Rotation and Angular Momentum

S. Turgut

April 11, 2016

There is a fundamental connection between the angular momentum (a mechanical quantity) and the rotation operators (which is a symmetry group). The angular momentum is the generator of rotations. Because of this connection, many properties of angular momentum operators can be derived from the properties of the rotation group. We will derive these properties (mainly the commutation relations) starting only from this fact. In this way we will be able to deal not only with the orbital angular momentum but also with all the other kinds of angular momenta, like the intrinsic angular momentum of particles called spin. In the following section, we will concentrate mainly on the rotations themselves. After that, we will apply these to quantum mechanical systems.

Rotations

Rotations are transformations of vectors in the 3D real space. It is possible to describe this as a transformation of the components of vectors. Consider a certain rotation operation. Let us denote the operation by \( R \). It represents a rotation that transforms any given vector \( \vec{V} \) to a rotated vector \( \vec{V}' \),

\[
R : \vec{V} \rightarrow \vec{V}' = R\vec{V} .
\]

We can also consider \( R \) as a transformation of the cartesian components of the vectors, in this way, we can think of \( R \) as a 3 × 3 matrix.

\[
V_k \rightarrow V'_k = \sum_\ell R_{k\ell} V_\ell .
\]
\[ \mathbf{V} = \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} \rightarrow \mathbf{V'} = \begin{bmatrix} V'_x \\ V'_y \\ V'_z \end{bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{bmatrix} \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix}. \]

Here \( R_{kl} \) are the elements of the 3x3 rotation matrix \( R \) which completely represents the rotation operation. I am going to write the equation above as \( \mathbf{V'} = R\mathbf{V} \) to simplify the notation.

Rotations do not change the length of the vectors (\( \mathbf{V'} \cdot \mathbf{V'} = \mathbf{V} \cdot \mathbf{V} \)). Using this, we can show that the matrix \( R \) has to be an orthogonal matrix, i.e., the inverse of the rotation can be expressed as the matrix transpose, \( R^{-1} = R^T \).

\[ \mathbf{V'} \cdot \mathbf{V'} = \mathbf{v'}^T \mathbf{v'} = \mathbf{v}^T R^T R \mathbf{v} = \mathbf{v}^T \mathbf{v} \implies R^T R = I \]

The orthogonality also implies that the dot product of two vectors does not change under rotation,

if \( \mathbf{V'} = R\mathbf{V} \text{ and } \mathbf{W'} = R\mathbf{W} \) then \( \mathbf{V'} \cdot \mathbf{W'} = \mathbf{V} \cdot \mathbf{W} \).

Geometrically this means that if all vectors are rotated, the angles between the vectors remain the same.

Because of the invariance of the geometrical relation between vectors, it looks like we can easily relate the rotation of a cross product of two vectors (such as \( \mathbf{V} \times \mathbf{W} \)) to the product of the rotated vectors. Unfortunately there might be a complication if the rotation involves an inversion. We call a rotation proper if \( \det(R) = +1 \). Otherwise if \( \det(R) = -1 \), then it is called an improper rotation. Inversion, with a rotation matrix \( J = -I \) (in other words, the transformation \( \mathbf{V} \rightarrow \mathbf{V'} = J\mathbf{V} = -\mathbf{V} \)) is an example of improper rotations. All improper rotations can be considered as a proper rotation followed by an inversion (If \( R \) is improper, then \( R = JS \) where \( S \) is proper). The transformation rule of the cross products is

if \( \mathbf{U} = \mathbf{V} \times \mathbf{W} \) then \[ \begin{cases} \mathbf{U'} = \mathbf{V'} \times \mathbf{W'} \text{ (proper rotations)}, \\ \overline{\mathbf{U'}} = -\mathbf{V'} \times \mathbf{W'} \text{ (improper rotations)}, \end{cases} \]

where primed vectors are rotated ones.

Inversion is an important symmetry operation in quantum mechanics which is related to parity. But, it has to be treated separately from proper rotations. For this reason, we leave out inversion and all the improper rotations from now on. The two associated symmetry groups are called O(3) and SO(3).
The *orthogonal group* $O(3)$ contains all $3 \times 3$ orthogonal matrices. In other words, it contains all proper and improper rotations.

The *special orthogonal group* $SO(3)$ contains all $3 \times 3$ orthogonal matrices with determinant $+1$. As a result, $SO(3)$ contains only the proper rotations.

Our subject will be $SO(3)$. Note that we define $SO(3)$ as a “subset” of $3 \times 3$ matrices. But, it is much more fruitful to think of as an abstract set of symmetry operations. Later, we will find a way to imagine $SO(3)$ as a 3-dimensional manifold. The most important feature of $SO(3)$ is that it is connected, in other words, any rotation can be obtained as a successive application of small rotations. More on this later.

**Axis-Angle Description of Rotations**

Let $R$ be a proper rotation. Consider its eigenvalues. Using the fact that $R$ can be consider as a $3 \times 3$ orthogonal matrix, we can say a few things about the eigenvalues. First, $R$ is actually a *unitary* matrix (a unitary matrix which happens to have only real entries). We know that all eigenvalues of unitary matrices have modulus 1. So, if $\lambda_i$ ($i = 1, 2, 3$) denotes the eigenvalues of $R$, then we have

$$|\lambda_1| = |\lambda_2| = |\lambda_3| = 1 .$$

The reality of $R$ has also an consequence. It implies that the complex conjugate of any eigenvalue is also an eigenvalue (how can you show this?).

For all $i$, $\lambda_i^* = \lambda_j$ for some $j$.

Finally, since $R$ is proper, we have

$$\lambda_1 \lambda_2 \lambda_3 = \det R = +1 .$$

All of the above together implies that the eigenvalues of $R$ are of the form

$$\lambda_1 = 1 , \quad \lambda_2 = e^{i\theta} , \quad \lambda_3 = e^{-i\theta} ,$$

where $\theta$ is some angle with $0 \leq \theta \leq \pi$. Convince yourself that this is the only possibility. The angle $\theta$ will be called as the “rotation angle”. (Convince yourself that if $\theta = 0$, then $R$ must be the identity.)
The most important conclusion that we can draw from the analysis above is this: every proper rotation has +1 as one of the eigenvalues. Therefore, there is a vector $\hat{n}$ which satisfies

$$R\hat{n} = \hat{n}.$$  

If $R$ is not identity (i.e., $\theta \neq 0$), then the eigenvalue 1 is not repeated and hence there is only one corresponding eigenvector. This shows that $\hat{n}$ can be chosen as a real vector. We will always normalize $\hat{n}$. With this condition, there are only two different ways of defining this vector, i.e., either $\hat{n}$ or $-\hat{n}$ both can serve as a normalized real eigenvector for the eigenvalue +1.

Of course, the real importance of this direction is its being the rotation axis. Note that $R\hat{n} = \hat{n}$ implies that the vector $\hat{n}$ is left invariant by the rotation. Only the vectors parallel to $\hat{n}$ are left invariant by $R$. Note that $-\hat{n}$ as well as $\hat{n}$ might easily serve as a rotation axis. Rotation axis is really an line, not a direction.

Consider now the 2D plane perpendicular to $\hat{n}$. As $R$ preserves angles, all vectors in this plane must be transformed into other vectors in the same plane. As a result, we can think of $R$ as a rotation on this 2D plane. As such rotations can be described by a single rotation angle, we will define the rotation angle of $R$ to be that angle. This is the same as the angle $\theta$ that we have met in the eigenvalues above. However, in this case, we will let $\theta$ to be positive or negative.

We will use the right hand rule in defining the angle: If your thumb points along $\hat{n}$, the other four fingers show the direction of rotation. If the rotation angle $\theta$ is positive, then rotation is carried out along the four fingers. If $\theta$ is negative, rotation is in opposite direction. We can completely specify a rotation matrix by specifying the axis $\hat{n}$ and the angle $\theta$. I am going to write this relation as $R = R(\theta, \hat{n})$.

A few points to note down: Rotations around a fixed $\hat{n}$ by angles $\theta$ and $\theta + 2\pi$ are identical. Also, rotation around $\hat{n}$ by $\theta$ is the same thing as a rotation around $-\hat{n}$ by angle $-\theta$. To remove the last arbitrariness, we will use one single vector quantity

$$\vec{\theta} = \theta \hat{n}$$

to describe any rotation and write $R = R(\vec{\theta})$ for the rotation matrix. In summary, $\vec{\theta}$ is a vector, whose length gives the
rotation angle, whose direction specifies the rotation axis. Note that \( \bar{\theta} = 0 \) corresponds to no rotation (\( R \) is identity matrix, \( R = I \)). Therefore, small rotations correspond to values of vector \( \bar{\theta} \) close to 0. The \( 2\pi \) arbitrariness for the rotation angle still remains, so if the length of \( \bar{\theta} \) is \( 2\pi, 4\pi, \) etc., it represents no rotation again.

The rotation \( R(\hat{z}, +\pi/6) \) which is a 30° rotation in positive direction around the +z axis. The blue lines show the rotated x and y axes. The rotation \( R(\hat{z}, -\pi/6) \) which is a 30° rotation in negative direction around the +z axis. This can also be considered as a 30° rotation in positive direction around the −z axis.

Since every rotation \( R \) can be identified by a real vector \( \bar{\theta} \), the rotations depend on 3 real parameters, i.e., the three components of \( \bar{\theta} \). This means that the group SO(3) is a 3-dimensional manifold. There are other possible parameterizations as well. For example, the three Euler angles are frequently used to represent notations. But, for our analysis, \( \bar{\theta} \) is the best parameterization.

**Action of Rotations on Vectors**

We do not need to calculate the matrix elements of the rotation matrix \( R(\bar{\theta}) \) as a function of \( \bar{\theta} \), as this is not needed. However, we need a simple expression for rotations to be able to deal with certain relations that we will meet later. Consider a rotation \( R = R(\bar{\theta}) \). We will find a simple way for expressing how \( R \) acts on vectors. To be able to do this, we start with a reference frame.
Let \( \hat{e}_1 \) and \( \hat{e}_2 \) be two perpendicular unit vectors in the plane perpendicular to the rotation axis. We will choose these two vectors in such a way that \( \hat{e}_1, \hat{e}_2, \hat{n} \) forms a right-handed frame (i.e., \( \hat{e}_1 \times \hat{e}_2 = \hat{n}, \hat{e}_2 \times \hat{n} = \hat{e}_1 \), etc.). The action of the rotation on these vectors can be expressed as

\[
R\hat{n} = \hat{n}, \\
R\hat{e}_1 = \hat{e}_1 \cos \theta + \hat{e}_2 \sin \theta, \\
R\hat{e}_2 = -\hat{e}_1 \sin \theta + \hat{e}_2 \cos \theta.
\]

At this point, it is instructive to check that the vectors

\[
\hat{e}_\pm = \frac{\hat{e}_1 \pm \hat{e}_2}{\sqrt{2}}
\]

are the eigenvectors associated with the complex eigenvalues of \( R \).

Now, consider an arbitrary vector, \( \vec{V} \). We would like to find an expression for \( \vec{V}' = R\vec{V} \). For this purpose, let us think of \( \vec{V} \) as a sum of a parallel component \( \vec{V}_\parallel \), which is parallel to the rotation axis, and a perpendicular component, \( \vec{V}_\perp \), which lies in the perpendicular plane. It is straightforward to see that

\[
\vec{V}_\parallel = \hat{n}(\hat{n} \cdot \vec{V}), \\
\vec{V}_\perp = \vec{V} - \hat{n}(\hat{n} \cdot \vec{V}), \\
\vec{V} = \vec{V}_\parallel + \vec{V}_\perp.
\]

Under a rotation, \( \vec{V}_\parallel \) remains the same while \( \vec{V}_\perp \) changes to some other vector in the perpendicular plane.

It is also straightforward to see that

\[
\hat{n} \times \vec{V} = \hat{n} \times \vec{V}_\perp
\]
is another vector in the plane perpendicular to the axis. In fact, the three
vectors, $\mathbf{V}_\perp, \mathbf{n} \times \mathbf{V}, \mathbf{n}$ forms a right-handed frame of mutually perpendicular
vectors. The first two are not unit vectors, but it is easy to see that they
have the same length. Therefore, it is straightforward to express the rotation
in terms of these vectors. The following relation follows naturally.

$$ R\mathbf{V} = \mathbf{V}_\parallel + \mathbf{V}_\perp \cos \theta + \mathbf{n} \times \mathbf{V} \sin \theta. $$

If we express it in terms of the original parameters, we have

$$ R\mathbf{V} = \mathbf{n}(\mathbf{n} \cdot \mathbf{V}) + (\mathbf{V} - \mathbf{n}(\mathbf{n} \cdot \mathbf{V})) \cos \theta + \mathbf{n} \times \mathbf{V} \sin \theta. $$

Luckily, we won’t need to use this complicated expression. To simplify it, we
convert it to a power series expansion in $\mathbf{\bar{\theta}}$. If you do the calculations, you
will see that the final expression is really simple,

$$ R\mathbf{V} = \mathbf{V} + \mathbf{\bar{\theta}} \times \mathbf{V} + \frac{1}{2!} \mathbf{\bar{\theta}} \times (\mathbf{\bar{\theta}} \times \mathbf{V}) + \frac{1}{3!} \mathbf{\bar{\theta}} \times (\mathbf{\bar{\theta}} \times (\mathbf{\bar{\theta}} \times \mathbf{V})) + \cdots. $$

Usefulness of the definition of the vector $\mathbf{\bar{\theta}}$ is apparent in this equation. If the
whole series is summed up, it will give the exact result for any finite value of $\mathbf{\bar{\theta}}$.

However, in most of our applications, we only need the expression for
small rotations where $|\mathbf{\bar{\theta}}| \ll 2\pi$. In that case, the first two terms of the
expansion above is enough for our purposes.

$$ R\mathbf{V} \approx \mathbf{V} + \mathbf{\bar{\theta}} \times \mathbf{V} \quad \text{(small angle approximation).} \quad (1) $$

**Properties of SO(3)**

Let us consider the composition of two rotations. We write $R = R_2R_1$ for a
rotation where $R_1$ is applied first and $R_2$ is applied next. The matrix for the
whole rotation, $R$, can be easily calculated by a simple matrix multiplication.
Unfortunately, the relationship between the rotation axes and angles are
complicated. We first cover the simplest cases. First, if the rotation axes are
the same, then the rotation angles are simply added up.

$$ R(\theta_2\mathbf{n})R(\theta_1\mathbf{n}) = R(\theta_2\mathbf{n} + \theta_1\mathbf{n}). $$
In this case, you can see that the rotation matrices also commute; it doesn’t matter which one is applied first. As a special application of this, we can show that any proper rotation can be obtained by a series of small rotations,

\[ R(\vec{\theta}) = R\left(\frac{\vec{\theta}}{2}\right)^2 = \cdots = R\left(\frac{\vec{\theta}}{N}\right)^N = \cdots \]

Also, using this rule you can show that the inverse of a \( \vec{\theta} \) rotation corresponds to a rotation by \( -\vec{\theta} \), i.e., \( R(\vec{\theta})^{-1} = R(-\vec{\theta}) \). A second simple case is for small rotations. By using equation (1), you can show that

\[ R(\vec{\theta}_2)R(\vec{\theta}_1) \approx R(\vec{\theta}_1 + \vec{\theta}_2) \quad \text{(if } \vec{\theta}_1 \text{ and } \vec{\theta}_2 \text{ are small.)} \]

This equation is valid only up to first order in angles. If you also include the second order terms, you will discover that the left and right hand sides of this equation are different. You will find more on this below.

**Noncommutativity:** An important feature of the rotation group is its noncommutativity. This is best seen by looking at finite rotations. As an example consider two rotations. Let \( R_1 = R(\vec{\theta}_1) \) and \( R_2 \) be two rotation matrices. Consider the action of \( R_2R_1 \) on a vector \( \vec{V} \),

\[
R_2R_1 \vec{V} = R_2(R_1 \vec{V}) = R_2(\vec{V} + \vec{\theta}_1 \times \vec{V} + \frac{1}{2!} \vec{\theta}_1 \times (\vec{\theta}_1 \times \vec{V}) + \cdots)
\]

\[
= R_2\vec{V} + (R_2 \vec{\theta}_1) \times (R_2 \vec{V}) + \frac{1}{2!} (R_2 \vec{\theta}_1) \times ((R_2 \vec{\theta}_1) \times (R_2 \vec{V})) + \cdots
\]

\[
= R(\vec{R}_2 \vec{\theta}_1)(R_2 \vec{V})
\]

The trick we have used above is simply the fact that the cross product of two vectors are rotated to the cross product of rotated vectors. Note that there is no approximation involved in here. This shows that \( R_2R_1 = R_2R(\vec{\theta}_1) = R(\vec{R}_2 \vec{\theta}_1)R_2 \). There is another way of expressing this

\[
R_2R(\vec{\theta}_1)R_2^{-1} = R(\vec{R}_2 \vec{\theta}_1)\]

which can be interpreted in many ways. For example, similarity transform of a rotation is another rotation with the same angle but with a different axis, etc.
It can be seen that in the general case we have $R_2 R_1 \neq R_1 R_2$ since $\vec{\theta}_1 \neq R_2 \vec{\theta}_1$, in general. Basically, if the rotation axis of $R_2$ is different than that of $R_1$, then these two rotations do not commute.

**Problem:** There is one exception to this rule. Suppose that $R_1 = R(\vec{\theta}_1)$ and $R_2 = R(\vec{\theta}_2)$ are two non-trivial rotations (i.e., $R_i \neq I$) with axes in different directions (i.e., $\vec{\theta}_1$ is not parallel to $\vec{\theta}_2$). If these rotations commute, $R_1 R_2 = R_2 R_1$, then both of these are $180^\circ$ rotations ($|\theta_1| \equiv |\theta_2| \equiv \pi$) and the two axes are perpendicular ($\vec{\theta}_1 \perp \vec{\theta}_2$).

To see the degree of noncommutativity we investigate the following expression for two rotation matrices $R_1$ and $R_2$

$$S = R_2 R_1 R_2^{-1} R_1^{-1}$$

If $R_1$ and $R_2$ commute, the matrix $S$ will be equal to identity matrix. If they don’t, then $S$ will be different from identity. We are going to assume that $R_1$ and $R_2$ are small rotations (as this is all we need). Let $R_1 = R(\vec{\theta}_1)$ and $R_2 = R(\vec{\theta}_2)$ where $|\vec{\theta}_1|, |\vec{\theta}_2| \ll 2\pi$. First we note that $R_2 R_1 R_2^{-1} = R(R_2 \vec{\theta}_1)$.

Also, note that $R_1^{-1} = R(-\vec{\theta}_1)$. As a result,

$$S = R(R_2 \vec{\theta}_1) R(-\vec{\theta}_1) \approx R(R_2 \vec{\theta}_1 - \vec{\theta}_1)$$

Using equation (1) we can write $R_2 \vec{\theta}_1 \approx \vec{\theta}_1 + \vec{\theta}_2 \times \vec{\theta}_1$. All of these lead to

$$S = R_2 R_1 R_2^{-1} R_1^{-1} \approx R(\vec{\theta}_2 \times \vec{\theta}_1) \quad (2)$$

The equation (2) is an important equation for us since it will lead directly to the angular momentum commutation relations. By calculating the omitted terms correctly, you can check that it is correct up to second order terms. It shows the degree of noncommutativity for small rotations. Basically, if you change the order of application of two rotations, you will be off by a rotation about $\vec{\theta}_2 \times \vec{\theta}_1$,

$$R_2 R_1 \approx R(\vec{\theta}_2 \times \vec{\theta}_1) R_1 R_2$$

Again, it is important to remember that the above is valid only for small rotations.

**Rotations of states**

Now we will see how we rotate the quantum mechanical states. We are going to consider an arbitrary physical system which can be anything. It can be
composed of several particles. Each particle may have an internal structure and spin. Particles might be created or annihilated. We are not going to specify anything about the nature of the system, so that the results we will obtain are as general as possible.

We know that the states of any such system are represented as normalized kets in a certain Hilbert space. Let $|\psi\rangle$ be any such state. Let us rotate the system in this state by a certain rotation with a rotation matrix, $R$. The system will now be in a different state. The new state is also represented by a normalized ket in the Hilbert space. Let us call the new state as $|\psi'\rangle$. We postulate the existence of a linear operator $\mathcal{D}(R)$ that relates the new, rotated state $|\psi'\rangle$ to the old state $|\psi\rangle$.

$$|\psi\rangle \rightarrow |\psi'\rangle = \mathcal{D}(R) |\psi\rangle .$$

Basically, $R$ takes care of the rotations in the 3 dimensional vectors and $\mathcal{D}(R)$ does the same for states. Since we require $\mathcal{D}(R)$ to map normalized kets into normalized kets (i.e., if $\langle \psi | \psi \rangle = 1$ then $\langle \psi' | \psi' \rangle = 1$), this operator has to be unitary.

$$\mathcal{D}(R)\mathcal{D}(R)^\dagger = \mathcal{D}(R)^\dagger\mathcal{D}(R) = 1 .$$

Before we count down the properties of $\mathcal{D}(R)$, let us state one minimal property that it should satisfy. Let $\mathbf{V}$ be a vector observable for the system, i.e., a set of three hermitian operators $V_x$, $V_y$, and $V_z$ which represent the components of a vector observable. The vector in question may be position (operators $x$, $y$ and $z$), momentum (operators $p_x$, $p_y$, and $p_z$), internal magnetic dipole moment of a certain particle in the system or any other vector.
In all cases, there are three different operators that represent these three components. We know that $\langle V_x \rangle$ is the average value of the $x$-component of the vector in question in the state used. This value might change as the state is changed, but the operator $V_x$ remains the same (this is what we expect from the operators). As a result, the expectation values of the three components of the vector operator in the old and the rotated state are related by the rotation matrix

$$\langle \psi' | V_k | \psi' \rangle = \sum_\ell R_{k\ell} \langle \psi | V_\ell | \psi \rangle .$$

In a compact form we can write this as, $\langle \vec{V} \rangle' = R \langle \vec{V} \rangle$. Since this relation has to be valid for any $|\psi\rangle$ we should have

$$\mathcal{D}(R)^\dagger V_k \mathcal{D}(R) = \sum_\ell R_{k\ell} V_\ell \quad \text{or} \quad \mathcal{D}(R)^\dagger \vec{V} \mathcal{D}(R) = R \vec{V} \quad (3)$$

Properties of $\mathcal{D}(R)$:

- We will postulate that $\mathcal{D}(R)$ satisfies the same composition rule as the rotation matrices:

$$\mathcal{D}(R_2) \mathcal{D}(R_1) = \mathcal{D}(R_2 R_1) . \quad (4)$$

In mathematical terminology, we say that $\mathcal{D}$ is a representation of the rotation group SO(3).

Unfortunately, there are cases where this relation cannot be satisfied as above. As overall phase factors cannot change the physical state, at
best we can say in such cases is that composition rule is satisfied with extra phase factors, \( D(R_2)D(R_1) = e^{i\alpha}D(R_2R_1) \), where \( \alpha \) depends on \( R_i \). However, even in such cases, we will require \( D \) to satisfy Eq. (4) (i.e., \( \alpha = 0 \)) for small rotations. The behavior of \( D \) around small rotations is enough for us to determine its global properties.

- It follows then that for no rotation, the operator is identity: \( D(I) = 1 \).
- Since \( RR^{-1} = I \), we have \( D(R^{-1}) = D(R)^{-1} = D(R)^\dagger \).
- For \( R = \mathcal{R}(\vec{\theta}) \) we will write \( D(\vec{\theta}) \) for \( D(R) \) to simplify the notation. We can see that \( D(\vec{\theta})^\dagger = D(\vec{\theta})^{-1} = D(-\vec{\theta}) \).
- Similarly, we have \( D(\vec{\theta}) = D(\vec{\theta}/N)^N \).
- Now let us consider the case of small rotations. We will define an operator \( J_n \) which we call the generator of rotations around an axis \( \hat{n} \) by

\[
D(\theta\hat{n}) \approx 1 - \frac{i}{\hbar}\theta J_n \quad (\text{for } \theta \ll 2\pi)
\]

We will also call \( J_n \) as the component of angular momentum along unit vector \( \hat{n} \).

- Using \( D(\theta\hat{n})^\dagger = D(-\theta\hat{n}) \) we can see that \( J_n \) is hermitian, \( J_n^\dagger = J_n \).
- There are actually only three independent generators since for small rotations we can write (using \( \vec{\theta}\hat{n} = \theta n_x\hat{x} + \theta n_y\hat{y} + \theta n_z\hat{z} \))

\[
\mathcal{R}(\vec{\theta}) \approx \mathcal{R}(\theta n_x\hat{x})\mathcal{R}(\theta n_y\hat{y})\mathcal{R}(\theta n_z\hat{z})
\]

\[
\rightarrow D(\vec{\theta}) \approx D(\theta n_x\hat{x})D(\theta n_y\hat{y})D(\theta n_z\hat{z})
\]

\[
\left(1 - \frac{i}{\hbar}\theta J_n\right) \approx \left(1 - \frac{i}{\hbar}\theta n_x J_x\right)\left(1 - \frac{i}{\hbar}\theta n_y J_y\right)\left(1 - \frac{i}{\hbar}\theta n_z J_z\right)
\]

\[
\rightarrow J_n = n_x J_x + n_y J_y + n_z J_z = \hat{n} \cdot \vec{J}.
\]

The fact that there are three independent generators follows from the fact that the vector variable \( \vec{\theta} \) that represents the rotations have three independent components. Here, we also see that these three generators can be considered as the components of a vector, so that the generator along \( \hat{n} \) is given as the dot product of \( \hat{n} \) and \( \vec{J} \). We will show below that \( \vec{J} \) is a vector by the standards of equation (3) as well.
• For a finite rotation $\vec{\theta} = \theta \hat{n}$, we have

$$ R(\vec{\theta}) = R\left(\frac{\vec{\theta}}{N}\right)^N $$

$$ \rightarrow D(\vec{\theta}) = D\left(\frac{\vec{\theta}}{N}\right)^N = \lim_{N \to \infty} \left(1 - \frac{i}{\hbar N} J_n\right)^N $$

$$ = \exp\left(-\frac{i}{\hbar} \theta J_n\right) $$

Therefore we have an exact expression for any finite rotation

$$ D(\vec{\theta}) = \exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{J}\right) $$

Note the similarity of this expression to the expressions of the translation and time development operators, which are defined in a similar way.

• Next, we will restate Eq. (3), the transformation rule for vectors, in terms of the generators. We consider small rotations again. The following equality is exact up to the first power of $\vec{\theta}$.

$$ D(\vec{\theta}) \hat{\vec{V}} D(\vec{\theta})^\dagger \approx \hat{\vec{V}} + \vec{\theta} \times \hat{\vec{V}} $$

$$ \exp\left(\frac{i}{\hbar} \vec{\theta} \cdot \vec{J}\right) \hat{\vec{V}} \exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{J}\right) \approx $$

$$ \hat{\vec{V}} + \frac{i}{\hbar} [\vec{\theta} \cdot \vec{J}, \hat{\vec{V}}] + \cdots \approx $$

$$ \rightarrow [\vec{\theta} \cdot \vec{J}, \hat{\vec{V}}] = -\hbar \vec{\theta} \times \hat{\vec{V}} $$

We will use the notation $V_m = \hat{m} \cdot \vec{V}$ to represent the component of the vector $\vec{V}$ along unit vector $\hat{m}$. Taking the dot product of the equation above with $\hat{m}$ and using $\vec{\theta} = \theta \hat{n}$ we get $[J_n, V_m] = -\hbar \hat{m} \cdot (\hat{n} \times \vec{V}) = -\hbar \vec{V} \cdot (\hat{m} \times \hat{n}) = +\hbar \vec{V} \cdot (\hat{n} \times \hat{m})$. Note that three distinct components are involved in this commutation relation. These are the alternative
forms of writing the same equations

\[
\begin{align*}
[J_n, V_m] &= i\hbar (\hat{n} \times \hat{m}) \cdot \vec{V}, \\
[J_i, V_j] &= i\hbar \sum_k \epsilon_{ijk} V_k, \\
[V_i, J_j] &= i\hbar \sum_k \epsilon_{ijk} V_k.
\end{align*}
\]

These equations tell us that the operator \( \vec{V} \) is a vector, \textit{it transforms like a vector under rotations}. It is completely equivalent to Eq. (3). In fact, you can easily obtain Eq. (3) starting from the commutation relations ab

- There are also observables that does not change after rotations. Simplest example is energy if it is a rotationally invariant system. We say that an operator \( A \) transforms like a scalar or \( A \) is rotationally invariant if its expectation value does not change under rotations, i.e., \( \langle A \rangle' = \langle A \rangle \) where prime denotes the rotated state. This is equivalent to \( \mathcal{D}(R)^\dagger A \mathcal{D}(R) = A \) which implies that \([ A, \mathcal{D}(R) ] = 0 \) for any rotation \( R \). Expressing this in terms of generators we have

\[
[J_i, A] = 0 \quad \text{if } A \text{ transforms like a scalar.} \quad (6)
\]

The converse is also true: if an operator \( A \) commutes with all components of \( \vec{J} \), i.e., if Eq. (6) is satisfied, then you can easily show that \( A \) transforms like a scalar.

- The dot product of two vectors is an obvious candidate of a quantity that transform like a scalar. You can show that if \( \vec{V} \) and \( \vec{W} \) are two vectors (i.e., satisfying Eq. (5)), then

\[
[J_i, \vec{V} \cdot \vec{W}] = 0.
\]

(Note that since \( \vec{V} \) and \( \vec{W} \) are operators, \( \vec{V} \cdot \vec{W} \) is not the same thing as \( \vec{W} \cdot \vec{V} \). This is an important relationship that we will use a lot. An important subcase is the squared length of a vector: \( V^2 = \vec{V} \cdot \vec{V} = V_x^2 + V_y^2 + V_z^2 \). We again have \([ J_i, V^2 ] = 0 \). From this we can easily see that \( r^2 = x^2 + y^2 + z^2 = \vec{r} \cdot \vec{r} \) and \( p^2 = \vec{p} \cdot \vec{p} = p_x^2 + p_y^2 + p_z^2 \) commute with all components of \( \vec{J} \). It also follows that any function of these, such as
\[ r = \sqrt{r^2}, \text{ commute with } \vec{J}. \] As a simple application, the Hamiltonian of a particle under a central potential

\[ H = \frac{\dot{p}^2}{2m} + V(r) \]

is rotationally invariant.

- For two vector operators \( \vec{V} \) and \( \vec{W} \), the cross product \( \vec{U} = \vec{V} \times \vec{W} \) is a vector operator by the meaning of Eq. (5). Showing this takes a little bit more time than the previous item, but it is straightforward. As an example, the operator \( \vec{L} = \vec{r} \times \vec{p} \), transforms like a vector.

- We see that the commutation relations, Eqs. (5) and (6), are two central relations that contain deep geometrical meaning. It tells us about the behavior of certain observables under rotations. Similar relations can be obtained for tensors with higher rank (such as quadrupole moments with rank 2) but we are not going to go into these.

- Finally, we use equation (2) to extract the commutation relations between components of the angular momentum.

\[
\begin{align*}
\mathcal{D}(\vec{\theta}_2)\mathcal{D}(\vec{\theta}_1)\mathcal{D}(-\vec{\theta}_2)\mathcal{D}(-\vec{\theta}_1) & \approx \mathcal{D}(\vec{\theta}_2 \times \vec{\theta}_1) \\
\exp\left(-\frac{i}{\hbar} \vec{\theta}_2 \cdot \vec{J}\right)\mathcal{D}(\vec{\theta}_1)\exp\left(+\frac{i}{\hbar} \vec{\theta}_2 \cdot \vec{J}\right)\mathcal{D}(-\vec{\theta}_1) & \approx \mathcal{D}(\vec{\theta}_2 \times \vec{\theta}_1) \\
\left(\mathcal{D}(\vec{\theta}_1) - \frac{i}{\hbar} \left[ \vec{\theta}_2 \cdot \vec{J}, \mathcal{D}(\vec{\theta}_1) \right] \right)\mathcal{D}(-\vec{\theta}_1) & \approx \mathcal{D}(\vec{\theta}_2 \times \vec{\theta}_1) \\
1 + \left(\frac{-i}{\hbar}\right)^2 \left[ \vec{\theta}_2 \cdot \vec{J}, \vec{\theta}_1 \cdot \vec{J} \right] & \approx 1 - \frac{i}{\hbar} \vec{\theta}_2 \times \vec{\theta}_1 \cdot \vec{J} \\
\left[ \vec{\theta}_2 \cdot \vec{J}, \vec{\theta}_1 \cdot \vec{J} \right] & = i\hbar \vec{\theta}_2 \times \vec{\theta}_1 \cdot \vec{J}
\end{align*}
\]

where in the derivation, only the second or lower order terms are kept; all higher order terms are dropped. Using various values for \( \vec{\theta}_2 \) and \( \vec{\theta}_1 \), we can obtain all commutation relations. First we rewrite the last equation in unit vectors

\[ [\hat{n} \cdot \vec{J}, \hat{m} \cdot \vec{J}] = i\hbar (\hat{n} \times \hat{m}) \cdot \vec{J}. \]

Basically, the commutator of two components of the angular momentum gives the third component.
• Same relations are often written as

\[ [J_i, J_j] = i\hbar \sum_s \epsilon_{ijk} J_k \]  \hspace{1cm} (7)

Note that these relations are of the form of Eq. (5). Therefore, they tell us that \( \vec{J} \) is a vector! It transforms like vectors under rotations.

• Since we now know that \( \vec{J} \) is a vector, we can immediately write the following

\[ [J_i, J^2] = 0 \]  

Examples

Up to now, we have said nothing about the physical system in question. We know that, whatever the system is, there is an operator \( \vec{J} \) that describes the behavior of this system under rotations. The only thing we know about them is the commutation relations expressed in the equations (7), (5) and (6). However, we don’t yet know anything about their mathematical composition as operators. In here we will give a few examples that will show what \( \vec{J} \) is.

First consider a system composed of a single spinless particle. We can describe the state of this particle with a position-space wavefunction, \( \psi(\vec{r}') \). Let \( \psi \) be one particular state of this particle. Let us rotate this particle by a rotation \( R \). Its state will now be different which we call \( \psi' \). What can we say about the relationship between \( \psi' \) and \( \psi \)? The correct way of thinking is like this: Let \( \vec{a} \) be a special position where \( \psi \) has a special behavior. For example, probability density for \( \psi \) might be a maximum at \( \vec{a} \), or the wavefunction might be zero there. We can easily see that the rotated wavefunction \( \psi' \) has the same feature at the rotated point, namely \( R\vec{a} \). This rule can be easily satisfied by taking the value of the wavefunctions at these two points to be the same, i.e.,

\[ \psi'(R\vec{a}) = \psi(\vec{a}) \]  

Using \( \vec{r}' = R\vec{a} \) we get the transformation rule

\[ \psi'(\vec{r}') = \psi(R^{-1}\vec{r}') \]  

At first sight, it seems like this rule is the opposite of what we expect. However, it is the correct one satisfying the condition that we stated above,
namely the features of the wavefunctions are rotated. You can also show that it correctly satisfies the composition rule. (Consider two rotations, \( R_1 \) is applied first, \( R_2 \) is applied next. The whole rotation has to be \( R_2 R_1 \).) Let

\[
\psi \xrightarrow{R_1} \psi' \xrightarrow{R_2} \psi''.
\]

The relationship between the functions are given as follows

\[
\psi''(\vec{r}') = \psi'(R_2^{-1}\vec{r}') \\
= \psi(R_1^{-1}(R_2^{-1}\vec{r}')) \\
= \psi((R_2 R_1)^{-1}\vec{r}'),
\]

which shows what we mean.

Now, we consider small rotations by parameter \( \vec{\theta} \) where \( |\vec{\theta}| \ll 2\pi \). Then

\[
\psi'(\vec{r}') = \langle \vec{r}'|\psi' \rangle \\
= \langle \vec{r}'|D(\vec{\theta})|\psi \rangle \\
\approx \langle \vec{r}'|\psi \rangle - \frac{i}{\hbar} \langle \vec{r}'|\vec{\theta} \cdot \vec{J}|\psi \rangle \\
= \psi(\vec{r}') - \frac{i}{\hbar} \vec{\theta} \cdot (\vec{J} \psi)(\vec{r}')
\]

On the other hand,

\[
\psi'(\vec{r}') = \psi(R^{-1}\vec{r}') \\
\approx \psi(\vec{r}' - \vec{\theta} \times \vec{r}') \\
\approx \psi(\vec{r}') - (\vec{\theta} \times \vec{r}') \cdot \vec{\nabla}' \psi(\vec{r}') .
\]

This shows that

\[
\vec{\theta} \cdot (\vec{J} \psi)(\vec{r}') = -i\hbar(\vec{\theta} \times \vec{r}') \cdot \vec{\nabla}' \psi(\vec{r}') \\
= -i\hbar \vec{\theta} \cdot (\vec{r}' \times \vec{\nabla}' \psi(\vec{r}')) \\
= \vec{\theta} \cdot (\vec{r} \times \vec{p} \psi)(\vec{r}')
\]

We can see that, therefore \( \vec{J} = \vec{r} \times \vec{p} \) for this particle. We will call \( \vec{L} = \vec{J} = \vec{r} \times \vec{p} \) the orbital angular momentum of the particle in question. We see that
\( \vec{L} \) generates rotations of the positional degree of freedom of the particle. The following commutation relations are naturally satisfied

\[
[L_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k ,
\]

\[
[L_i, x_j] = i\hbar \sum_k \epsilon_{ijk} x_k ,
\]

\[
[L_i, p_j] = i\hbar \sum_k \epsilon_{ijk} p_k .
\]

It is a useful exercise to derive all of these commutation relations starting from the canonical commutation relations \([x_i, p_j] = i\hbar \delta_{ij}\) only and not using the rotation related ideas we have presented in here.

Next, consider \( N \) spinless particles. The state of the system is represented by a wavefunction that depends on \( N \) positions, \( \psi = \psi(\vec{r}_1', \vec{r}_2', ..., \vec{r}_N') \). If under a rotation \( R \), the wavefunction changes to \( \psi' \), then we naturally expect that

\[
\psi'(\vec{r}_1', \vec{r}_2', ..., \vec{r}_N') = \psi(R^{-1}\vec{r}_1', R^{-1}\vec{r}_2', ..., R^{-1}\vec{r}_N') ,
\]

i.e., positions of all particles are rotated independently. Using the same derivation, we get the generator as

\[
\vec{J} = \vec{r}_1 \times \vec{p}_1 + \vec{r}_2 \times \vec{p}_2 + \cdots + \vec{r}_N \times \vec{p}_N = \vec{L}_1 + \vec{L}_2 + \cdots + \vec{L}_N .
\]

Note that, \( \vec{L}_1 \) generates rotations of first particle only. It does not have any effect on other particles. To rotate the whole system, we need the generator of each and every particle. So, the real generator, \( \vec{J} \), is the sum of the orbital angular momenta of each particle. From mechanics, we used to call \( \vec{J} \) as the total angular momentum of the system and we will keep using this name. We will state a general rule in here: If a system is composed of several subsystems, then the angular momentum of the whole is equal to the sum of angular momenta of the subsystems.

Next we consider a single particle with an “internal degree of freedom”. Such a particle has a positional state and an independent “internal state”. An easier to understand example would be Hydrogen atom considered as a single particle. This atom can be at different positions due to its positional degree of freedom. But it can also have internal excitations like 1s or 2s states. The following two states are different from each other: “the atom is at \( \vec{r}' \) and in 1s level” and “the atom is at \( \vec{r}' \) and in 2s level”. Therefore, to represent
the complete state of such a particle we have to represent both its positional  
behavior and its internal behavior. Particles with spin like electrons have  
such internal degrees of freedom. For example the electron has an internal  
magnetic dipole moment, a vector property, which exist independently from  
the positional degree of freedom. To do a complete rotation of an electron  
we have to rotate it both internally and in positional space.

I will assume that the particle in question has two internal levels that I will  
show by $|\uparrow\rangle$ and $|\downarrow\rangle$ (no meaning is attached to these symbols at this point).  
There are then two position eigenkets located at the same position, $|\vec{r}',\uparrow\rangle$  
and $|\vec{r}',\downarrow\rangle$ (in other words, the position eigenvalues are 2-fold degenerate).  
These states for all possible values of $\vec{r}'$ form an orthonormal basis for the  
Hilbert space. Therefore we can express any state $|\psi\rangle$ as a superposition of  
these,

$$|\psi\rangle = \int d^3\vec{r}' \left( f_1(\vec{r}') |\vec{r}',\uparrow\rangle + f_2(\vec{r}') |\vec{r}',\downarrow\rangle \right),$$

which means that we have to use two wavefunctions, $f_1(\vec{r}')$ and $f_2(\vec{r}')$ to  
describe such a particle quantum mechanically. We then have a matrix-  
wavefunction representation of the states of this particle,

$$\psi = \begin{bmatrix} f_1(\vec{r}') \\ f_2(\vec{r}') \end{bmatrix}.$$

Such an expression is called a two-component spinor. We have never met such  
kinds of expressions before because we have always considered the positional  
and spin degrees of freedom of electrons to be independent from each other.  
However, in reality we have to use such a spinor to describe the electrons.

Now, we do a rotation $R$. In that case, the wavefunctions are changed  
into $f'_1$ and $f'_2$ and we would like to express them in terms of the original  
wavefunctions $f_1$ and $f_2$. In general such a relationship looks like

$$\psi \rightarrow \psi' = \begin{bmatrix} f'_1(\vec{r}') \\ f'_2(\vec{r}') \end{bmatrix} = \begin{bmatrix} a f_1(R^{-1}\vec{r}') + b f_2(R^{-1}\vec{r}') \\ c f_1(R^{-1}\vec{r}') + d f_2(R^{-1}\vec{r}') \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} f_1(R^{-1}\vec{r}') \\ f_2(R^{-1}\vec{r}') \end{bmatrix}$$

The matrix in front does the rotation of internal state, while the argument  
$R^{-1}\vec{r}'$ takes care of the rotation in positional space. Here, the internal rotation  
is represented by a 2x2 matrix

$$D_{\text{spin}}(R) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$
The generator of these internal rotations will be called \textit{spin angular momentum}, $\mathbf{S}$. We won’t say anything about the matrix elements of the rotation operator and spin at this stage; all will be derived later. The whole rotation operator is

$$D(R) = D_{\text{spin}}(R) D_{\text{space}}(R) = D_{\text{space}}(R) D_{\text{spin}}(R),$$

which means that the internal and positional rotations are carried out one after another. The generator of rotations is then

$$\mathbf{J} = \mathbf{S} + \mathbf{r} \times \mathbf{p} = \mathbf{S} + \mathbf{L}.$$  

As a result, we can consider the internal and positional degree of freedoms as two subsystems of a whole. Each subsystem has its own rotations and its own generators. The total angular momentum is then the sum of individual angular momenta. From these we can see that spin and orbital angular momentum commute with each other, $[S_i, L_j] = 0$, while each of these angular momenta satisfy the basic commutation relations,

$$[S_i, S_j] = i\hbar \sum_k \epsilon_{ijk} S_k \quad \text{and} \quad [L_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k.$$

It is straightforward to extend all of these to $N$ particles with spin. The angular momentum for the whole system is then the sum of all orbital and spin angular momenta,

$$\mathbf{J} = \mathbf{S}_1 + \cdots + \mathbf{S}_N + \mathbf{L}_1 + \cdots + \mathbf{L}_N.$$

For rotational symmetry, the total $\mathbf{J}$ is the relevant operator, as the dynamics will be invariant under rotations if the \textit{whole} system is rotated.