Lectures on Elementary Quantum Mechanics

G P Sastry
Department of Physics and Meteorology
Indian Institute of Technology
Kharagpur

January 9, 2026

Abstract

Autumn Semester 1995. Notes by Indrajit Mitra. Edited, revised, and typeset in LaTeX 2_ε by Aniket Basu[1].

[1]To report errata, please write to aniket.basu@gmail.com

Contents

1 The Uncertainty Principle

Question: Why does a 1s-electron in a hydrogen atom not fall into the nucleus, even though its total angular momentum vanishes? [2]

Answer: Let us estimate the kinetic energy of an electron confined to a length scale of the order of Δx:

$$\frac{(\Delta p)^2}{2m} \sim \left(\frac{\hbar}{2\Delta x}\right)^2 /2m \tag{1}$$

Note: This expression is non-relativistic. For nuclear dimensions, $\Delta x \sim 1$ fermi $\sim 10^{-15}$m. For this value of Δx, the average kinetic energy is of the order of 100 MeV, at which scale the effects of relativity become important and the formula no longer applies. With this energy, an electron will simply fly off – it can never fall into the nucleus. (*Aside:* The order of energy of electrons in β-decay is, however, a few KeV, so electrons are *not present* in the nucleus as such. For α-particles, however, average kinetic energy is of the order of 10 KeV, and this is what is found experimentally.) In the first Bohr orbit, the average kinetic energy is of the order of 1 eV, comparable to the binding energy of the hydrogen atom, as we should have expected. For molecules, typical binding energies will be of the order of 10^{-3} eV by the same estimates.

The Uncertainty Principle can be shown to be due to the form of the Fourier transform, and wave-particle duality. Consider a pulse

$$f(x) = \begin{cases} A & |x| < a \\ 0 & \text{otherwise} \end{cases}$$

where A, a are constants.

Insert a diagram of this pulse here.

Here $\Delta x \sim a$.

Then the Fourier transform is

$$\begin{aligned} g(k) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx \\ &= \frac{1}{2\pi}\int_{-a}^{a} Ae^{-ikx}dx \\ &= \frac{A\sin ka}{k\pi} \end{aligned} \tag{2}$$

Observe that in the numerator we have a sine, while in the denominator we have a monotonically increasing function, producing oscillations that die out as we go further from the origin. *Insert a diagram here.*

In k we have a definite spread, because of a superposition of closely spaced frequencies, thus we have an uncertainty in frequency.

The first zero occurs when $ka = \pi$, so

$$\Delta k = \frac{\pi}{a}$$

[2] This question was originally put to GPS by Rupak Mohapatra.

so from $\hbar k = p$, the de Broglie relation, we have

$$\Delta p \sim \frac{\hbar \pi}{a}$$

We started off with the assumption that

$$\Delta x \sim a$$

so

$$\Delta x \Delta p \sim h$$

If we identify the pulse as a photon state, we have the uncertainty principle in this relation.

Similarly, considering a function $f(t)$ and its Fourier transform $F(\omega)$, we can have $\Delta\omega\Delta t \sim \pi$. Identifying this pulse as a photon state, we again have $\Delta E = \hbar\Delta\omega$, so

$$\Delta E \Delta t \sim h$$

If we take any E measurement in a short time interval Δt, the uncertainty ΔE will be large. Any measurement process disturbs the system, and if Δt is small, the system does not have time to settle down, giving a value of E different from the value we actually wanted to measure.

2 Formalism of Quantum Mechanics

We set up a system in a certain state in lab space. Note that the definition of a state in quantum mechanics is different from that in classical mechanics. Whereas in classical mechanics, the positions and momenta of all constituent particles specify the state of the system, in quantum mechanics the uncertainty relation sets certain limits on the accuraccy with which we can know the position and the momentum of a particle at the same time.

So we go from classical lab space to *Hilbert space*: a mathematical, complex, linear vector space.

There is a one-to-one correspondence between the physical space and the "ket" in Hilbert space. The ket specifies the state in Hilbert space and since it behaves as a vector in real space, the ket vector is also known as a state vector, and is represented by

$$|\Psi\rangle$$

However, the one-to-one correspondence of the ket with a state in the lab space has an exception: if this ket

$$|\Psi\rangle$$

specifying a certain state is multiplied by a complex number c, we do not get a new physical state (just as multiplying a vector by a scalar does not change its direction).

Consider the ammonia molecule at low tempterature. All vibrational and rotational degrees of freedom are frozen, save the flip-flop: *insert a diagram here.*

Here we have two possible states and we say that its state vector lies in a two dimensional Hilbert space. Let one state be represented by $|\phi\rangle$ and the other by $|\psi\rangle$. Now, if c_1, c_2 are complex numbers, and operation combining these two kets linearly is permitted:

$$|\eta\rangle \;=\; c_1|\phi\rangle \;+\; c_2|\psi\rangle$$

$|\eta\rangle$ is a new ket in the Hilbert space, and must correspond to a new state of the system. This is absolutely alien to classical mechanics and is the heart of quantum mechanics. This is the *Superposition Principle.*

As the temperature is raised we have the rotational degrees excited and the energy levels split up further. The Hilbert space becomes 6-dimensional and the energy level diagram looks like: *insert diagram here.*

Just as a new state is formed by the superposition of two different states, we can also say that *a state may be resolved into several "base states".* The number of ways this can be done is infinite, and is again analogous to *the resolution of a vector into components along several basis vectors.*

Another operation defined is that of the *adjoint* of a ket, denoted by $\overline{|\phi\rangle}$. This is also denoted as $\langle\phi|$ and known as a "bra"-vector. If there is a complex number multiplied with the ket and we take the adjoint of the whole, we get a bra multiplied by the complex conjugate of the complex number, that is:

$$\overline{|\phi\rangle} = \langle\phi|$$

and

$$\overline{c_1|\phi\rangle} = c_1^*\langle\phi|$$

If $|\eta\rangle = c_1|\phi\rangle + c_2|\psi\rangle$, then $\langle\eta| = c_1^*\langle\phi| + c_2^*\langle\psi|$.

Thus every ket can get associated with a bra and consequently we can associate every state in lab space with a bra vector in a Hilbert space that is different from the Hilbert space of the kets.

Note that addition of a bra and a ket is meaningless (as they belong to different spaces).

Multiplication of two bra vectors or of two ket vectors is also meaningless.

But the product of a bra and a ket is defined and is denoted by $\langle\phi|\psi\rangle$. Physically, this means the probability amplitude of a system in a state ϕ to be found in a state ψ. To obtain the probability we take the the magnitude squared of this complex number.

$\langle\phi|\psi\rangle$ is thus just like the scalar product of two vectors, with *one exception*: it is the complex conjugate of $\langle\psi|\phi\rangle$, which also makes sense, as while calculating the probability this gives the same magnitude.

The scalar product $\langle\psi|\psi\rangle$ is special, as it is always real and non-negative. This is called the *norm* of ψ and is usually set to 1.

The scalar product of two base kets is zero. In this case the two kets concerned are said to be *orthogonal* to each other.

A ket whose norm is zero does not represent a physical state.

A non-normalizable ket is one that has an infinite norm. We have to be slightly careful in its interpretation.

2.1 Operators in Hilbert Space

An operator $\hat{A}$ acts on a ket $|\psi\rangle$ to givea new ket in general, $|\phi\rangle$, suppose:

$$\hat{A}|\psi\rangle = |\phi\rangle$$

We shall be exclusively concerned with linear operators. The two tests for linearity of an operator are the following:

$$\hat{A}\left(c_1|\psi\rangle + c_2|\phi\rangle\right) = \hat{A}c_1|\psi\rangle + \hat{A}c_2|\phi\rangle \tag{3}$$

and

$$\hat{A}\left(c|\psi\rangle\right) = c\hat{A}|\psi\rangle \tag{4}$$

In an operator product $\hat{A}\hat{B}$, we usually understand $\hat{B}$ to be the first operator, and it acts on whatever appears on the right:

$$\hat{A}\hat{B}|\psi\rangle = \hat{A}|\eta\rangle \tag{5}$$

$$= |\phi\rangle \tag{6}$$

where $|\eta\rangle = \hat{B}|\psi\rangle$ and $|\phi\rangle = \hat{A}|\eta\rangle$.

If $\hat{A}|\psi\rangle = \hat{B}|\psi\rangle$ for any arbitrary ket $|\psi\rangle$, then we say that $\hat{A} = \hat{B}$.

Two operators $\hat{A}$ and $\hat{B}$ are said to *commute* if $\hat{A}\hat{B} = \hat{B}\hat{A}$.

We define:

$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ to be the *commutator* of $\hat{A}$ and $\hat{B}$,

and

$\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$ to be the *anticommutator* of $\hat{A}$ and $\hat{B}$.

Just as every ket has a *dual*, a bra, every operator $\hat{A}$ has an *adjoint* operator, which is constructed thus: the dual of $\hat{A}|\phi\rangle$ is a new bra which is antilinear in $|\phi\rangle$ and so is like a bra formed by operating on $\langle\phi|$ by another operator from the right denoted by $\hat{A}^\dagger$. Thus we get a defining equation for the adjoint of $\hat{A}$:

$$\overline{\hat{A}|\phi\rangle} = \langle\phi|\hat{A}^\dagger$$

Then

$$\overline{\langle\psi|\hat{A}|\phi\rangle} = \langle\phi|\hat{A}^\dagger|\psi\rangle$$

It may sometimes so happen that $\hat{A} = \hat{A}^\dagger$. Then we say that $\hat{A}$ (or $\hat{A}^\dagger$ for that matter) is a *self-adjoint* or Hermitian operator (named after the French mathematician Charles Hermite). Such operators, as we will come to see, have a very important role to play in quantum mechanics.

If $\hat{U}$ is an operator, such that $\hat{U}\hat{U}^\dagger = \hat{I}$, where $\hat{I}$ is the identity operator which operates on a ket (or bra) to give the same ket (or bra, respectively), we say that $\hat{U}$ is a *unitary* operator. This, too, has a vital role to play in quantum mechanics.

If $\hat{A}^\dagger = -\hat{A}$ we say that $\hat{A}$ (or equivalently $\hat{A}^\dagger$) is an *anti-Hermitian* operator.

In quantum mechanics, we do not often encounter anti-Hermitian operators. More commonly, one encounters Hermitian operators. There are, of course, operators, which are neither Hermitian nor anti-Hermitian: the *ladder* operators, better known as the *creation* and the *annihilation* operators. These happen to be important in the theory of the harmonic oscillator, and are of central importance in quantum field theory.

2.2 Eigenvalue Problem

If we can find a ket so that when the operator $\hat{A}$ acts on it, we get the same ket with a multiplicative number, then this ket is said to be an eigenket at $\hat{A}$ and the number is said to be an eigenvalue:

$$\hat{A}|a\rangle = a|a\rangle$$

Here $|a\rangle$ is the eigenket, and a is the number that is the eigenvalue. There may be several eigenkets with a single eigenvalue. Then we say that we have a *degenerate* eigenvalue. If several eigenkets $|a\rangle$ with different values of a exist, we say we have a spectrum of eigenvalues.

2.2.1 Properties of Eigenvalues and Eigenvectors

1. *Eigenvalues of Hermitian operators are real numbers.*

Let $|a\rangle$, $|a'\rangle$ be two distinct eigenkets of $\hat{A}$: then

$$\begin{aligned} \langle a'|\hat{A}|a\rangle &= \langle a'|a|a\rangle \\ &= a\langle a'|a\rangle \end{aligned} \tag{7}$$

so, taking the complex conjugate of this, we have:

$$\langle a|\hat{A}^\dagger|a'\rangle = a^*\langle a|a'\rangle \tag{8}$$

or, because $\hat{A}$ is Hermitian,

$$\langle a|\hat{A}|a'\rangle = a^*\langle a|a'\rangle \tag{9}$$

Similarly, by exchanging a and a', we can show that

$$\langle a'|\hat{A}|a\rangle = a'^*\langle a'|a\rangle \tag{10}$$

Thus

$$(a - a'^*)\langle a'|a\rangle = 0 \tag{11}$$

For $a' = a$, therefore,

$$(a - a^*)\langle a|a\rangle = 0 \tag{12}$$

If the state $|a\rangle$ exists, then its norm is non-vanishing, so we must have

$$a = a^* \tag{13}$$

that is, the eigenvalues of a Hermitian operator are real.

2. *Eigenkets belonging to different eigenvalues are othogonal to each other.*

From

$$(a - a'^*)\langle a'|a\rangle = 0 \tag{14}$$

if a is different from a', since the eigenvalues here are real, firstly,

$$(a - a')\langle a'|a\rangle = 0 \tag{15}$$

and therefore

$$\langle a'|a\rangle = 0 \tag{16}$$

That is eigenkets belonging to different eigenvalues of a Hermitian operator are orthogonal to each other.

Often, we normalize kets so that $\langle a|a\rangle = 1$, so that

$$\langle a'|a\rangle = \delta_{a'a} \tag{17}$$

3. We now make an assumption which holds good only for non-degenerate eigenkets. We assume that eigenkets of a Hermitian operator of significance form a *complete set*.

   Definition of a *complete set*: If, for an arbitrary state $|\psi\rangle$, we can express it as a linear combination of a set of eigenkets $|a\rangle$ of some Hermitian operator $\hat{A}$, we say that these eigenkets form a complete set. Then

$$|\psi\rangle = \sum_a c_a |a\rangle \tag{18}$$

   Under these circumstances we say that c_a are the components along all eigenkets $|a\rangle$ and the complete set forms a basis. To obtain the coefficients c_a, multiply $|\psi\rangle$ by the eigenbra $\langle a'|$.

$$\begin{aligned} \langle a'|\psi\rangle &= \langle a'| \sum_a c_a |a\rangle \\ &= \sum_a c_a \delta_{a'a} \\ &= c_{a'} \end{aligned} \tag{19}$$

   so

$$c_a = \langle a|\psi\rangle \tag{20}$$

   Thus c_a represents the probabiltiy amplitude of a system in a state $|\psi\rangle$ jumping to a base state $|a\rangle$.

4. *Closure Relation*

   Since the base states form a complete set, we should expect that

$$\sum_a |c_a|^2 = 1 \tag{21}$$

   assuming that

$$\langle \psi|\psi\rangle = 1 \tag{22}$$

   Now we actually prove this: The last equation above implies that

$$\sum_a c_a^* \langle a|c_a|a\rangle = 1 \tag{23}$$

   or

$$\sum_a c_a c_a^* \langle a|a\rangle = 1 \tag{24}$$

   or

$$\sum_a |c_a|^2 = 1 \tag{25}$$

   as claimed. This is known as a *closure relation*.

5. *Completeness condition*

We now represent this relation in a different way:

$$\begin{aligned} |\psi\rangle &= \sum_a c_a |a\rangle \\ &= \sum_a |a\rangle c_a \\ &= \sum_a |a\rangle\langle a|\psi\rangle \end{aligned} \tag{26}$$

Thus, for arbitrary $|\psi\rangle$, we may identify

$$\sum_a |a\rangle\langle a| = \hat{I} \tag{27}$$

where $\hat{I}$ is the identity operator. This is known as the *completeness condition.*

Now we return to our physical lab space. The Hermitian operators have a one-to-one correspondence with dynamical variables like energies, momentum, position, etc. They are then called, for instance, Energy operator, momentum operator, etc.

When we prepare a system in a certain state, then if we pass it through a measuring device, we shall measure only eigenvalues of the observable we are measuring, that is, our device will show only eigenvalues.

If we prepare a system in one eigenstate, then we shall always measure the eigenvalue corresponding to this state.

If we prepare it in any intermediate state, then we shall observe any one of the eigenvalues with definite probabilities for each, the probability amplitudes are equal to c_a.

3 Representations

In the A-basis we have three relations:

1. Orthonormality: $\langle a|a'\rangle = \delta_{aa'}$
2. Closure: $\sum_a |a\rangle\langle a| = \hat{I}$
3. Eigenvalue problem: $\hat{A}|a\rangle = a|a\rangle$

Let us consider the case when the eigenvalues $a, a', \ldots$ are all discrete. We take an arbitrary state $|\psi\rangle$ and represent it in the A-basis.

$$|\psi\rangle = \hat{I}|\psi\rangle = \sum_a |a\rangle\langle a|\psi\rangle \tag{28}$$

Let us define $\psi_a \equiv \langle a|\psi\rangle$. Then

$$|\psi\rangle = \sum_a |a\rangle\psi_a \tag{29}$$

In the following discussion we shall consider only two-state systems. Examples are abundant:

1. The ammonia maser
2. H_2^+ ion: This is a system consisting of two protons and one electron. It can exist either as one or the other of the following two states: the electron predominantly in the neighbourhood of proton 1 or proton 2.
3. The benzene molecule

In (29), $\sum_a \psi_a|a\rangle$ is represented by a single column matrix called the *Representative* or loosely the *Wavefunction* of the state $|\psi\rangle$ in the A-basis, denoted by $[\psi]_A$. So

$$|\psi\rangle = \begin{pmatrix} \psi_a \\ \psi_{a'} \end{pmatrix} \tag{30}$$

As usual, the conjugate of this relation gives us

$$\langle\psi| = \begin{pmatrix} \psi_a^* & \psi_{a'}^* \end{pmatrix} \tag{31}$$

So

$$\langle\psi|\psi\rangle = \begin{pmatrix} \psi_a^* & \psi_{a'}^* \end{pmatrix} \begin{pmatrix} \psi_a \\ \psi_{a'} \end{pmatrix} = |\psi_a|^2 + |\psi_{a'}|^2 \tag{32}$$

Since $|\psi\rangle$ is normalised, we get

$$|\psi_a|^2 + |\psi_{a'}|^2 = 1 \tag{33}$$

Similarly the scalar product $\langle\phi|\psi\rangle$ would be represented as

$$\langle\phi|\psi\rangle = \begin{pmatrix} \phi_a^* & \phi_{a'}^* \end{pmatrix} \begin{pmatrix} \psi_a \\ \psi_{a'} \end{pmatrix} = \phi_a^*\psi_a + \phi_{a'}^*\psi_{a'} \tag{34}$$

3.1 Representatives of Operators

By definition, an operator $\hat{F}$ operates on a ket $|\psi\rangle$ to give a new ket $|\phi\rangle$. So

$$|\phi\rangle = \hat{F}|\psi\rangle \tag{35}$$

To get numbers we multiply by $\langle a|$ on both sides:

$$\langle a|\phi\rangle = \langle a|\hat{F}|\psi\rangle \tag{36}$$

$$\Rightarrow \langle a|\phi\rangle = \langle a|\hat{F}\hat{I}|\psi\rangle = \sum_{a'}\langle a|\hat{F}|a'\rangle\langle a'|\psi\rangle \tag{37}$$

This is easily represented in the form of matrices:

$$\begin{pmatrix} \phi_a \\ \phi_{a'} \end{pmatrix} = \begin{pmatrix} F_{aa} & F_{aa'} \\ F_{a'a} & F_{a'a'} \end{pmatrix}\begin{pmatrix} \psi_a \\ \psi_{a'} \end{pmatrix} \tag{38}$$

Again, identifying $|\phi\rangle$ and $|\psi\rangle$ with column matrices, we see that operating by an operator $\hat{F}$ amounts to multiplying the column matrix by a square matrix to get a new column matrix. The above relation is also denoted as

$$[\phi]_A = [\hat{F}]_A[\psi]_A \tag{39}$$

Again, the adjoint of $\hat{F}$, namely $\hat{F}^\dagger$ is the same square matrix $[\hat{F}]_A$ with rows and columns interchanged, and the complex conjugates of the matrix elements taken:

$$[F^\dagger]_A = \begin{pmatrix} F^*_{aa} & F^*_{a'a} \\ F^*_{aa'} & F^*_{a'a'} \end{pmatrix} \tag{40}$$

It can be shown from the definition of operator products that

$$[\hat{F}\hat{G}]_A = [\hat{F}]_A[\hat{G}]_A \tag{41}$$

Now that we know the representation of $\hat{F}$ in the A-basis, let us find out what is the representation of $\hat{A}$ itself in its own basis.

$$\begin{aligned} [\hat{A}]_A &= \begin{pmatrix} \langle a|\hat{A}|a\rangle & \langle a|\hat{A}|a'\rangle \\ \langle a'|\hat{A}|a\rangle & \langle a'|\hat{A}|a'\rangle \end{pmatrix} \\ &= \begin{pmatrix} a\langle a|a\rangle & a'\langle a|a'\rangle \\ a\langle a'|a\rangle & a'\langle a'|a'\rangle \end{pmatrix} \\ &= \begin{pmatrix} a & 0 \\ 0 & a' \end{pmatrix} \end{aligned} \tag{42}$$

Thus the representation of a dynamical variable in its own basis is a diagonal matrix with the eigenvalues as the diagonal elements. If there were degeneracy present, we would

have had eigenvalues repeated along the diagonal. But $[\hat{A}]_B$ is not a diagonal matrix. So to find the eigenvalues of an operator $\hat{F}$ we diagonalise the matrix.

So long we have confined our attention to a two-dimensional discrete Hilbert space. But if the Hilbert space is not discrete and is infinite dimensional, we proceed in the following way. We start off with a big Hilbert space:

$$|\psi\rangle = \begin{pmatrix} \psi_a \\ \psi_{a'} \\ \psi_{a''} \\ \vdots \end{pmatrix} \tag{43}$$

Since these components are all numbers (assumed real, for the sake of convenience) we can plot a histogram, with a labelling the horizontal axis. But when the Hilbert space becomes infinite dimensional, we have, firstly, an infinite number of bars in the histogram. Also, in the limit, the bars shrink in width, and the histogram is replaced by a function. Now we can no longer talk about $\psi_a, \psi_{a'}, \ldots$ since $a, a', \ldots$ are closely spaced. So as our continuous basis if we choose x where $-\infty < x < \infty$, we have $\psi(x)$. This is the *wave function.*

An example of a continuous basis may be a bead sliding on a string. The ammonia maser is *not* an example of a continuous coordinate basis as there are only two eigenvalues of x, nothing in between.

The three basic relations that we derived for the discrete basis have their counterparts in the continuous basis as:

1. Orthonormality: $\langle x|x'\rangle = \delta(x - x')$
2. Closure: $\int_{-\infty}^{\infty} |x\rangle\langle x| dx = \hat{I}$
3. Eigenvalue equation $\hat{x}|x\rangle = x|x\rangle$

4 Angular Momentum

We now focus our attention on problems in the three dimensional lab space. For this we introduce the concept of *angular momentum*. As a starting point in its definition, we turn to the classical definition first. Classically it is given by the so-called cross-product of position $\vec{r}$ and momentum $\vec{p}$. Let us see what we get if we define, by analogy:

$$\hat{L} = \hat{x} \times \hat{p} \tag{44}$$

Let us first evaluate the various commutator $[\hat{L}_x, \hat{L}_y]$. For this we shall need the fundamental commutation relation $[\hat{x}, \hat{p}_x] = i\hbar$.

Then,

$$\begin{aligned}
[\hat{L}_x, \hat{L}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x - \hat{x}\hat{p}_z] \\
&= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] - [\hat{y}\hat{p}_z, \hat{x}\hat{p}_z] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] \\
&= [\hat{y}\hat{p}_z, \hat{z}]\hat{p}_x + \hat{z}[\hat{y}\hat{p}_z, \hat{p}_x] - 0 - 0 + [\hat{z}\hat{p}_y, \hat{x}]\hat{p}_z + \hat{x}[\hat{z}\hat{p}_y, \hat{p}_z] \\
&= [\hat{y}, \hat{z}]\hat{p}_z\hat{p}_x + \hat{y}[\hat{p}_z, z]\hat{p}_x + 0 + 0 + \hat{x}[\hat{z}, \hat{p}_z]\hat{p}_y + \hat{x}\hat{z}[\hat{p}_y, \hat{p}_z] \\
&= 0 - i\hbar\hat{y}\hat{p}_x + i\hbar\hat{x}\hat{p}_y + 0 \\
&= i\hbar(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x)
\end{aligned} \tag{45}$$

In short,

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z \tag{46}$$

Similarly we have $[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}x$ and $[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y$.

There is one difficulty with this method of defining angular momentum. The spin, for instance, cannot be described this way. However, the beauty is that there also the commutation relations remain exactly the same. It is only for *Orbital Angular Momentum* that the former definition applies.

Therefore, as our our general definition of *Angular Momentum* we take the commutation relations themselves: the angular momentum operator is defined as an operator $\hat{\mathbf{J}}$ which obeys the commutation relations:

$$\begin{aligned}
[\hat{J}_x, \hat{J}_y] &= i\hbar\hat{J}_z \\
[\hat{J}_y, \hat{J}_z] &= i\hbar\hat{J}_x \\
[\hat{J}_z, \hat{J}_x] &= i\hbar\hat{J}_y
\end{aligned} \tag{47}$$

and, furthermore, $\mathbf{J}$ is an observable.

We define another operator $\hat{\mathbf{J}}^2$ as

$$\hat{\mathbf{J}}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \tag{48}$$

Let us find out the properties of this operator, that is, its commutation relations. We shall now drop the hats from now on, and the meanings will be clear apparent from their

usage.

$$\begin{aligned}[J^2, J_x] &= [J_x^2, J_x] + [J_y^2, J_x] + [J_z^2, J_x] \\ &= 0 + [J_y, J_x]J_y + J_y[J_y, J_x] + [J_z, J_x]J_z + J_z[J_y, J_x] \\ &= 0 - i\hbar J_z J_y - i\hbar J_y J_z + i\hbar J_y J_z + i\hbar J_z J_y \\ &= 0\end{aligned} \tag{49}$$

Thus, similarly, $[J^2, J_y] = [J^2, J_z] = 0$ so the eigenvalues of $\mathbf{J}^2$ are in general degenerate.

Now, we make the following assumption: The simultaneous eigenkets of J^2 and J_z together form a complete set. For our purposes, J_y or J_x would have done equally well, but the choice of J_z is dictated by convention. This gives us the *Angular Momentum Basis*.

The problem now at hand is to find the eigenvalues and eigenkets of J^2 and J_z, so we have to solve the following equations:

$$J^2|\chi\rangle = \hbar^2\lambda|\chi\rangle \tag{50}$$

$$J_z|\chi\rangle = \hbar m|\chi\rangle \tag{51}$$

Here λ, m are both real (*why*?), and to make them dimensionless (*how*?) $\hbar^2, \hbar$ are inserted.

Both the eigenkets are the same since the operators commute. Since we have the same $|\chi\rangle$ belonging to two operators, we change our notation a bit. and write $|\chi\rangle = |\lambda, m\rangle$. Since λ is degerate, we shall have several eigenkets all belonging to the same eigenvalue λ. For a particular value of λ there is an upper limit and a lower limit on the value of m. Let us find the limits:

$$J^2 = J_x^2 + J_y^2 + J_z^2 \tag{52}$$

so

$$\langle\lambda, m|J^2|\lambda, m\rangle = \langle\lambda, m|J_x^2|\lambda, m\rangle + \langle\lambda, m|J_y^2|\lambda, m\rangle + \langle\lambda, m|J_z^2|\lambda, m\rangle \tag{53}$$

Since, J_x is self-adjoint, J_x^2 is a positive operator, and its expectation value must be non-negative, and so for J_y and J_y^2.

$$\lambda\hbar^2 = \text{nonnegative quantity} + \text{nonnegative quantity} + m^2\hbar^2$$

So

$$\lambda \geqslant m^2 \tag{54}$$

Thus the upper and lower limits of m are fixed at let us say, j and k, and we label the corresponding λ by λ_j. That would mean we will have to write our eigenkets as $|\lambda_j, m\rangle$, which would be very clumsy unless some clever guy got rid of the λ altogether to give $|j, m\rangle$. Now we define two new operators as follows:

$$J_+ = J_x + iJ_y \tag{55}$$

$$J_- = J_x - iJ_y \tag{56}$$

Note that $J_+^\dagger = J_-$. A few lines of algebraic manipulations are sufficient to show the following results:

$$[J^2, J_+] = 0 \tag{57}$$

$$[J^2, J_-] = 0 \tag{58}$$

and

$$[J_z, J_\pm] = \pm\hbar J_\pm \tag{59}$$

Now let us see what happens if we operate $J^2 J_+$ on $|j, m\rangle$:

$$J^2 J_+|j, m\rangle = J_+ J^2|j, m\rangle = J_+(\hbar^2\lambda_j|j, m\rangle) = \hbar^2\lambda_j(J_+|j, m\rangle) \tag{60}$$

Thus $J_+|j, m\rangle$ is an eigenket of J^2 belonging to the eigenvalue $\hbar^2\lambda_j$, so operating J_+ on an eigenket $|j, m\rangle$ gives us a new eigenket belonging to the same eigenvalue. Let us find out what the new ket $J_+|j, m\rangle$ is in terms of $|j, m\rangle$, by first acting on it by J_z:

$$\begin{aligned} J_z J_+|j, m\rangle &= ([J_z, J_+] + J_+ J_z)\,|j, m\rangle \\ &= \hbar J_+|j, m\rangle + J_+ m\hbar|j, m\rangle \end{aligned} \tag{61}$$

or

$$J_z\,(J_+|j, m\rangle) = \hbar(m + 1)\,(J_+|j, m\rangle) \tag{62}$$

so again we find that J_+ operating on the simultaneous eigenket of J_z and J^2, namely, $|j, m\rangle$ gives another eigenket belonging to a different eigenvalue, namely, $\hbar(m + 1)$. So effectively J_+ *raises* $|j, m\rangle$ belonging to an eigenvalue m to a new eigenket $|j, m+1\rangle$ with eigenvalue $m+1$. Thus J_+ is called a **Raising Operator**, and similarly, J_- a **Lowering Operator**. But it must be remembered that this movement up and down the "ladder" is restricted to the single constant eigenvalue λ_j.

So

$$J_+|j, m\rangle = c|j, m + 1\rangle \tag{63}$$

$$J_-|j, m\rangle = b|j, m - 1\rangle \tag{64}$$

where c, b are factors to be determined. We are now in a position to find the limits of the ladder, namely, j and k.

Since J_+ takes $|j, m\rangle$ upward by one step, applying it at the top of the ladder gives us a null ket:

$$J_+|j, j\rangle = 0 \tag{65}$$

Operating J_- on a null ket also gives a null ket, so

$$J_- J_+|j, j\rangle = 0 \tag{66}$$

Rewriting the operator $J_- J_+$ in terms of J^2 and J_z, we get:

$$\begin{aligned} J_- J_+ &= (J_x - iJ_y)(J_x + iJ_y) && (67)\\ &= J_x^2 + J_y^2 + i[J_x, J_y] && (68)\\ &= J^2 - J_z^2 - \hbar J_z && (69) \end{aligned}$$

Thus

$$J_- J_+ |j, j\rangle = 0 \tag{70}$$

implies that

$$J^2|j, j\rangle - J_z^2|j, j\rangle - \hbar J_z|j, j\rangle = 0 \tag{71}$$

or

$$(\lambda_j \hbar^2 - j^2\hbar^2 - j\hbar^2)|j, j\rangle = 0 \tag{72}$$

Since $|j, j\rangle \neq 0$, we must have $\lambda_j = j(j + 1)$

Similarly repeating this procedure for the lower bound leads eventually to the discovery (some steps are omitted here for now, will be filled later) that the only viable solution is $k = -j$.

Thus we have found out that given a particular value of λ_j, it has $2j+1$ values running from $+j$ to $-j$ where the value of j may be obtained from solving $j(j + 1) = \lambda_j$.

Let us find out the values of j in a few special cases:

1. *Single Step Ladder:* $j = 0, k = -j = 0$: This corresponds to $\lambda_j = 1$. Here $|0, 0\rangle$ is both at the top and the bottom of the ladder.

2. *Two-step ladder:* Here the states must be restricted to $|j, j\rangle$ and $|j, -j\rangle$. Now J_- acting on the first of these must give the latter, so $j - 1 = -j$, solving which gives $j = \frac{1}{2}$, and $\lambda_j = \frac{3}{4}$.

3. **Note:** We *cannot* have $j = \frac{1}{3}$, since repeatedly operating with J_- on $|\frac{1}{3}, \frac{1}{3}\rangle$ will give $|\frac{1}{3}, -\frac{2}{3}\rangle$, then $|\frac{1}{3}, -1\rangle$, then $|\frac{1}{3}, -\frac{4}{3}\rangle$, and so on, but *never* $|\frac{1}{3}, -\frac{1}{3}\rangle$ which is needed to terminate the *ladder*.

4. $(n + 1)$*-step ladder*: After n applications of J_- we reach $|j, j - n\rangle$, which is $|j, -j\rangle$, so $j - n = -j$, which in turn implies that $j = \frac{n}{2}$, where $n = 0, 1, 2, 3, \ldots$. Thus j can have values $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$.

4.1 Matrix Representation of Angular Momentum

Sometimes it is easier to work with matrix representations of operators. For this purpose we will need to know the representations of J^2, J_x, J_y, J_z in the *angular momentum basis*. For that we shall need to know the values of b and c in the equations

$$J_+|j, m\rangle = \hbar\, c|j, m + 1\rangle \tag{73}$$

and

$$J_-|j, m\rangle = \hbar\, b|j, m - 1\rangle \tag{74}$$

Taking the conjugate of the first of these we get

$$\langle j, m|J_- = \hbar\, c\langle j, m + 1| \tag{75}$$

Thus

$$\langle j,m|J_-J_+|j,m\rangle = \hbar^2\ c^2\langle j,m+1|j,m+1\rangle \tag{76}$$

or

$$\langle j,m|J^2 - J_z^2 - \hbar J_z|j,m\rangle = \hbar^2 c^2 \tag{77}$$

or

$$\lambda_j - m^2 - m = c^2 \tag{78}$$

or

$$c = \sqrt{j(j+1) - m(m+1)} \tag{79}$$

or equivalently,

$$c = \sqrt{(j-m)(j+m+1)} \tag{80}$$

Similarly,

$$b = \sqrt{(j+m)(j-m+1)} \tag{81}$$

4.2 Orbital Angular Momentum

Next we turn to a special case of angular momentum, namely, *orbital angular momentum.*

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}} \tag{82}$$

Since $\hat{L}$ involves $\hat{x}, \hat{y}, \hat{z}$ it has a coordinate representation. However, in general, angular momentum in quantum mechanics does not necessarily always allow a coordinate representation.

For orbital angular momentum, instead of expressing the operators in x, y, z we write them in r, θ, ϕ.

Let us first recall the transformation rules:

$$x = r\sin\theta\sin\phi \tag{83}$$

$$y = r\sin\theta\cos\phi \tag{84}$$

$$z = r\cos\theta \tag{85}$$

$$\frac{\partial}{\partial r} = \frac{\partial}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial}{\partial z}\frac{\partial z}{\partial r}$$

$\therefore$

$$\frac{\partial}{\partial r} = \sin\theta\cos\phi\frac{\partial}{\partial x} + \sin\theta\sin\phi\frac{\partial}{\partial y} + \cos\theta\frac{\partial}{\partial z} \tag{86}$$

Similarly

$$\frac{\partial}{\partial \theta} = r\cos\theta\cos\phi\frac{\partial}{\partial x} + r\cos\theta\sin\phi\frac{\partial}{\partial y} - r\sin\theta\frac{\partial}{\partial z} \tag{87}$$

and

$$\frac{\partial}{\partial \phi} = -r\sin\theta\sin\phi\frac{\partial}{\partial x} + r\cos\theta\cos\phi\frac{\partial}{\partial y} \tag{88}$$

Thus we can write