qmd212

Rotations and Angular Momentum

Robert B. Griffiths
Version of 27 September 2012

Contents

---

1 Unitary Transformations

★ Coordinate transformations play an important role in all branches of physics. Those that occur most frequently in quantum mechanics are represented by *unitary* operators on the quantum Hilbert space.

• In fact, one can regard every unitary operator as representing in some sense a coordinate transformation, though this is not always the most helpful point of view.

★ Reminder. A unitary operator U is one such that

$$U^\dagger U = I = U U^\dagger. \tag{1}$$

• On a finite dimensional Hilbert space one of these equalities implies the other, so only one has to be checked.

★ Unitary operators *preserve inner products*, and thus *preserve norms*:

$$|\psi'\rangle = U|\psi\rangle, \quad |\phi'\rangle = U|\phi\rangle, \quad \langle\phi'|\psi'\rangle = \langle\phi|\psi\rangle, \quad \|\psi'\| = \sqrt{\langle\psi'|\psi'\rangle} = \|\psi\| = \sqrt{\langle\psi|\psi\rangle}. \tag{2}$$

□ Exercise. Prove it.

◦ In fact, as sophisticates can show you, on a complex space a linear operator which preserves all norms will also preserve all inner products. But the key thing is to remember is that a unitary operator preserves all inner products (and therefore all norms).

• This is analogous to the fact that translations and rotations in ordinary three-dimensional space preserve distances between points, and also angles between intersecting lines.

• Note that while a simple application of the unitary U is needed to transform a ket, the corresponding transformation on an operator,

$$A' = U A U^\dagger, \tag{3}$$

requires the use of both the unitary U and its adjoint $U^\dagger$, the inverse of U. That this is "natural" follows from the observation that for a simple dyad,

$$|\psi'\rangle\langle\phi'| = U\big(|\psi\rangle\langle\phi|\big)U^\dagger \tag{4}$$

if we use the definitions of $|\psi'\rangle$ and $|\phi'\rangle$ in (2), along with the observation that $(U|\phi\rangle)^\dagger = \langle\phi|U^\dagger$.

◦ A helpful mnemonic for remembering (3) is to think of U as moving an object from a previous location to a new location, $U^\dagger$ as the reverse, moving the object from the new to the old location, and A as some operation carried out on the object at the old location. Then A' consists in first transporting the object from the new location to the old location using $U^\dagger$, carrying out the operation A at the old location, and transporting the result back to the new location using U. To make this more concrete, think of A as the operation of sawing a block of wood into two pieces.

★ Two orthonormal bases $\{|b_j\rangle\}$ and $\{|c_j\rangle\}$ for the same Hilbert space can be regarded as alternative coordinate systems which can be related by a unitary transformation

$$U = \sum_j |c_j\rangle\langle b_j|; \quad |c_j\rangle = U|b_j\rangle, \quad [c_j] = U[b_j]U^\dagger, \tag{5}$$

where we use the abbreviation $[\psi] = |\psi\rangle\langle\psi|$ for a the projector on the space spanned by the normalized ket $|\psi\rangle$. Thus the basis kets are transformed by U, and the corresponding projectors using the pair U and $U^\dagger$, as in (3) and (4).

□ Exercise. Show that the operator U defined in (5) is unitary.

• It follows from the definition in (2) that the result of applying a unitary operator to *any* orthonormal basis is a collection of kets that form an orthonormal basis. Conversely, if one has a linear operator that maps *some* orthonormal basis into an orthonormal basis, it must be a unitary operator in light of the construction in (5).

2 Rotations in Space

★ Rotating an orthogonal coordinate system in three dimensions yields another orthogonal coordinate system.

2.1 Two dimensions

★ For simplicity, start with a vector $\vec{v}$ in two dimensions, Fig. 1(a), which is rotated in what is called an *active rotation* by an angle θ to produce a vector $\vec{v}\,'$. The components are related by

$$\vec{v}\,' = R\vec{v}, \quad \begin{pmatrix} v'_x \\ v'_y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} v_x \\ v_y \end{pmatrix}. \tag{6}$$

The 2×2 matrix R is a unitary matrix with real components, thus a real orthogonal matrix.

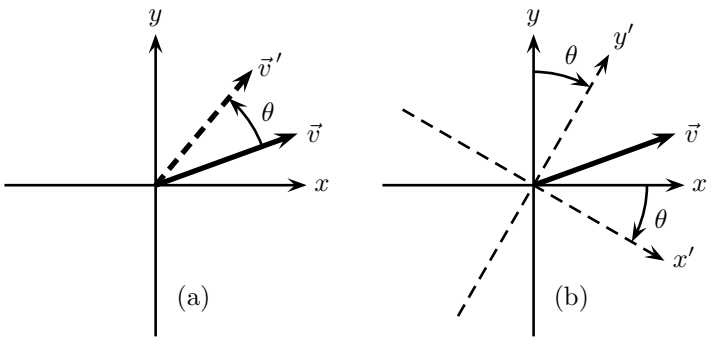

Figure 1: (a) Vector rotated by angle θ, coordinates fixed. (b) Vector fixed, coordinates rotated by angle $-\theta$. In both cases the new components of the vector are related to the old ones by (6).

• A mnemonic for the minus sign in the matrix is to consider a vector initially along the y axis, which after a small rotation will have a negative x component.

★ One may prefer to leave the vector in place and rotate the coordinate system. This is called a *passive rotation*. See Fig. 1(b). Let v'_x, v'_y be the coordinates of the unmoved vector in the new coordinate system. Then (6) is again correct, but notice that in order to achieve this we had to rotate the coordinate system in the opposite sense to that used for the vector in Fig. 1(a).

• The difference between active and passive transformations can be the source of endless confusion, and there seems to be no remedy. The author of these notes will henceforth use *active* transformations: the coordinate system is nailed down and *stays there*.

◦ Here is a little exercise that may help in seeing why the active and passive rotations seem to move in opposite directions. Let us adopt Dirac notation to the matter at hand, and write

$$\vec{v} = |v\rangle, \quad v_x = \langle \hat{x}|v\rangle, \quad v_y = \langle \hat{y}|v\rangle \tag{7}$$

in a fairly obvious notation, where $\hat{x}$ and $\hat{y}$ are unit vectors along the x and y axes, respectively. Then write

$$\vec{v}\,' = |v'\rangle = R|v\rangle, \quad v'_x = \langle \hat{x}|v'\rangle = \langle \hat{x}|R|v\rangle = \langle \hat{x}|Rv\rangle = \langle R^\dagger \hat{x}|v|.\rangle \tag{8}$$

And of course there is a similar way to write v'_y. The point is that v'_x comes about *either* from applying R to $\vec{v} = |v\rangle$ or from applying $R^\dagger$ (in the case of a real matrix $R^\dagger$ is the same as the transpose R^{T}) to the coordinate vector $\hat{x}$. The point is that R and $R^\dagger$, whose product is the identity I, are rotations in the opposite direction.

2.2 Three dimensions

★ Rotations in two dimensions commute with each other: $R(\theta)R(\bar{\theta}) = R(\bar{\theta})R(\theta)$. But in three dimensions this is no longer the case: in general, two rotations *do not commute*, and it makes a difference in which order they are applied.

□ Exercise. Lay down a book in front of you in such a way that you you are looking at the front cover, with the spine of the book on the left. Rotate the book 90° counterclockwise looking down on it from above, and then take the rotated book in your right hand and rotate it 90° towards you. Next do the two 90° rotations in the opposite order. Notice the difference?

★ We can represent rotations by 3×3 matrices which when acting on a column vector produce a new column vector. (Remember, these are active rotations: the vector gets pushed to some other position.)

★ Example. The matrix

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9}$$

represents a rotation by an angle θ in the positive direction about the z axis.

◦ This should be pretty obvious, since this is the rotation of the x and y components of the vector that we encountered earlier in (6), and the z component is held fixed.

◦ Positive direction: Put the thumb of your right hand upwards along the z axis; your fingers curl in the direction of the rotation when θ is positive.

□ Exercise. Write down the matrices $R_x(\theta)$ and $R_y(\theta)$ for rotations of an angle θ about the x and about the y axes.

□ Exercise. Show that $R_x(\pi/2)R_z(\pi/2)$ is not equal to $R_z(\pi/2)R_x(\pi/2)$.

★ Euler angles. Any rotation in 3 dimensions can be carried out using the following sequence: Rotate by an angle α about the z axis, an angle β about the x axis and an angle γ about (once again!) the z axis. The triple (α, β, γ) is known as the *Euler angles*.

◦ Warning! This is only one of the ways in which Euler angles can be defined. Note in particular that we are rotating about axes fixed in space. It is also possible to attach axes to the object being rotated and make

subsequent rotations about these axes fixed to the object. And even if one fixes the axes in space (which is what we will be doing) there are other choices of three axes to rotate about.[1]

◦ Fortunately, in the remainder of these notes we shall not use Euler angles. . . .

★ Instead of the Euler angles, a rotation in three dimensions can be represented by a single vector $\vec{\omega}$, interpreted in the following way. The direction of $\vec{\omega}$ gives the axis of rotation, and the magnitude of the rotation (in radians) by the length $|\vec{\omega}|$ of this vector. If we write $\vec{\omega} = \omega\hat{n}$, where $\hat{n}$ is a unit vector, then positive ω corresponds to the direction your fingers curl when your right thumb is in the direction $\hat{n}$.

◦ Note that in this description there are three real parameters, the three components of $\vec{\omega} = (\omega_x, \omega_y, \omega_z)$, just as there are three independent Euler angles.

◦ It is convenient to restrict $|\vec{\omega}|$ to the range from 0 to π.

□ Exercise. Why does this suffice?

3 Quantum Rotations

3.1 Introduction

★ Suppose we have a description of a quantum object using some ket in a Hilbert space. The description may require the use of a coordinate system. Thus in the case of a spin-half particle the ket $|z^+\rangle$ tells us that the spin angular momentum is positive in the z direction. But in writing this down, we must know that the z direction is. Suppose that it is upwards, towards the ceiling. But now what about $|x^+\rangle$? That requires choosing some horizontal direction for x. With that we are done, because the y direction will be perpendicular to both x and z, and the triple x, y, z forms a right-handed coordinate system. Thus we know what $|y^+\rangle$ means, and also $|w^+\rangle$ for w any direction in space.

• And suppose we change our coordinate system? If the new x axis is directed towards the ceiling, what we earlier called $|z^+\rangle$ ought now to be $|x^+\rangle$ in the new coordinate system.

• In discussing changes of coordinate system, passive transformations seem more "natural", but in these notes we will stick to active transformations: we think of rotating the particle while leaving the coordinate system firmly attached to the laboratory.

◦ But the laboratory is attached to the earth, and the earth is rotating. Well, let us do our experiments quickly. It didn't take very long for a silver atom to get from one end of the Stern-Gerlach magnet to the other. . . .

• And it really is possible to "rotate" the spin direction of a spin-half particle; this can be done by applying a suitable magnetic field for a suitable time interval.

• There are lots of very elegant things to say about rotations and angular momentum in quantum mechanics. These notes are restricted to the basics which every student of the subject should know (preferably by heart).

★ FIRST VERY IMPORTANT FORMULA. The unitary quantum mechanical operator which represents an active rotation $\vec{\omega}$ can be written as

$$R(\vec{\omega}) = \exp[-i\vec{\omega} \cdot \vec{J}], \tag{10}$$

where $\vec{\omega} \cdot \vec{J} = \omega_x J_x + \omega_y J_y + \omega_z J_z$, and J_x, J_y and J_z are (dimensionless) *angular momentum operators*. They are Hermitian operators, and this implies that $R(\vec{\omega})$ is a unitary operator. In fact,

$$R^\dagger(\vec{\omega}) = \exp[+i\vec{\omega} \cdot \vec{J}], \tag{11}$$

represents a rotation about the same axis by the same amount, but in the opposite sense. Thus it is the inverse of $R(\vec{\omega})$: $R^\dagger(\vec{\omega})R(\vec{\omega}) = I$.

◦ Angular momentum has the same dimensions as Planck's constant, and in quantum mechanics it is convenient to measure angular momentum in units of $\hbar$. If one does that the angular momentum operators are dimensionless. However, if one uses dimensioned angular momentum operators (e.g., expressed in units of kg m^2/s) then it is necessary to rewrite (10) as

$$R(\vec{\omega}) = \exp[-i\vec{\omega} \cdot \vec{J}/\hbar], \tag{12}$$

[1] Various conventions are discussed in H. Goldstein, *Classical Mechanics*, 2d edition (Addison-Wesley, 1981). See 'Euler angles' in the index.

in order for the exponent to be dimensionless. Usually it is pretty obvious where to insert extra factors of $\hbar$ into formulas if one is using dimensioned operators, and as the formulas are simpler with $\hbar$ missing (or, as one often says, in units in which $\hbar = 1$), the $\hbar$'s will be omitted in what follows.

★ SECOND VERY IMPORTANT FORMULA

$$[J_x, J_y] = J_x J_y - J_y J_x = iJ_z; \quad [J_y, J_z] = iJ_x; \quad [J_z, J_x] = iJ_y. \tag{13}$$

• These *commutation relations* of the angular momentum operators should be committed to memory. It suffices to remember the first, and then keep in mind that the others are obtained by cyclic permutation of $xyz \to yzx \to zxy$.

3.2 Exponentials of operators

• The first important formula will do us no good unless we know what to do with exponentials of operators. So a brief interlude on that subject.

★ A matrix M can be diagonalized by a unitary transformation U if it can be written as:

$$M = UDU^\dagger, \quad D = \begin{pmatrix} m_1 & 0 & \dots & 0 \\ 0 & m_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m_n \end{pmatrix}. \tag{14}$$

The diagonal matrix D has the eigenvalues of M on its diagonal. If $f(z)$ is some numerical (perhaps complex) function that is defined for each of the m_j, then we can *define*

$$f(D) := \begin{pmatrix} f(m_1) & 0 & \dots & 0 \\ 0 & f(m_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(m_n) \end{pmatrix}, \quad \text{and } f(M) := Uf(D)U^\dagger. \tag{15}$$

◦ Comment. While we will not need it for discussing rotations, the same trick will work for a matrix which can be diagonalized by a similarity transformation, $M = SDS^{-1}$, where S is any invertible matrix, that is, a matrix which has an inverse S^{-1}, with $SS^{-1} = S^{-1}S = I$. Just replace U by S and $U^\dagger$ by S^{-1} in the above.

◦ Comment. There can be troublesome cases. E.g., if $f(z) = \ln(z)$ and one of the eigenvalues of M is 0. From this perspective $f(z) = e^z$ is particularly nice, since it is defined, and in fact analytic, for *any* (complex) value of z.

★ A second way to define $f(M)$ is to use a power series expansion for $f(z)$, assuming it has one. In particular, we can write

$$\exp[M] = I + M + M^2/2! + M^3/3! + \cdots, \tag{16}$$

where M^2, M^3, etc. are matrix products. The series for $\exp[M]$ converges for any finite matrix M.

□ Exercise. Show that (16) agrees with (15) if M can be diagonalized and $f(z) = e^z$, by showing that M^n has a simple form in terms of D.

★ While the preceding discussion is in terms of matrices it works in the obvious way for operators. Given the spectral form of a normal operator A with eigenvalues α_j we can define $f(A)$ using:

$$A = \sum_j \alpha_j P_j, \quad f(A) = \sum_j f(\alpha_j) P_j, \tag{17}$$

where $\{P_j\}$ is the corresponding decomposition of the identity. Or if $f(z)$ has a power series expansion, one can use that, as in (16).

4 Spin Half

4.1 Kets

★ Let w be a direction in space corresponding to the polar angles θ and ϕ, where θ is the angle between this direction and the z axis, and ϕ the angle between this direction projected on the x, y plane and the x axis. (This is the usual choice of polar angles.) The kets $|w^+\rangle$ and $|w^-\rangle$ corresponding to a spin angular momentum of $+1/2$ and $-1/2$ (in units of $\hbar$) in the w direction are then defined to be:

$$|w^+\rangle = \cos(\theta/2)|z^+\rangle + e^{i\phi}\sin(\theta/2)|z^-\rangle, \quad |w^-\rangle = \sin(\theta/2)|z^+\rangle - e^{i\phi}\cos(\theta/2)|z^-\rangle, \tag{18}$$

where $|z^+\rangle$ and $|z^-\rangle$ are the orthonormal kets corresponding to the z component of angular momentum being positive and negative, respectively.

□ Exercise. Check that $|w^+\rangle$ and $|w^-\rangle$ defined in this way are normalized and orthogonal to each other.

∘ In quantum information theory $|z^+\rangle$ and $|z^-\rangle$ are written as $|0\rangle$ and $|1\rangle$.

• Frequently used:

$$|x^+\rangle = \frac{|z^+\rangle + |z^-\rangle}{\sqrt{2}}, \quad |x^-\rangle = \frac{|z^+\rangle - |z^-\rangle}{\sqrt{2}}, \quad |y^+\rangle = \frac{|z^+\rangle + i|z^-\rangle}{\sqrt{2}}, \quad |y^-\rangle = \frac{|z^+\rangle - i|z^-\rangle}{\sqrt{2}}. \tag{19}$$

• Note that (18) employs a particular choice of overall phase; other choices are possible. The choice of overall phase does not affect the physics, but can cause confusion if one starts a problem using one convention and then switches to another.

★ Bloch sphere. There is a convenient geometrical representation of the states of a spin-half particle by means of a sphere of unit radius centered at the origin. Let w be a point on the surface of the sphere with polar coordinates θ, ϕ. Then one thinks of this point as associated with the projector $|w^+\rangle\langle w^+|$ onto the ray corresponding to the ket $|w^+\rangle$ as defined in (18).

• Because it represents projectors, the Bloch sphere picture is not affected by the choice of the overall phase used for defining the kets in (18).

• Two states of a spin-half particle are *orthogonal* if the two projectors correspond to *antipodes* on the Bloch sphere, points at opposite ends of a diameter passing through the center of the sphere. For example, $[z^+] = |z^+\rangle\langle z^+|$ is represented by the north pole and $[z^-]$ by the south pole.

∘ The fact that antipodes are 180° apart when viewed from the center of the sphere is somewhat confusing, as the corresponding states are orthogonal, which one usually thinks of as perpendicular, i.e., with an angle of 90° between them. Just one of many instances where there is an annoying factor of 2 needed for interpreting Bloch sphere angles. One just has to get used to this. The advantages of this geometrical representation far outweigh the annoyances.

• The points on the surface of the Bloch sphere correspond to *pure states*, one-dimensional rays in the Hilbert space. Points in the interior are used to represent density operators corresponding to mixed states, which lie outside the scope of these notes.

4.2 Angular momentum operators

★ The matrices of the angular momentum operators for spin half are traditionally denoted by S_x, S_y, and S_z in place of J_x, etc. Their matrices in the standard $\{|z^+\rangle, |z^-\rangle\}$ basis (or $\{|0\rangle, |1\rangle\}$ in quantum information theory) are the Pauli matrices apart from a factor of $1/2$:

$$2S_x = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad 2S_y = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad 2S_z = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{20}$$

∘ Note that the order is z^+ (top row, first column) followed by z^- (bottom row, second column). Thus $\langle z^+|\sigma_z|z^+\rangle = +1$, $\langle z^+|\sigma_y|z^-\rangle = -i$.

• The Pauli matrices have the following useful properties. They are Hermitian, something which is true of any angular momentum operator ($J_x = J_x^\dagger$). But in addition, and these are special for spin half, each Pauli matrix is unitary, and its square is the identity matrix I. Furthermore,

$$\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z, \quad \sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x, \quad \sigma_z\sigma_x = -\sigma_x\sigma_z = i\sigma_y. \tag{21}$$

□ Exercise. Use matrix multiplication to check the first of the equations in (21)

□ Exercise. Show that (21) implies that (13) holds for spin half.

□ Exercise. Write the projectors $[z^+]$ and $[z^-]$ in terms of I and σ_z, and check that the square of each projector is equal to itself, using properties of σ_z. Write $[x^+]$, $[x^-]$, $[y^+]$, and $[y^-]$ as 2×2 matrices.

4.3 Rotations

★ Let us start by considering a particular example: a rotation by an angle ω about the z axis. What will this do to the ket $|x^+\rangle$? Let us apply (10) and see what happens:

$$R(\omega\hat{z})|x^+\rangle = e^{-i(\omega/2)\sigma_z}|x^+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\omega/2} & 0 \\ 0 & e^{+i\omega/2}\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\omega/2}\\ e^{+i\omega/2}\end{pmatrix}. \tag{22}$$

∘ Here we used the fact that σ_z, and thus $-i(\omega/2)\sigma_z$ is a diagonal matrix in order to set its exponential equal to the diagonal matrix with exponentials of its eigenvalues on the diagonal.

• Suppose that $\omega = \pi/2$, in which case we might expect the rotation to map $|x^+\rangle$ onto $|y^+\rangle$. However, the final column vector in (22) is $(e^{-i\pi/4}, e^{i\pi/4})^{\mathrm{T}}$ rather than the $(1, i)^{\mathrm{T}}$ we might have expected, see (19). [To save space in writing it out it is convenient to transpose the column vector to a row vector; note that the transpose, indicated by $^{\mathrm{T}}$, involves no complex conjugation, unlike the adjoint †.] Is something wrong? No, because we can extract an overall factor of $e^{-i\omega/2}$ from the column vector on the right side of (22) to leave $(1, e^{i\omega})^{\mathrm{T}}$, which is just what one would expect: set $\theta = \pi/2$ and $\phi = \omega$ in (18). So the rotation is doing what we expect.

□ Exercise. Show that with $\omega = \pi/2$ matrix multiplication yields the result

$$R[(\pi/2)\hat{z}]\,\sigma_x\,R^\dagger[(\pi/2)\hat{z}] = \sigma_y. \tag{23}$$

Does this seem sensible? Might one have expected an extra phase factor in this case, say $e^{i\epsilon}\sigma_y$, in place of σ_y? Why or why not?

★ Because σ_x is not diagonal, we cannot evaluate something like $R(\omega\hat{x})|y^+\rangle$ by simply taking exponentials of the diagonal elements of $-i(\omega/2)\sigma_x$. However, in this case we can sum the power series in (16) in closed form. The first few terms are

$$\exp[-i(\omega/2)\sigma_x] = I - i\frac{(\omega/2)}{1!}\sigma_x - \frac{(\omega/2)^2}{2!}(\sigma_x)^2 + i\frac{(\omega/2)^3}{3!}(\sigma_x)^3 + \cdots. \tag{24}$$

Now use the fact that $(\sigma_x)^2 = I$, so $(\sigma_x)^3 = \sigma_x$, and so forth. This allows one to rewrite the right side of (24) as something times I plus something times σ_x, and one recognizes the power series for sine and cosine, so

$$\exp[-i(\omega/2)\sigma_x] = \cos(\omega/2)I - i\sin(\omega/2)\sigma_x. \tag{25}$$

□ Exercise. Use (25) to find what a rotation by $\pi/2$ around the x axis does to $|x^+\rangle$ and $|y^+\rangle$. Is it what you expect?

□ Exercise. Use (25) and matrix multiplication to work out what a rotation of $\pi/2$ about the x axis does to σ_y.

★ The same trick can be applied to obtain an explicit matrix $R(\vec{\omega})$ for *any* $\vec{\omega}$. First write $\vec{\omega} = \omega\hat{n}$ for a unit vector $\hat{n} = (n_x, n_y, n_z)$. Next define

$$\sigma_{\hat{n}} = \hat{n}\cdot\vec{\sigma} = n_x\sigma_x + n_y\sigma_y + n_z\sigma_z. \tag{26}$$

The rest can be left as an exercise.

□ Exercise. Write out $\sigma_{\hat{n}}$ in (26) as a 2×2 matrix and evaluate its square by matrix multiplication. Then use the same strategy employed for (24), and obtain $R(\vec{\omega})$ as a 2×2 matrix. Check that the answer is what you expect for $\vec{\omega}$ along the z axis, and along the x axis.

★ Thinking of the Bloch sphere as a rigid object rotated (we are using active rotations) in three-dimensional space is a very helpful way of visualizing the different possible unitary operations which are

possible for spin half. Up to an overall phase, i.e., $e^{i\epsilon}U$ in place of U, *every* unitary matrix on a 2-dimensional (complex) Hilbert space can be thought of as a *proper* rotation of the Bloch sphere.

□ Exercise. Establish this result. [Hint. First find the most general form of a 2×2 unitary matrix. Then compare with $R(\vec{\omega})$ obtained in a previous exercise.]

• A proper rotation is one which does not involve any reflection. The transformation which in three dimensions maps $+x$ into $+x$, $+y$ into $+y$, but z into $-z$ is an example of an improper rotation. Any improper rotation can be performed by first carrying out a parity operation, $x \to -x$, $x \to -y$, $z \to -z$, followed by a proper rotation.

5 Larger Values of Spin

5.1 Kets and angular momentum operators

★ A very elegant theory of angular momentum and rotations can be based upon the two VERY IMPORTANT FORMULAS of Sec. 3.1 along with one additional idea. The operator

$$J^2 := J_x^2 + J_y^2 + J_z^2, \tag{27}$$

the square of the total angular momentum, has the property that it is Hermitian and commutes with all three components of angular momentum:

$$[J^2, J_x] = [J^2, J_y] = [J^2, J_z] = 0. \tag{28}$$

□ Exercise. Check that (28) is a consequence of (13).

• This means that it is possible to simultaneously diagonalize J^2 and *one* of the three components J_x, J_y, J_z. The traditional choice is to diagonalize J^2 and J_z. Thus one can find a basis of the Hilbert space consisting of kets which are simultaneously eigenkets of J^2 and J_z.

★ There is a very elegant algebraic approach, found in many textbooks, which uses nothing but the commutation relations (13) in order to determine the possible values of the eigenvalues of J^2 and J_z. Here is a somewhat abbreviated presentation.

★ Define the *raising* and *lowering* operators—the reason for these names will appear shortly—

$$J_+ := J_x + iJ_y, \quad J_- := J_x - iJ_y, \tag{29}$$

• Note that J_+ and J_- are Hermitian conjugates of each other, $J_- = (J_+)^\dagger$, $J_+ = (J_-)^\dagger$.

• The following are immediate consequences of (13) and (27):

$$[J_z, J_+] = J_+, \quad [J_z, J_-] = -J_-, \tag{30}$$

$$J_+J_- = J^2 - J_z^2 + J_z, \quad J_-J_+ = J^2 - J_z^2 - J_z. \tag{31}$$

□ Exercise. Check that these are correct.

• It turns out in view of the following argument that the possible eigenvalues of J^2 are of the form $j(j+1)$ where j is either a nonnegative integer or half integer, thus the possibilities are

$$j = 0,\ 1/2,\ 1,\ 3/2,\ 2,\ 5/2,\ \ldots. \tag{32}$$

Hence the possible eigenvalues of the operator J^2 are $0,\ 3/4,\ 2,\ 15/4,\ \ldots$.

◦ Spin half means $j = 1/2$, and the eigenvalues of $J^2 = S^2 = (\sigma_x^2 + \sigma_y^2 + \sigma_z^2)/4$ are indeed $1/2(1+1/2) = 3/4$.

★ Consider a ket $|j, m\rangle$ which is simultaneously an eigenstate of J^2 and J_z:

$$J^2|j,m\rangle = j(j+m)|j,m\rangle, \quad J_z|j,m\rangle = m|j,m\rangle, \tag{33}$$

so we are labeling the ket with the eigenvalues of J^2 and J_z, except that in place of $j(j+1)$ we use j as a label.

• Now consider the ket $J_+|j,m\rangle$, and use (30) to show that

$$J_z\big(J_+|j,m\rangle\big) = J_+J_z|j,m\rangle + [J_z,J_+]|j,m\rangle = J_+\big(m|j,m\rangle + |j,m\rangle\big) = (m+1)\big(J_+|j,m\rangle\big). \tag{34}$$

◦ Conclusion: $J_+|j,m\rangle$ is an eigenstate of J_z with eigenvalue $m+1$. Thus applying J_+ has increased the eigenvalue of J_z by 1, justifying the name "raising operator".

◦ As for J^2, it is fairly obvious that the $J_+|j,m\rangle$ is an eigenvector with the same eigenvalue as $|j,m\rangle$, namely $j(j+1)$.

□ Exercise. Or is it so obvious? Provide a short argument.

• What can be done once can be done twice, or as often as we like. Thus $J_+^2|j,m\rangle$ will be an eigenstate of J_z with eigenvalue $m+2$, $J_+^3|j,m\rangle$ with eigenvalue $m+3$, and so forth, all with the same eigenvalue $j(j+1)$ of J^2.

★ But something is wrong here. We cannot construct a state $|j,m\rangle$ with m as large as we want. In fact we cannot construct a state with $m > \sqrt{j(j+1)}$, for the following reason. A simultaneous eigenstate of J^2 and J_z is also an eigenstate of $J_x^2 + J_y^2$,

$$\big(J_x^2 + J_y^2\big)|j,m\rangle = \big(J^2 - J_z^2\big)|j,m\rangle = \big(j(j+1) - m^2\big)|j,m\rangle. \tag{35}$$

But $J_x^2 + J_y^2$ is a positive operator and cannot have negative eigenvalues. Consequently

$$|m| \le \sqrt{j(j+1)}. \tag{36}$$

◦ That J_x^2 is a positive (nonnegative) operator follows from the fact that J_x is Hermitian with real eigenvalues, and thus J_x^2 is Hermitian with positive (nonnegative) eigenvalues. The same for J_y^2. That the sum $Q+R$ of two positive operators Q and R is positive follows from the fact that for any $|\psi\rangle$

$$\langle\psi|Q+R|\psi\rangle = \langle\psi|Q|\psi\rangle + \langle\psi|R|\psi\rangle \ge 0. \tag{37}$$

• What, then, is wrong with our argument that if we simply continue to apply J_+ time after time we can create states with the eigenvalue of J_z as large as we want? What is wrong is that an eigenvector must, by definition, be *nonzero*.

◦ By contrast, an *eigenvalue* can very well be zero.

★ Let us assume that $|j,m\rangle$ is normalized. Applying J_+ to it will yield another ket, which in general is not normalized. Let us write

$$J_+|j,m\rangle = c|j,m+1\rangle. \tag{38}$$

Then by assuming that both $|j,m\rangle$ and $|j,m+1\rangle$ are normalized, we can compute c by equating the inner products of the two sides of (38)

$$\langle j,m|J_-J_+|j,m\rangle = \langle j,m|\big(J^2 - J_z^2 - J_z\big)|j,m\rangle = j(j+1) - m(m+1) = |c|^2\langle j,m+1|j,m+1\rangle = |c|^2, \tag{39}$$

where we have used (31). This equation determines c up to a phase, and is customary to choose the phase to be positive, and write (38) in the form

$$J_+|j,m\rangle = \sqrt{j(j+1) - m(m+1)}\;|j,m+1\rangle. \tag{40}$$

• Thus if $m = j$ we have $J_+|j,m\rangle = 0$. It is this fact that prevents us from generating kets with indefinitely large values of m. In fact, the fact that we cannot keep increasing m means that *there must be* a value of m such that $m = j$.

★ In a similar way one can show that J_- applied to a ket $|j,m\rangle$ yields a constant times a ket $|j,m-1\rangle$, or perhaps 0, and just as in (39) one can determine the absolute value of this constant. Choosing a positive phase leads to

$$J_-|j,m\rangle = \sqrt{j(j+1) - m(m-1)}\;|j,m-1\rangle. \tag{41}$$

• Observe that in this case the right side vanishes for $m = -j$, and this fact will prevent the repeated application of J_- leading to values of m so negative that they violate (36).

★ It is important to notice what we have just done. Starting with a *single* ket $|j,m\rangle$ which is an eigenket of both J^2 and J_z we have produced a *cluster* of $2j+1$ nonzero kets, with labels $m=-j,-j+1,\dots j$, orthogonal to one another, with the property that we can generate the entire cluster starting with just one, any one, of these kets and applying suitable angular momentum operators a suitable number of times. The set of all linear combinations of the kets in the cluster forms a $2j+1$-dimensional subspace of the Hilbert space, which we will henceforth refer to as a "cluster space".

◦ In the language of group theory we are dealing with an *irreducible representation* of the rotation group.

★ Using these $2j+1$ orthonormal kets one can write down matrices for the operators J_x, J_y and J_z. The usual choice is to order rows (descending) and columns (left to right) in the order of decreasing m.

• The matrix of J_z is diagonal, with $\langle j,j|J_z|j,j\rangle = +j$ in the upper left corner.

• The matrices for J_+ and J_- follow from (40) and (41):

$$\begin{aligned}\langle j,m'|J_+|j,m\rangle &= \sqrt{j(j+1)-m(m+1)}\;\delta_{m',m+1},\\ \langle j,m'|J_-|j,m\rangle &= \sqrt{j(j+1)-m(m-1)}\;\delta_{m',m-1}.\end{aligned} \tag{42}$$

• The matrices for J_x and J_y can be written in terms of those for J_+ and J_- by inverting (29)

$$J_x = (J_+ + J_-)/2, \quad J_y = (J_+ - J_-)/2i \tag{43}$$

★ In the case of $j=1/2$, spin half, the kets $|\frac{1}{2},+\frac{1}{2}\rangle$ and $|\frac{1}{2},-\frac{1}{2}\rangle$ are just the $|z^+\rangle$ and $|z^-\rangle$ states encountered earlier, and the raising and lowering operators in this basis (using S in place of J) are:

$$S_+ = \begin{pmatrix}0&1\\0&0\end{pmatrix}, \quad S_- = \begin{pmatrix}0&0\\1&0\end{pmatrix}. \tag{44}$$

□ Exercise. Check that these matrices agree with the definitions $S_+ = S_x + iS_y$, $S_- = S_x - iS_y$, with S_x and S_y defined in terms of the usual Pauli matrices.

□ Exercise. Show that (40) and (41) are correct for spin half.

□ Exercise. Work out the matrices for J_+ and J_- in the case $j=3/2$.

5.2 Spin one

★ Here are eigenkets and operators for $j=1$, spin one, where we use S_x, etc., in place of J_x, etc. Note that the choice of phase for eigenkets is not unique. Matrices and column vectors use eigenvectors $|m\rangle$ of S_z with m decreasing: the top row or first column corresponds to $m=1$.

$$\begin{aligned}|z^+\rangle &= \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad |z^0\rangle = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad |z^-\rangle = \begin{pmatrix}0\\0\\1\end{pmatrix},\\ |x^+\rangle &= \frac{1}{2}\begin{pmatrix}1\\\sqrt{2}\\1\end{pmatrix}, \quad |x^0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix}, \quad |x^-\rangle = \frac{1}{2}\begin{pmatrix}1\\-\sqrt{2}\\1\end{pmatrix},\\ |y^+\rangle &= \frac{1}{2}\begin{pmatrix}1\\i\sqrt{2}\\-1\end{pmatrix}, \quad |y^0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix}, \quad |y^-\rangle = \frac{1}{2}\begin{pmatrix}1\\-i\sqrt{2}\\-1\end{pmatrix},\end{aligned} \tag{45}$$

$$\begin{aligned}S_z &= \begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}, \quad S_x = \frac{1}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}, \quad S_y = \frac{1}{\sqrt{2}}\begin{pmatrix}0&-i&0\\i&0&-i\\0&i&0\end{pmatrix},\\ S_+ &= \sqrt{2}\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}, \quad S_- = \sqrt{2}\begin{pmatrix}0&0&0\\1&0&0\\0&1&0\end{pmatrix}.\end{aligned} \tag{46}$$

□ Exercise. Check that the matrices for S_+ and S_- (J_+ and J_-) agree with (40) and (41), respectively

□ Exercise. Check that $[S_x, S_y] = iS_z$ by multiplying out the matrices.

• Consider a rotation about the z axis. From (46) we see that

$$R(\omega\hat{z}) = e^{-i\omega S_z} = \begin{pmatrix} e^{-i\omega} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\omega} \end{pmatrix} \tag{47}$$

If we set $\omega = \pi/2$ and apply this to $|x^+\rangle$, (45), the result is $-i\,|y^+\rangle$, a phase times $|y^+\rangle$, which is what we might expect.

□ Exercise. What might you expect $R(\omega\hat{z})S_xR(-\omega\hat{z})$ to be equal to? Multiply out the matrices and see if you're right.

□ Exercise. What are the eigenvalues of $R(\omega\hat{y})$? A physicist's answer (plausible but not a proof) will suffice.

□ Exercise. In the case of spin half the projectors onto $m = 1/2$ and $m = -1/2$ can be written as $[z^+] = S_z + I/2$; $[z^-] = -S_z + I/2$. Can you find a comparable formula for $[z^+]$, $[z^0]$, $[z^-]$ for spin 1? You may need to use S_z^2 in addition to S_z.

6 Hamiltonian Invariant Under Rotations

★ An important application of the theory of quantum angular momentum is in classifying the energy eigenstates of systems for which the Hamiltonian H is *invariant under rotations*, that is, H commutes with all the rotation operators:

$$[J_x, H] = 0 = [J_y, H] = [J_z, H]. \tag{48}$$

• A consequence of (48) is that

$$R(\vec{\omega})HR^\dagger(\vec{\omega}) = R(\vec{\omega})HR(-\vec{\omega}) = e^{-i\vec{\omega}\cdot\vec{J}}He^{+i\vec{\omega}\cdot\vec{J}} = H, \tag{49}$$

for any $\vec{\omega}$, that is, H is left unchanged by any rotation.

□ Exercise. Prove (49) is a consequence of (48). [Hint. Consider a particular choice of $\vec{\omega}$. Show that $\vec{\omega}\cdot\vec{J}$ is a Hermitian operator that commutes with H. Hence one can find an orthonormal basis in which both are diagonal. Write out the matrices involved in (49) in this basis and check the final equality.]

★ Because an invariant H commutes with J^2 and J_z, there is an orthonormal basis of the Hilbert space in which all three Hermitian operators are diagonal.

• In particular this means the following. Let $\{P_k\}$ be the decomposition of the identity associated with H in the sense that

$$H = \sum_k \epsilon_k P_k; \quad \epsilon_k \neq \epsilon_{k'} \text{ for } k \neq k', \tag{50}$$

i.e., P_k projects onto the eigenspace $\mathcal{P}_k$ of H associated with energy eigenvalue ϵ_k. Then it is possible to choose an orthonormal basis of the subspace $\mathcal{P}_k$ consisting of simultaneous eigenkets of J^2 and J_z.

• Let $|j, m\rangle$ be one of these eigenkets, using the notation of Sec. 5.1. Then we know that it belongs to a cluster of $2j + 1$ kets which are produced from it by successive applications of J_+ and/or J_-. All of the kets in this cluster must have the same energy, because H commutes with the angular momentum operators.

◦ So the dimension $d_k = \mathrm{Tr}(P_k)$ of the subspace $\mathcal{P}_k$ cannot be less than $2j + 1$ if there is at least one ket present which is an eigenstate of J^2 with eigenvalue $j(j + 1)$.

◦ Of course d_k could be larger, as there might be kets corresponding to two different values of j in this energy eigenspace. Or there could be two linearly-independent clusters having the same j, assuming $d_k \geq 2(2j + 1)$.

★ There is a very interesting folk theorem that says that there is always *only one* cluster present, corresponding to a *single* specific value of j, in any give energy eigenspace, i.e., in any $\mathcal{P}_k$.

◦ Thus if, for example, $d_k = 3$, then $j = 1$ for this k.

• A folk theorem is one that is true in all instances except those in which it is false. This particular folk theorem works very well in that experiments indicate that it holds for a large number of situations in both atomic and nuclear physics.

◦ A glaring exception to the folk theorem is the nonrelativistic hydrogen atom, the energy spectrum of which is worked out in many textbooks. Except for the ground state, the energy eigenspaces are bigger than what the folk theorem would lead one to expect. (As for the real, relativistic hydrogen atom.... The discussion lies outside the scope of these notes.)

★ Readers may have encountered the somewhat crazy notation used in atomic physics whereby a given energy level might carry the label $^2\mathrm{P}_{3/2}$. Each of the three symbols, 2, P, and 3/2, has something to do with angular momentum. For the full explanation the reader should look at a book on atomic physics. For present purposes we note that the final subscript 3/2 denotes the j value: the level is $2j + 1 = 4$-fold degenerate, i.e., the dimension of this subspace of the Hilbert space is 4, and the different states are mapped into linear combinations of one another when one rotates the atom.