classical relation $J_z = Rp$ to write:

$$p = J_z/R = -i\hbar\frac{\partial}{\partial\theta}\frac{1}{R} = -i\hbar\frac{\partial}{\partial R\theta} = -i\hbar\frac{\partial}{\partial x}. \tag{176}$$

This new, linear view-point still carries some remembrance of its angular origin since the points $x = \pm L/2$ correspond to $\theta = -\pi$ and π which are physically equivalent. In particular the wave function $\psi_a(\theta)$ should vary smoothly as θ crosses between 2π and 0. Thus, we should expect that $\psi_l(-L/2) = \psi_l(+L/2)$ and $\partial\psi_l(-L/2)/\partial x = \partial\psi_l(+L/2)/\partial x$.

Since the operators J_z and p are proportional, their eigenfunctions will be the same and their eigenvalues will differ only by the factor of $1/R$. Thus, the J_z eigenvalue of $m\hbar$ will correspond to an eigenvalue $m\hbar/R = m2\pi\hbar/L$

of p. Using the linear variable x and these new eigenvalues, $p_m = m2\pi\hbar/L$, we can rewrite the Eqs. 162 and 163:

$$\widetilde{\psi}_{p_m} = \frac{1}{\sqrt{L}} \int_{-L/2}^{+L/2} dx \; \psi(x) e^{-ip_m x/\hbar} \tag{177}$$

$$\psi(x) = \frac{1}{\sqrt{L}} \sum_{m=-\infty}^{+\infty} \widetilde{\psi}_{p_m} e^{+ip_m x/\hbar}. \tag{178}$$

Here the definite momentum amplitude $\widetilde{\psi}_{p_m}$ is the same as the original, definite J_z amplitude $\widetilde{\psi}_m$ and, for simplicity, we have dropped the subscript l. Equation 178 is said to express $\psi(x)$ as a Fourier series while Eq. 177 determines the coefficients $\widetilde{\psi}_{p_m}$ which appear in that series.

The final step that should be taken so that we have standard, single-particle quantum mechanics is to take the limit $L \to \infty$. Then our particle will be able to move over all of space between $x = -\infty$ and $x = +\infty$. As in the case of the limit $j \to \infty$, we must adjust the quantities appearing in Eqs. 177 and 178 so that this limit is well defined. As $L \to \infty$ the momentum values $p_m = m2\pi\hbar/L$ become more densely spaced, with vanishing separation $p_{m+1} - p_m = 2\pi\hbar/L$. Thus, we expect the individual amplitudes $\widetilde{\psi}_{p_m}$ should vanish but the sum of those amplitudes for a range of momenta between p and $p + \Delta p$ to remain finite. Thus, as before, we define:

$$\widetilde{\psi}(p) = \widetilde{\psi}_{p_m} \sqrt{\frac{L}{2\pi\hbar}}. \tag{179}$$

This choice implies that:

$$\sum_{m=pL/2\pi\hbar}^{(p+\Delta p)L/2\pi\hbar} |\widetilde{\psi}_{p_m}|^2 \approx \int_{p}^{p+\Delta p} |\widetilde{\psi}(p)|^2 dp. \tag{180}$$

In terms of the new quantity $\widetilde{\psi}(p)$, Eqs. 177 and 178 take a form in which the limit $L \to \infty$ can be taken:

$$\widetilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dx \; \psi(x) e^{-ipx/\hbar} \tag{181}$$

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dp \widetilde{\psi}(p) e^{+ipx/\hbar}. \tag{182}$$

These relations between the functions $\widetilde{\psi}(p)$ and $\psi(x)$ are conventionally described by saying that $\widetilde{\psi}(p)$ and $\psi(x)$ are Fourier transforms of each other.

A General form for the angular momentum operators

To begin this discussion we will take a more general point of view. Imagine that we have a complex vector space $\mathcal{H}$ with a large dimension (possibly infinite). We can then try to isolate a small, finite-dimensional subspace of vectors which transform into each other when rotations are performed. If we can find such a set then, at least when making rotations, we can examine this smaller subset of vectors in isolation and ignore the larger, often more complicated space in which they were found.

As in the case of spin-1/2, we will study the three operators J_x, J_y and J_z, (or, equivalently J_1, J_2 and J_3) which generate rotations. These were discussed earlier and, up to a factor of $\hbar/2$ took the form of the Pauli matrices for the case of a two-dimensional vector space of states. Now we would like to study the general case and instead of guessing possible forms for these matrices we will develop an approach which actually works them out exhaustively so that we discover all possible cases.

The first step in doing this is to consider a natural operator which can be formed from J_x, J_y and J_z, the operators

$$J^2 = J_x^2 + J_y^2 + J_z^2. \tag{183}$$

Since this is the length of a "vector" we should expect that its value will not be changed by rotations. In quantum mechanics, rotations are constructed from the individual J_i's so saying that J^2 is not changed by rotations is equivalent to the statement that first applying the operator J_x, for example, and then applying J^2 will give exactly the same result instead J^2 is applied first and then J_x. Or in equations:

$$J^2 J_x |\psi\rangle = J_x J^2 |\psi\rangle \tag{184}$$

for all states $|\psi\rangle$. Or $J^2 J_x - J_x J^2 = [J_x J^2] = 0$. This can be verified directly as is requested in problem #169.

Thus, if we find a state $|j\rangle$ which is an eigenstate of J^2:

$$J^2 |j\rangle = j(j+1)\hbar^2 |j\rangle \tag{185}$$

applying one of the operators J_i will produce a new state, $J_i|j\rangle$ which is also an eigenstate of J^2 with the same eigenvalue. This is easy to demonstrate

by exchanging the order of J^2 and J_i:

$$J^2\Big(J_i|j\rangle\Big) = J_i\Big(J^2|j\rangle\Big) = J_i\Big(j(j+1)\hbar^2|j\rangle\Big) = j(j+1)\hbar^2\Big(J_i|j\rangle\Big). \qquad (186)$$

Here we are writing the eigenvalue of the operators J^2 in the conventional form $j(j+1)\hbar^2$ which looks somewhat odd at present but will turn out to be a convenient form in few paragraphs below. At least it has the right dimension and is positive if we assume that j is a pure, non-negative number.

We will choose to work only with states in our larger vector space $\mathcal{H}$ which have a specific value of the quantum number j. The operator J_z acts within this set of states and, from the discussion above, we know that it cannot change the value of j. Within this subspace of states with the same value j we can find a basis of states which have definite values for the operator J_z. (The fact that such a basis can be found will be discussed later.) Here we will pick one such basis state $|j, m\rangle$ with the J^2 eigenvalue $j(j + 1)\hbar^2$ and the eigenvalue $m\hbar$ of J_z:

$$J^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle \quad J_z2|j, m\rangle = m\hbar|j, m\rangle \qquad (187)$$

The next step contains the essential idea needed to classify all of the different types of rotation matrices which exist. By using the commutation relations among the J_i we can easily show that when acting on an eigenstate of J_z with eigenvalue $m\hbar$ the operator $J_x + iJ_y$ creates a different eigenstate of J_z with eigenvalue increased by $\hbar$ while $J_x - iJ_y$ decreases the eigenvalue by $\hbar$:

$$J_z(J_x + iJ_y)|j, m\rangle = (m + 1)\hbar(J_x + iJ_y)|j, m\rangle \qquad (188)$$
$$J_z(J_x - iJ_y)|j, m\rangle = (m - 1)\hbar(J_x - iJ_y)|j, m\rangle. \qquad (189)$$

This is easy to demonstrate by simply reordering the operators. Consider the $J_x + iJ_y$ case:

$$J_z(J_x + iJ_y)|j, m\rangle = \Big([J_z, J_x + iJ_y]\Big)|j, m\rangle + (J_x + iJ_y)J_z|j, m\rangle \quad (190)$$
$$= \Big(iJ_y + J_x\Big)|j, m\rangle + (J_x + iJ_y)m\hbar|j, m\rangle \qquad (191)$$
$$= (m + 1)\hbar(J_x + iJ_y)|j, m\rangle. \qquad (192)$$

For this reason the operators $J_x \pm iJ_y$ are called ladder operators moving us up and down a ladder of eigenstates of J_z with eigenvalues increased or

decreased by $\hbar$. Of course, as discussed above, neither J_x nor J_y can change the eigenvalue of J^2 so these states are properly labeled with the fixed value of j.

Even more remarkable is that we can also determine the norm of the state $(J_x + iJ_y)|j, m\rangle$:

$$
\begin{align}
||(J_x + iJ_y)|j,m\rangle||^2 &= ((J_x + iJ_y)|j,m\rangle, (J_x + iJ_y)|j,m\rangle) \tag{193} \\
&= (|j,m\rangle, (J_x - iJ_y)(J_x + iJ_y)|j,m\rangle) \tag{194} \\
&= (|j,m\rangle, \left(J_x^2 + J_y^2 + i[J_x, iJ_y]\right)|j,m\rangle) \tag{195} \\
&= (|j,m\rangle, \left(J^2 - J_z^2 - \hbar J_z\right)|j,m\rangle) \tag{196} \\
&= \left(j(j+1) - m^2 - m\right)\hbar^2 \tag{197} \\
&= (j + 1 + m)(j - m)\hbar^2, \tag{198}
\end{align}
$$

where in going from Eq. 193 and to Eq. 194 we use the fact that the operators J_x and J_y are Hermitian while the i must change sign when moved from the left to the right argument of the inner product. Thus we can construct a chain of states $|j, m'\rangle$, $m' = m + 1$, $m + 2$, ... from our original state $|j, m\rangle$ joined by the raising operator $J_x + iJ_y$.

Likewise we can apply the lowering operator $J_x - iJ_y$ and show that this will decrease the value of J_z from $\hbar m$ to $\hbar(m - 1)$. We can also repeat steps very similar to those in Eq. 193 and following to find the norm of the resulting lowered state:

$$||(J_x - iJ_y)|j,m\rangle||^2 = (j + 1 - m)(j + m)\hbar^2. \tag{199}$$

So far it appears that we can create states with arbitrarily positive and arbitrarily negative values of J_z. This makes no physical sense (and cannot occur mathematically either) because for any state $|\psi\rangle$

$$\langle\psi|J_z^2|\psi\rangle \leq \langle\psi|J^2|\psi\rangle = \hbar^2 j(j+1). \tag{200}$$

Thus, for our construction to have a chance to be meaningful, these operations of raising and lowering much each terminate which can only happen if there are largest and smallest (most negative) value of m, m_{max} and m_{min} such that the next state in the series has zero norm. Thus, m_{max} and m_{min} must make the corresponding norms computed in Eqs. 198 and 199 vanish:

$$(j + 1 + m_{\mathrm{max}})(j - m_{\mathrm{max}}) = 0 \quad \text{and} \quad (j + 1 - m_{\mathrm{min}})(j + m_{\mathrm{min}}) = 0. \tag{201}$$

Since we must require $m_{\max} > m_{\min}$, these equation can be solved by $m_{\max} = j$ and $m_{\min} = -j$. Adding these two equation together we discover $2j = m_{\max} - m_{\min}$ which is an integer. Thus, j must be an integer multiple of half-integers. The total number of states connected by the operators $J_x \pm J_y$ is thus $m_{\max} - m_{\min} + 1 = 2j + 1$.

We can completely determine all three operators as $2j+1 \times 2j+1$ matrices if we decide on the phases of the $2j+1$ basis states that we have just obtained by raising and lowering our starting state $|j, m\rangle$. We will do this by requiring:

$$(J_x + iJ_y)|j, m\rangle = \sqrt{(j + 1 + m)(j - m)}|j, m + 1\rangle \qquad (202)$$

$$(J_x - iJ_y)|j, m\rangle = \sqrt{(j + 1 - m)(j + m)}|j, m - 1\rangle \qquad (203)$$

where the second equation is actually implied by the first if a Hermitian conjugate is taken. This choice of the relative phase of the $2j + 1$ states $\{|j, m\rangle\}$, $-j \le m \le j$ is called the Condon and Shortley phase convention.

Finally, we will use these equations to write out explicitly what the three matrices are for the two cases $j = 1/2$ and $j = 1$. First when $j = 1/2$ the number of states is $2j + 1 = 2$ and the matrices will be 2×2. The easiest is J_z which is diagonal with eigenvalues $\pm \hbar/2$:

$$J_z = \begin{pmatrix} +\frac{\hbar}{2} & 0 \\ 0 & -\frac{\hbar}{2} \end{pmatrix}. \qquad (204)$$

To determine J_x and J_y we can start with the raising and lowering operators are can be easily deduced from their action on our states $|j, \pm 1/2\rangle$:

$$J_x + iJ_y = \begin{pmatrix} 0 & \hbar \\ 0 & 0 \end{pmatrix} \quad J_x - iJ_y = \begin{pmatrix} 0 & 0 \\ \hbar & 0 \end{pmatrix}. \qquad (205)$$

since $\sqrt{(j + 1 \pm m)(j \mp m)}$ evaluates to one for $j = 1/2$ and $m = \mp 1/2$. We can then add and subtract these equations to find:

$$J_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad J_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad (206)$$

which provides a derivation of our earlier expression using the Pauli matrices σ_i: $J_i = (\hbar/2)\sigma_i$, $1 \le i \le 3$.